Handbook of Discourse Processes
This page intentionally left blank
Handbook of Discourse Processes

Edited by

Arthur C. Graesser
University of Memphis

Morton Ann Gernsbacher
University of Wisconsin, Madison

Susan R. Goldman
University of Illinois, Chicago
Contents

Preface vii

1 Introduction to the Handbook of Discourse Processes
   Arthur C. Graesser, Morton A. Gernsbacher,
   and Susan R. Goldman 1

2 Genres, Registers, and Contexts of Discourse
   Allen D. Grimshaw 25

3 Text Comprehension
   Rolf A. Zwaan and Murray Singer 83

4 Processes of Interactive Spoken Discourse: The Role
   of the Partner
   Michael F. Schober and Susan E. Brennan 123

5 Classroom Discourse
   Courtney B. Cazden and Sarah W. Beck 165

6 Learning From Traditional and Alternative Texts:
   New Conceptualizations for the Information Age
   Patricia A. Alexander and Tamara L. Jetton 199

7 Theories and Methods in Mediated Communication
   Steve Whittaker 243

8 Narrative Discourse
   David Bloome 287
9  Literary Discourse
   David S. Miall  321
10  Nonliteral Speech Acts in Text and Discourse
    Raymond W. Gibbs, Jr.  357
11  Discourse Development
    Michael Bamberg and Luke Moissinac  395
12  Discourse in Computational Linguistics and Artificial Intelligence
    Johanna D. Moore and Peter Wiemer-Hastings  439
13  Quantitative Cognitive Models of Text and Discourse Processing
    Peter W. Foltz  487

   Author Index  525
   Subject Index  543
The idea of editing *Handbook of Discourse Processes* originated in the summer of 1998 in Madison, Wisconsin. Morton Gernsbacher was busy hosting a plethora of conferences in Madison in a 10-day time span in July. Madison was the Mecca that year for thousands of researchers to attend such annual meetings as the Society for Text and Discourse, the Cognitive Science Society, and the American Association of Artificial Intelligence. Morton Gernsbacher was the President of the Society for Text and Discourse, and Susan Goldman was about to become President-Elect. The idea of a *Handbook* emerged in a conversation between Judith Amsel and Art Graesser while Art was scanning the LEA books and journals, wondering what book to recommend to a graduate student who had never heard of the field of discourse processes. After lamenting over this obvious, if not embarrassing, oversight in the field, Judy Amsel (of Lawrence Erlbaum) simply asked “Why don’t you edit a *Handbook of Discourse Processes*?” And so we did. Art talked with Morton, Morton and Art talked with Susan, and by evening it was decided we would go for it. So now, 5 years later, here it is.

This is an appropriate time for the first *Handbook of Discourse Processes*. The field is experiencing its 25th anniversary, marked by the launching of the journal *Discourse Processes* in 1978. The Society for Text and Discourse is now running strong in its lucky 13th year.

We have several people to thank for assisting us in making this *Handbook* a reality. Of course, there are the chapter authors. It was
their willingness to participate in this *Handbook* project that made it a reality. We thank Judy Amsel for her brilliant suggestion, and Sara Scudder, Susan Barker, Bill Webber, and others at Lawrence Erlbaum Associates for their assistance at various stages of production. We are indebted to Shulan Lu and Amy Adcock who took on the important task of preparing the subject and author index. Deep heartfelt gratitude goes to Larry Erlbaum himself, who has always provided generous support and consistent enthusiasm for our community of researchers. We toast and dedicate this *Handbook* to Larry.

The process of editing this book was indirectly facilitated by funding from various research grants. We greatly appreciate the grants to the first editor from the Department of Defense Multidisciplinary University Research Initiative (MURI) administered by the Office of Naval Research (N00014-00-1-0600), from ONR (N00014-00-1-0917), and from the National Science Foundation (SBR 9720314, SES 9977969, REC 0106965, REC 0126265), to the second editor from National Institutes of Health (R01 NS29926 and F33 DC05365), and to the third editor from the National Science Foundation (REC 0126265), and the Spencer Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of DoD, ONR, NSF, NIH, or the Spencer Foundation.

—Art Graesser
Morton Ann Gernsbacher
Susan Goldman
Introduction to the Handbook of Discourse Processes

Arthur C. Graesser
University of Memphis

Morton A. Gernsbacher
University of Wisconsin, Madison

Susan R. Goldman
University of Illinois at Chicago

One way to organize an introduction to a handbook is to divide it into the past, present, and future. This is precisely the approach we have adopted. We start with a short history that documents how the field of discourse processes emerged. Next we describe the current trends in investigating discourse processes. We end with our forecast of how the field is destined to evolve into the future or how we hope it will evolve. The purpose of this introduction is to set the stage for the subsequent chapters rather than to provide a comprehensive overview of the field.

THE PAST: THE BIRTH OF DISCOURSE PROCESSES

The multidisciplinary field of discourse processes does not have a long history. It was officially launched in 1978 with the publication of the first issue of the journal Discourse Processes. If we accept that event and year as the christening, this year marks the 25th anniversary of the field—a fitting time to publish the first Handbook of Discourse Processes. The founder and first editor of Discourse Processes was Roy O. Freedle, a research scientist at Educational Testing Service. Freedle had the vision of the multidisciplinary field in the early 1970s. He was struck by a convergence of interest from researchers in several fields...
who were passionately investigating connected discourse in either print or oral form. Freedle edited a book as early as 1972 with his colleague John Carroll (Carroll & Freedle, 1972) entitled Language Comprehension and the Acquisition of Knowledge. Freedle went on to edit the book series Advances in Discourse Processes, which had its first edited volume in 1977 and had dozens of volumes up through the late 1990s.

Researchers became interested in discourse when they became dissatisfied with the sentence or utterance as the unit of analysis in their investigations of language. It is quite correct that printed texts consist of a sequence of sentences, and that oral conversations consist, more or less, of a sequence of spoken utterances. However, discourse cannot be entirely reduced to sentences and utterances. Discourse has a context, cohesion, coherence, and rhetorical structure that weaves together and transcends the sentences/utterances. The context includes the participants in the acts of communication (speaker, listener, overhearer, reader, writer, narrator, narratee), the communication setting and goals, the pragmatic ground rules, the subject matter knowledge, and a host of other components that situate the discourse event (see chapter by Grimshaw). Most of this context is invisible in the sense that there are few features, elements, or constituents in the text to point to and designate “that’s the context.” Cohesion refers to the linguistic elements or features that signal how to connect constituents, such as connectives (and, in order to, although), punctuation, verb tense, and syntactic markers of given-new contrasts. Coherence refers to the alinguistic conceptual knowledge that interrelates constituents in the discourse, such as time, space, causality, goals, and agency (see chapter by Zwaan and Singer). The rhetorical structure specifies the global organization of discourse, such as setting–plot–moral, problem–solution, compare–contrast, claim–evidence, question–answer, and argue–counterargue. These levels all impose meaning and structure on individual sentences (or utterances) that go well beyond the compositional meaning of sentences in isolation. It is perfectly obvious, for example, that the meaning of a text is entirely different when we reorder the sentences. We also know that a sentence in isolation is frequently ambiguous, whereas a sentence in a naturalistic discourse context is rarely ambiguous.

Researchers became interested in discourse when they discovered some systematic discourse patterns, empirical findings, and processing mechanisms. These discoveries emerged rather quickly from several different fields in the 1970s just before the field was launched in 1978. Consider the following landmark contributions during 6 golden years of discovery between 1972 and 1978.
Text Linguistics

The structural grammars in linguistic theories of sentence syntax could be successfully applied to more global stretches of connected discourse (van Dijk, 1972). Principles of cohesion and coherence were identified in different types of discourse (Grimes, 1975; Halliday & Hasan, 1976).

Psychology

The propositional density of explicit text could predict reading times for sentences, whereas recall of text propositions was influenced by their level in a hierarchical structure (Frederiksen, 1975; Kintsch, 1974). The distinction between given and new information in a text predicted reading times for sentences (Haviland & Clark, 1974). Comprehension and memory for a vague text dramatically improved when there was a theme that clarified and integrated the sentences in the text (Bransford & Johnson, 1972). The structures generated by story grammars predicted recall and summaries of text (Mandler & Johnson, 1977; Rumelhart, 1975). Memory and inferences about texts describing frequently enacted activities (e.g., eating at a restaurant, attending a baseball game) were robustly fleshed out by the content of scripts, schemas, and other packages of generic knowledge (Bower, Black, & Turner, 1979; Spilich, Vesonder, Chiesi, & Voss, 1979).

Education

Memory for text constituents was influenced by the rhetorical organization of expository text (Meyer, 1975). Interactive models of reading were proposed that incorporated the importance of knowledge of discourse in the process of comprehension (Rumelhart & Ortony, 1977); these models contrasted with strictly bottom-up models (Gough, 1972). The impact of world knowledge on reading was vigorously explored after years of receiving minimal attention in literacy studies (Anderson, Spiro, & Montague, 1977). Comprehension was viewed by Rosenblatt (1978) as a transaction between an author and a reader through the medium of a text, as opposed to being a bottom-up extraction of language codes and meaning. Researchers working from anthropological and sociological perspectives were exploring discourse and communication in the classroom (McDermott, 1978; McDermott & Gospodinoff, 1979). For example, Sinclair and Coulthard (1975) and Mehan (1979) identified the prototypical structure of classroom interaction known as the IRE sequence: teacher Initiates, student
Responds, and teacher Evaluates. This was a predominant prototype, although detailed analyses of classroom interactions later revealed frequent improvisations and adaptive variations on the prototype in some classrooms (Erickson, 1982).

**Artificial Intelligence**

Computer models attempted to answer questions (Lehnert, 1978; Woods, 1977), generate inferences (Rieger, 1978; Schank, 1972), comprehend simple, script narratives (Schank & Abelson, 1977), and engage in mixed-initiative dialog while interacting in a toy world (Winograd, 1972). Colby, Weber, and Hilf (1971) developed a computer program called PARRY, which simulated the conversational contributions of a paranoid agent.

**Sociology and Communication**

Patterns of turn-taking in conversation were explored by Sacks, Schegloff, and Jefferson (1974). Grice (1975) identified the conversational postulates that underlie smooth conversation, including the cooperation principle and the maxims of quality, quantity, relation, and manner. Rommetveit (1974) proposed that a printed text is a structured, pragmatic, social interaction between an author and reader. Hymes (1972) analyzed the knowledge and competence that is needed for members of a speech community to participate effectively in communication events.

The theories, findings, and insights from the 6 golden years before 1978 furnished a solid foundation for the multidisciplinary field to flourish well into the future. In 1990, the Society for Text and Discourse (ST&D) was founded as a society dedicated to the field of discourse processes. *Discourse Processes* was designated as the official journal of this society. The ST&D has held annual meetings in North America and Europe throughout its 13 years of existence. The scope of ST&D embraces multiple disciplines as long as the discipline adopts a methodology that empirically tests the validity of its claims. Rigorous methods of empirical research have therefore dominated the research reported at ST&D and in the journal. These include qualitative as well as quantitative approaches to establishing evidence for claims. ST&D and *Discourse Processes* would not be the appropriate arena for literary criticism, a manifesto on the virtues of plain talk, and an interpretive analysis of psychoanalytic themes in *The Fountainhead*. The field called *discourse studies*, which emerged in the 1990s in Europe, does
not fall under the umbrella of discourse processes because rigorous scientific methods are not directly embraced as their dominant methodological and epistemological foundation.

The notion of process is a key ingredient of discourse processes. Researchers explore the processes of comprehending, producing, reproducing, composing, recalling, summarizing, and otherwise creating, accessing, and using discourse representations. So a theory of discourse processing does not simply translate a discourse excerpt into a representation (e.g., a set of propositions, clauses, hierarchical structure, or a matrix of features). The theory also needs to specify the operations of assembling, augmenting, rearranging, or disassembling the representations and contexts in which these operations occur. The processes and representations may be captured at different degrees of grain size and analytical specification. Most of the time the components of a theory are articulated verbally, but sometimes the researcher adopts a concise symbolic language (as in the chapter by Moore and Wiemer-Hastings) and sometimes a precise mathematical language (as in the chapter by Foltz).

1. INTRODUCTION

THE PRESENT: CURRENT TRENDS AND APPROACHES

The field of discourse processes is currently fueled by seven dominant approaches, which we label as (a) discourse psychology, (b) corpus analysis, (c) computational discourse, (d) discourse technologies, (e) conversation analysis, (f) hybrid qualitative and quantitative approaches, and (g) cultural foundations. Most of these approaches are hybrids of two or more disciplines, but some are confined to a single discipline. As we discuss later, our hope is that the field becomes more interdisciplinary instead of being merely multidisciplinary. Interdisciplinary research is a serious attempt to integrate research from two or more fields—a form of intellectual cross-breeding. For example, computational linguistics is a field that combines computer science and linguistics. Multidisciplinary research is a collection of disciplines that focuses on investigating a particular phenomenon; there may or may not be serious efforts for the disciplines to communicate. We ultimately argue that discourse processing research will become progressively more sophisticated to the extent that it shifts from being multidisciplinary to interdisciplinary. This shift requires developing an appreciation for multiple methods of establishing rigorous scientific claims.
Discourse Psychology

This is the most dominant approach to investigating discourse processing (Clark, 1996; Gernsbacher, 1994; Graesser, Gernsbacher, & Goldman, 1997; Graesser, Millis, & Zwaan, 1997; Kintsch, 1998). The majority of the articles in *Discourse Processes* would fall under the rubric of discourse psychology. Discourse psychology is represented in five of the chapters of this volume that cover text comprehension (Zwaan and Singer), language use (Schober and Brennan), learning from text (Alexander), nonliteral speech acts (Gibbs), and discourse development (Bamberg). Discourse psychologists test theories by collecting data from humans either during or after discourse comprehension or production. For example, the process of comprehending a text online (i.e., during comprehension) would be traced by collecting eye movements during reading. The gaze durations on individual words increase to the extent that the words involve complex cognitive processing. Regressive eye movements move backward to prior text constituents when there are comprehension difficulties or when there is a need to conceptually link a current text constituent to a prior text constituent. A more qualitative form of online data consists of think-aloud protocols, where comprehenders say whatever comes to mind as they comprehend text sentence by sentence (Pressley & Afflerbach, 1995; Trabasso & Magliano, 1996). In contrast to the online methods, offline methods tap the representation of text after comprehension is completed. Examples of off-line measures are recall protocols, summaries, decisions about test sentences on various criteria (true vs. false, old vs. new, main idea vs. minor idea), or ratings on test sentences on various dimensions. Discourse psychologists have collected dozens of measures of discourse processing in experimental designs that rigorously test theoretical claims (Haberlandt, 1994). The recruitment of factorial designs and appropriate control conditions are routinely adopted in an effort to rule out extraneous variables and determine the causes of particular discourse processes and representations. These researchers routinely conduct fine-grained statistical analyses to assess the generality of the empirical findings across humans and discourse samples.

The most recent research in discourse psychology has become progressively more interdisciplinary in several respects. One trend has been to develop quantitative models that more precisely specify predictions of theories (see chapter by Foltz). These efforts coordinate discourse psychology with mathematical models developed in the cognitive sciences. For example, Kintsch's (1998) construction-integration model incorporates neural networks, production rules, and latent se-
mantic analysis (i.e., high-dimensional semantic space) in his attempt to quantitatively pin down the mechanisms of particular processing modules. A second trend is to ground claims about discourse processing in mainstream theories in cognitive psychology; these are cognitive theories about perception, pattern recognition, memory, reasoning, and problem solving. Research in discourse psychology is now seriously criticized if it has a component that is not firmly grounded in a plausible cognitive mechanism. A third trend is to integrate discourse psychology with research in neuroscience as is discussed later (Gernsbacher & Kasckhak, in press). Thus, some researchers are recording evoked potentials and cerebral blood flow to identify the regions of the brain that participate in particular discourse processing mechanisms. A fourth trend is to measure readers on general cognitive abilities and correlate these measures with reading and discourse processes (Perfetti, 1985). This approach merges psychometrics with theoretical models in discourse psychology.

A fifth trend combines corpus analyses with data collected in psychology experiments. Discourse psychologists are sometimes criticized for writing their own discourse excerpts. They do this to exert experimental control over the stimulus and rigorously test theories by removing extraneous variables. The downside of this approach is that some of the discourse materials are unnatural or not representative of the normal discourse that humans encounter. When this happens, the researcher ends up studying how experimental participants adapt to bizarre textual materials rather than studying normal discourse processes and strategies. One way to guard against this problem is to collect or identify a naturalistic discourse corpus and assess whether the experimental discourse materials deviate from normal discourse patterns. Many of us are convinced that a corpus analysis should be routinely integrated with all approaches in discourse processes.

**Corpus Analysis**

This is the second most popular approach to investigating discourse processing (Biber, Conrad, & Reppen, 1998; Marcus, Santorini, & Marcinkiewicz, 1993). The researcher collects or identifies a corpus (sample) of naturalistic discourse excerpts that are relevant to the particular research question being investigated. The discourse corpus is analyzed by counting the frequency of discourse elements, categories, features, sequences, global patterns, or combinations of these linguistic/discourse entities. The frequencies can be normalized by counting the number of occurrences per 1,000 words or per time duration (which we call an *incidence score*). For example, one could count how
many metaphors occur per 1,000 words in a corpus of short stories. Researchers compare corpora on the incidence scores of various discourse features. For example, Graesser and Person (1994) reported that students ask .17 question per hour in a classroom on research methods, whereas the rate is 26.5 questions per hour in tutoring sessions on research methods.

One of the principle challenges in this line of research lies in selecting the discourse corpus. Some corpora are general as in the case of the Brown corpus (Francis & Kucera, 1982), the Santa Barbara corpus of spoken American English (DuBois, 2000), the Switchboard corpus of telephone conversations between strangers (Godfrey, Holliman, & McDaniel, 1992), or the new TalkBank corpus currently being developed (MacWhinney, 2001). The Brown corpus is a 1 million word collection selected from 500 written texts from different discourse genres (e.g., newspapers, novels, nonfiction). The Switchboard corpus has 2,430 conversations for 240 hours of speech and 3 million words. MacWhinney’s new TalkBank corpus (http://talkbank.org/) will have a diversity of registers of oral conversations (e.g., classrooms, tutoring, family discussions). These corpora are available electronically so it is easy for researchers to access and search through the corpus to compute incidence scores. Electronic availability also makes it possible for researchers to re-analyze the data from different perspectives or apply different criteria for segmenting the discourse. Some corpora are more specific as in the case of classroom discourse (see chapter by Cazden), tutoring (Graesser & Person, 1994), doctor–patient interactions, or arguments between spouses. Some theoretical claims are relevant to a restricted class of genres, registers, or contexts of discourse (see chapter by Grimshaw) so it is important to have the appropriate classes of texts represented in the corpus.

Researchers frequently collect their own corpus of discourse samples. This is quite appropriate when theoretical claims can be tested only by identifying a new corpus that satisfies various constraints of selection (designated as discourse class C). For example, one might want a corpus of tutorial dialogues that includes accomplished tutors of physics who tutor underachieving inner-city children. No available corpus satisfies this set of constraints, so a new corpus needs to be selected.

According to the hard-core scientific researchers, the discourse corpus needs to be sampled systematically and scientifically, rather than haphazardly or with bias. Three mistakes are frequently made when a corpus is collected to make claims about discourse class C. The first is that the researcher hand picks the discourse excerpts and uses them as examples to illustrate whatever theoretical point is being made.
This method suffers from a selection bias so absolutely no generalizations can be made about discourse class C. Indeed, nothing at all can be concluded because the examples might be unusual rather than representative of the full population of discourse samples in class C. The second mistake is that the method of selection is unspecified. That is, the researcher does not bother to identify discourse class C and how excerpts from class C were sampled. As a consequence, the scope of the generalizations is entirely indeterminate so researchers never know whether the researcher’s conclusions are relevant to another discourse corpus. Ideally, the science of selecting a discourse corpus should be on par with the science of selecting participants in public opinion polls. The third mistake is that the sample size is typically too small when a corpus is assembled. The reliability of an empirical test is dramatically different for a sample size of 10 versus 100 versus 1,000.

It is important to acknowledge that some discourse processing researchers do not insist on the hard-core scientific approach to sampling discourse excerpts. According to an alternative approach (sometimes called the qualitative approach), much can be learned about discourse from a detailed analysis of a single case, a small representative sample of cases, or a small nonrepresentative sample of cases (Mitchell, 1984). A single case might be particularly illuminating when it has patterns of discourse that challenge existing discourse theories or unveil the intricacies of complex discourse mechanisms. In presenting such illustrative cases, researchers need to provide the warrants for claims that these cases are truly illuminating from the standpoint of existing theory and research. They must also resist the temptation to overgeneralize and assume their conclusions apply to other discourse excerpts, genres, and registers.

Computational Discourse

Just as the field of computational linguistics combines computer science and linguistics, the computational discourse approach combines discourse processes and computer science (Allen, 1995; Jurafsky & Martin, 2000; chapter by Moore and Wiemer-Hastings). The researcher needs a sufficiently detailed understanding of a discourse processing module so that the researcher can program the computer to implement the mechanism. Some discourse modules are comparatively easy to implement. A lexicon is a list of words or morphemes, with each entry having a list of linguistic features (phonological, syntactic, semantic), semantic word senses, familiarity metrics, and so on (as in the case of Miller et al., 1990). It is also
straightforward to implement syntactic parsers, which automatically construct syntactic tree structures to sentences. Semantic analyses are difficult and not as reliable, however. For example, the assignment of noun phrases to case structure roles (e.g., agent, recipient, object, location) is well below 80% correct in the available computer systems (DARPA, 1995). There still is no computer system that can automatically assign text to proposition units with an acceptable degree of reliability. Automatic construction of most discourse representations is still out of reach. The binding of pronouns or noun phrases to prior discourse constituents (i.e., the anaphoric referent) hovers around 50% in most systems tested on naturalistic texts. There is no available system that automatically tracks the goals, beliefs, and common ground of speech participants or that automatically constructs the rhetorical composition of a text.

Nevertheless, there have been some notable successes in computational discourse. Most of these successes have used statistical models of discourse and world knowledge that induce discourse patterns from a large discourse corpus. The most successful statistical models are Bayesian Markov models, neural networks, and latent semantic analysis (Jurafsky & Martin, 2000; Landauer, Foltz, & Laham, 1998). Biber (1988) developed a system that impressively classifies texts into different discourse genres and registers on the basis of approximately 50 different linguistic/discourse features. The Automated Essay Grader uses the LSA technology to grade essays and is capable of assigning grades to essays as reliably as experts in English composition (Foltz, Gilliam, Kendall, 2000; see chapter by Foltz). AutoTutor is a computerized tutoring system that holds conversations in natural language with students on the topics of computer literacy and conceptual physics (Graesser, Person, Harter, & the Tutoring Research Group, 2001). AutoTutor evaluates the quality of students’ contributions (using LSA) almost as reliably as graduate student RAs; AutoTutor generates dialogue moves that are indistinguishable from human tutors in tests that have judges decide whether a dialogue move was generated by AutoTutor or a human tutor (Person, Graesser, & the Tutoring Research Group, 2002). Pennebaker and Francis (1999) developed a Linguistic Inquiry and Word Count (LIWC) system that analyzes narratives written by victims of traumatic events; the LIWC system can predict how well the person can cope with the trauma and the number of visits to a medical doctor. The Question Understanding Aid (QUAID) critiques survey questions on potential problems of question comprehension difficulty, such as unfamiliar technical terms, vague noun phrases, complex syntax, and working memory overload (Graesser, Wiemer-Hastings, Kreuz, Wiemer-Hastings, & Marques, 2000). These
success cases in computational discourse demonstrate that it is not too early to build automatic systems of discourse analysis. However, there is a large road ahead of us.

One of the challenges in computational discourse lies in evaluating how well the computer system is performing. One solution is to compare the computer output to human experts. For example, in the assessments of Foltz’s Automated Essay Grader, the grades generated by the computer had a correlation of .73 with an expert human grader, whereas a pair of human experts correlated .71. This is indeed impressive performance—a clear success case. The problems surface when human judges disagree and therefore there is no solid gold standard. All scientific analyses of discourse requires a modicum of agreement among discourse analysts regarding the (a) assignment of discourse excerpts to particular theoretical categories (e.g., speech act categories, genre categories, main vs. subordinate ideas) or (b) ratings of discourse excerpts on theoretical dimensions (e.g., hierarchical level, discourse focus, whether an author has an intention). That is why good discourse researchers report interjudge agreement scores in all of their discourse analyses. Unfortunately, some discourse categories and dimensions are so abstract, subtle, or murky that there is not a sufficient foundation for researchers to agree. Nevertheless, to be fair, this is not a unique problem for the discourse level of language analysis. Language experts do not show impressive agreement on judgments of the syntactic complexity of sentences, for example (Graesser et al., 2000).

**Discourse Technology**

We live in a world where people talk to machines as often as they talk to other humans. There is a technological challenge of how to design telecommunication systems, computer systems, devices, and other artifacts in a fashion that satisfies the constraints of discourse (see chapter by Whittaker). The history of such artifacts is replete with disasters. For example, Microsoft’s talking Paperclip was originally designed as an animated conversational agent that would help computer users interact with the software. So the Paperclip gave hints and recommendations as a helpful conversational partner. Unfortunately, there were no discourse processing researchers on the original design team, and the Paperclip ended up being irritating to most users. It butted in, uninvited, and it was virtually impossible to get rid of it. The design of animated conversational agents needs discourse researchers just as they need experts in speech synthesis, facial expressions, kinematics, computer graphics, and computer science. Similarly, auto-
mated telephone answering systems and other information appliances need the guidance of discourse processing theories. Discourse is at the heart of any sophisticated human–machine system.

Electronic technologies are revolutionizing day-to-day communication by providing new, virtual environments for interaction and learning (Goldman, 2001). Most of us are now immersed in a world of e-mail, asynchronous message posting systems (such as electronic forums and bulletin boards), and synchronous, multi-user chat rooms. These technologies are catalyzing the creation of hybrid discourse that reflects some of the features and informality of spoken discourse, but the formality of written discourse. These electronic environments, especially asynchronous ones, make thinking visible in ways that oral conversations are not because there is a printed trace of the discussion (Goldman, 1996). Indeed, emotional dimensions of messages can be made even more explicit than they might be in face-to-face conversations. For example, users frequently incorporate into their messages different icons to indicate emotional valence (😊). Icons such as these are embodied symbols of dimensions of messages that potentially are more subtle and difficult to identify in oral conversations. The print-based trace of the discourse creates opportunities for both participants and discourse analysts to reflect on the interactive construction of meaning from psychological, sociological, and cultural points of view. The existence of electronic environments for communication allows interactions to transcend traditional time and place constraints. Analyses of the discourse occurring in these environments can reveal how interactants use rhetorical devices and discourse moves to create and maintain coherence, continuity, and threads of the discourse. What remains to be seen are where the discourses of communication and learning in electronic environments converge with, versus diverge from, nonelectronic ones.

**Conversation Analysis**

Conversational Analysis (normally capitalized) is the school of thought that dissect in rich detail the timing and sequencing of language and discourse constituents in conversations (Sacks et al., 1974; Schegloff, 1997). There is a fine-grain analysis of moment-to-moment interaction and the sequences of linguistic/discourse actions that create meaning. For example, turn-taking patterns, interruptions, overlapping speech, hesitations, and pauses have frequently attracted the attention of CA researchers. CA researchers focus exclusively on naturalistic discourse. They routinely videotape the interactions to have a rich record of the
subtleties, including intonation, gesture, head nods, and facial expressions. Conversation timing and sequencing have been investigated in a multitude of discourse registers, such as dinner conversations, arguments, doctor–patient interaction, and business transactions.

Hybrid Qualitative and Quantitative Approaches

Qualitative approaches identify discourse categories, sequences, and patterns on various dimensions and grain sizes. The context of the discourse is carefully scrutinized and taken into consideration when the qualitative analyses are conducted. However, the hybrid approach does not stop with the qualitative analyses. The researcher takes additional steps by assessing the reliability between expert judges on the assignment of discourse categories/patterns to observations. The researcher collects frequencies of discourse units and patterns per unit time; these data are subsequently analyzed with statistics and other quantitative techniques routinely adopted by social scientists. The incidence of these units and patterns are then correlated with performance measures outside of the discourse arena. In essence, quantitative analyses are performed on qualitative data. The schism between qualitative and quantitative approaches to discourse processing is entirely dismissed by those pursuing the hybrid approach.

As one noteworthy example, the hybrid approach is currently popular in studies of learning and problem solving in the classroom and other learning environments. Researchers explore potential relationships between the processes manifest in the discourse and the learning outcomes of those processes. Discourse plays a dual function: It provides a window into learning processes and is an object of study. The discourse reveals how meaning is constructed individually and in social interaction (see chapter by Cazden and Berk). Discourse reveals the process of using language to further one’s thinking and deepen one’s understanding of specific content areas. For example, individuals who generate self-explanations while reading perform better on learning tasks than those who do not (Chi, de Leeuw, Chiu, & LaVancher, 1994; Coté, Goldman, & Saul, 1998; Trabasso & Magliano, 1996).

There have been substantial technological advances in the discourse and contextual data that can be captured in real time. This has accelerated the use of discourse analytic techniques to understand learning and social interaction. Prior to advances in electronic recording devices, data were limited to what an observer could check off or a researcher could write in field notes. Now there are video and audio records of interactions, which are processed by computer software in
real time. This opens the door for inventing new methodologies for understanding the construction of meaning and the processes of learning. For example, traditional Conversational Analysis was typically limited to relatively short stretches of discourse. Yet instructional interactions occur over much longer stretches of time, making it necessary to capture relationships among segments of discourse that may be separated by 30 to 40 minutes, days (in multiday instructional situations), and weeks (see Cazden & Berk chapter; Goldenberg & Patthey-Chavez, 1995; Saunders, Goldenberg, & Hamann, 1992).

**Culture and Discourse**

Cultures are defined in part by the discourse of members, so becoming a member of a culture means learning the discourse of that culture (Gee, 1996; Lave & Wenger, 1991). Failure to know the discourse of the culture presents barriers to acceptance and full participation in the culture. *Knowing the discourse* of a culture means knowing the norms that regulate participation (such as who speaks, when they speak, and how this is decided), as well as the acceptable content and form of the various discourses prevalent in the culture (see chapter by Bloome). For example, analyses of classroom discourse have shown that the language of teaching and learning is based on mainstream culture and often produces inequities in opportunities to learn for those who are outside mainstream culture (McDermott, 1978; Tharp & Gallimore, 1988). The participation structures of the classroom frequently are barriers to the participation of subgroups of individuals on the basis of ability, gender, and ethnicity. There are different styles of discourse associated with individual academic disciplines. For example, we can speak of the discourse of science and scientists, which is quite different from the discourse of history and historians (Goldman & Bisanz, 2002).

**THE FUTURE: THREE DIRECTIONS FOR GROWTH AND SURVIVAL**

Yogi Berra once warned us that we should never predict anything, especially the future. This wisdom has never stopped any academician from looking into the crystal ball and forecasting future trends. In this final section, we offer three directions that are either inevitable or profoundly needed for the field to survive. These are the integration of neuroscience with discourse research, the use of more advanced computer technologies for analyzing discourse, and a more pronounced shift from multidisciplinary to interdisciplinary research.
Integrating Neuroscience

When one of us (MAG) was in graduate school, she once sat in on an undergraduate entry-level cognition course to see how one of the department’s best instructors conveyed the excitement and exquisiteness of cognitive psychology to an audience of 19-year-olds. This professor began the first day by posing the following challenge: Imagine that you were sent to some far away planet, and your mission on this planet was to discern the workings of a mysterious structure (at least to you)—what here on Earth we might call a building. However, you were prohibited or otherwise unable to enter the structure. How would you discern what work went on inside the structure? With guidance, the students arrived at recommendations such as to first carefully observe what entered the structure and then carefully observe what exited the structure, and from those observations to infer what work must go on inside that structure. Thus, if one saw sheet metal, rubber, and glass entering the structure and one observed Subaru’s leaving the structure, one might infer something different than if one saw denim cloth, thread, and zippers entering the structure and men’s blue jeans coming out.

At this point, the professor skillfully introduced some of the rudimentary concepts of experimental design—how, if clever enough, one might manipulate certain aspects of the input to the structure while controlling as many extraneous variables as possible, then measure qualitatively or quantitatively the output and thereby allow sharper inferences of what type of work went on inside the structure. At this point, students suggested other covert techniques—for example, sneaking in at night or sometime when their presence would not be detected. In other words, observing the structure at rest. Trying to obtain a blueprint or floor plan of the structure was also suggested. Even bombing a section of the structure to see how that affects the output (an idea not so far removed from our U.S. military reconnaissance) was proposed. These approaches resemble, respectively, techniques of neuroanatomical inquiry: postmortem analyses, nonfunctional imaging, and lesion studies. However, as the professor pointed out, none of those approaches would portray the structure at work. Indeed, 20 years ago, when Gernsbacher sat in on this classroom activity, the opportunity to see the brain at work eluded discourse processing researchers.

Times have changed. We now have a chance to generate visual images of that mysterious structure during discourse processing by using functional brain imaging. With functional brain imaging, we can identify the changes in brain activity that occur during various functions. A handful
of studies have already used functional brain imaging techniques to examine brain activity during discourse comprehension.

For example, Gernsbacher, Robertson, and colleagues (Robertson et al., 2000) examined brain activity during discourse comprehension using functional magnetic resonance imaging (FMRI). In one experiment, participants read sentences connected into a discourse, unrelated sentences, and viewed nonalphabetic character strings. The creation of the discourse and disconnected sentences conditions was achieved by using the definite article the in one set of sentences and using indefinite articles (a or an) in another set. This subtle manipulation produced no change in left hemisphere activation, but activation specific to connected sentences was found in medial and superior frontal regions of the right hemisphere.

St. George, Kutas, Martinez, and Sereno (1999) asked participants to read both titled and untitled narratives. They found similar patterns of activation as in Robertson et al. (2000), but noted that the patterns of activation were stronger when the stories were untitled, particularly in the right hemisphere. Interestingly, Mazoyer et al. (1993) reported a study in which participants listened passively to narratives, and none of the right frontal regions reported by Robertson et al. (2000) or St. George et al. (1999) showed activation. The lack of right frontal activation was also observed in Tzourio, Nkanga-Ngila, and Mazoyer (1998). This suggests that the right hemisphere activation may only arise when the participant is actively comprehending a narrative, rather than passively processing speech sounds.

This handful of studies is only the beginning of what will hopefully be a productive direction of future research in discourse processing. Once again, this direction requires an interdisciplinary collaboration between discourse researchers and those who investigate brain mechanisms.

More Computer Technology

One of the challenges facing the field is keeping pace with new technologies that facilitate the collection and analysis of streams of discourse. At present, the ability to collect discourse data exceeds the capacity to process and analyze it. There are countless boxes of videotapes in hundreds of researchers’ offices waiting to be cued up for analysis. We need to design and develop tools that can make the analysis of discourse corpora more tractable and more quickly processed. These tools are needed at multiple points in the process, starting with transcription. We can imagine speech recognition technology that can take video or audio input and generate transcripts that translate speech to
text. Imperfect transcripts would even be useful; researchers could take these first-cut transcripts and then elaborate them with corrections, edits, nonverbal information, and annotations on various levels of discourse codes. There are prototype systems that allow annotation of video directly, which makes it possible to avoid the distortion that occurs when analyzing discourse from printed transcripts (see http://talkbank.org/ for available systems). Creating a database of video excerpts that might be shared by multiple researchers is the goal of MacWhinney’s efforts to develop TalkBank. Web access to such a database would create the ability for researchers with different theoretical perspectives and empirical approaches to examine and interpret the same observations.

The value of bringing multiple perspectives to the same event is illustrated by the special issue of Discourse Processes (1999, Volume 27, Issue 2) edited by Tim Koschmann (1999). There was a 6-minute segment of a problem-based learning group consisting of second-year medical students and a faculty coach. This segment was analyzed by five different sets of researchers whose assumptions and perspectives differed. The analyses set the stage for critical dialogue among researchers with different perspectives about the construction of meaning in this and related contexts (Green & McClelland, 1999).

Tools are also needed to support the laborious process of developing and applying analytic coding schemes. The development of such schemes goes through multiple iterations that follow the principles of grounded theory. We need computer-based applications that allow flexibility in the parsing (and reparsing) and coding (and recoding) of both printed and video data. Some discourse coding software, such as NUDST, have been in use for several years, but they have not had the flexibility required for sophisticated discourse analyses. For example, NUDST assumes a strictly hierarchical structure of language and discourse units, whereas it is widely acknowledged that discourse involves overlapping speech from different participants and networks of codes that deviate from strict lattice hierarchies. We need more methods that integrate video clips of facial expressions, gestures, and context with the speech in a software platform that is easy for discourse researchers to use.

Another class of tools allows researchers to construct graphic representations of the discourse and knowledge structures. These graphic representations consist of a set of nodes that refer to constituents (e.g., words, propositions, states, goals) and relations (arcs, links) that connect the constituents. Sets of nodes are frequently connected by probabilistic weights that signify excitatory (+) or inhibitory (−) activation, as in the case of Kintsch’s (1998) construction integra-
tion model and other neural network models. There needs to be easy methods of zooming into ensembles of nodes that form natural packages of knowledge or reflect particular perspectives (e.g., such as the speech acts of a particular speaker in a multiparty conversation). Graphical tools tremendously help researchers who routinely work with large spaces of knowledge.

From Multidisciplinary to Interdisciplinary Research

In multidisciplinary research, there is a collection of researchers who explore a particular phenomenon from the standpoint of the theories and methodologies of their own disciplines. For example, discourse coherence can be investigated from the perspective of text linguistics, discourse psychology, computational linguistics, artificial intelligence, Conversation Analysis, and so on. A researcher in one discipline may or may not embrace the wisdom collected from the other disciplines. It really does not matter. The important consequence of multidisciplinary research is that the phenomenon is studied from multiple viewpoints rather than a singular viewpoint. Sometimes valuable insights result from peeking over the shoulders of colleagues in sister disciplines. An anthropologist’s theory of power relationships between interactants in a conversation might provide new insights into a discourse psychologist’s interpretation of the language used in a learning situation. However, a discourse psychologist might take a look at a theory of text coherence in text linguistics and conclude that it has little relevance to the construction of coherent representations in the minds of humans.

Interdisciplinary research, in contrast, requires integration of the wisdom of two or more disciplines. Researchers work collaboratively on research projects so there needs to be a coordination of the perspectives of different disciplines. Disagreements surface and need to be resolved. This can be painful particularly when researchers are wedded to their own beliefs and have the illusion that the epistemological foundation of their own field is inherently superior to all others. Fortunately some researchers have the cognitive flexibility to conduct interdisciplinary research and stretch their horizons. There have been noteworthy collaborations between computer science and psychology (Schank & Abelson, 1977), between psychology and literature (Dixon, Bortolussi, Twilley, & Leung, 1993; Miall & Kuiken, 1994), between acting and psychology (T. Noice & H. Noice, 1997), and between text linguistics and psychology (van Dijk & Kintsch, 1983). When AutoTutor was recently developed, there was an interdisciplinary collaboration of approximately 12 faculty in psychology, computer science,

So why is it so important to conduct interdisciplinary research? Is multidisciplinary research unsatisfactory? We believe there are three major virtues in conducting interdisciplinary research. First, there is a deeper communication between researchers from different fields because the collaboration forces them to seek a meeting of the minds. Multidisciplinary research suffers from the complacency of researchers who wallow in their own autonomous, detached, narrow research niche. Second, the science of interdisciplinary research is deeper because the researchers from different fields must confront their differences, defend their positions, and negotiate compromises that accommodate the various trade-offs of alternative solutions. This process typically leads to more precise understanding of the phenomenon under consideration. Third, we believe that interdisciplinary research produces better solutions for practical problems, the majority of which do not respect disciplinary boundaries. This explains why funding agencies have set such a high priority for interdisciplinary research. Interdisciplinary collaborations are painful when the barriers between disciplines are strong, but these collaborations cut quickly through the thickets of unknown frontiers and sometimes, perhaps by accident, help solve difficult problems of humankind.

ACKNOWLEDGMENTS

This research was partially supported by grants to the first author from the Department of Defense Multidisciplinary University Research Initiative (MURI) administered by the Office of Naval Research (N00014-00-1-0600) and the National Science Foundation (SBR 9720314, SES 9977969, REC 0106965, REC 0126265); to the second author from National Institutes of Health (R01 NS29926 and F33 DC05365); and to the third author from the Spencer Foundation and the National Science Foundation (REC 0113669). Any opinions, findings, conclusions, or recommendations expressed in this chapter are those of the author and do not necessarily reflect the views of the funding agencies.

REFERENCES


1. INTRODUCTION


This page intentionally left blank
For the past 20 years or so, I have been reading manuscripts for *Discourse Processes*, a multidisciplinary journal with a tilt toward experimental social psychology. In retrospect, it seems that my most frequent and consistent complaints about submissions I read were that authors were careless in defining both independent and dependent variables, often indifferent to questions of representativeness and comparability, and often acted as if they had forgotten the critical importance of *ceteris paribus* considerations in behavioral research. When the editors asked me to write this chapter and gave me the suggested title, I did not initially realize that I was being asked to suggest remedies for some of my complaints—or at the very least specification of my view of some of the problems. The assignment has required that I look closely at (sometimes contradictory) definitions of *discourse, genres, registers, and contexts of discourse*. It has also constrained me to articulate some reasons why these matters are of great moment in research on language/discourse in use in social contexts.

**DISCOURSE: SOME DICTIONARY DEFINITIONS**

The term *discourse* has gone through complex definitional vicissitudes in its evolution over the past seven centuries. In its first several hundred years of use, the term developed semantic variety, which it had not had in its French roots. On the other hand, many semantic fea-
FIG. 2.1. The text should be read linearly, like any sentence. Following linguistic convention lists enclosed in curly brackets are options; one word should be selected from each list. Readers who perform this exercise, and I urge all readers to do so, will discover that selection of one or another of the choices will change meaning(s) of the text to varying degree. They will also see that selection of one or another choice from a set of choices will sometimes also, and also to varying degree, constrain other selections. Such modest changes can signal differences in genre or register and can, themselves, constitute contexts of text.
tures disappeared over the centuries, and by the 19th century the prevailing sense had become “a spoken or written treatment of a subject, in which it is handled or discussed at length; a dissertation, treatise, homily, sermon, or the like” (OED). Similar variation obtained for definitions of the verb. In recent years, the term seems to have receded from use in public “discourse” and to have evolved toward one diffuse and one specific technical definition.

The first perspective sees “a” discourse as something like an ideological “bundle”, a subculture, or even an arena of special interaction. Thus, there are “discourses” of feminism, environmentalism, struggle against oppression (of a number of specific and named varieties), individualism, sexism, and Marxism. I have not heard of a discourse (in this sense) of adolescence; if there is not such, there plausibly could be. I doubt, however, that there is or could be a “discourse of infancy,” although there may be a number of discourses of childrearing. All of this seems to imply that these “discourses” require some sort of self-consciousness and reciprocal awareness among participants in them.

The greatest bulk of students of such discourses appears to be involved in the burgeoning field of culture/cultural studies. A small subset, students of ideologies, are more akin in their goals and methods to scholars who subscribe to the more specific technical contemporary definition of discourse.

This second perspective sees discourse as spoken or written text in a language, intended for use in the accomplishment of social ends of users (speakers, hearers, writers, readers). According to such a definition, both a single word shouted in warning and a lengthy written legal brief are examples of discourse, but neither the hypothetical examples made up by linguists nor lists of lexical items are. The critical criterion is neither length nor degree of formality, but as Brown and Yule (1983) observed, it is “primarily pragmatic.” If language is a uniquely human attribute, discourse is the language in use that allows human social life. Essentially all humans use discourse; a small subset of us have chosen to study how discourse works. This handbook is about how that study can best be done. My concern is to show how the understanding of discourse and discourse processes requires identification of constituent elements of and constraints on the phenomenon. In the three main sections, I discuss the two constituent elements of discourse (viz. genre

---

1Among obsolete and archaic meanings listed in the OED are (a) “course, process, succession of events or actions,” (b) armed combat, (c) the act of comprehension, (d) a talk or conversation, (e) rumor, and (f) narrative or account.

2Both the linguists’ examples and lexical lists may be embedded in and thus become discourse. See any linguistic monograph or, for a particularly charming lexical list example, Barth’s (1960) *The Sot-Weed Factor*. 

---
and register) and one constraint on discourse—namely, context—which make up the title of this chapter. Before turning to those specifics, however, I briefly consider several matters that, in addition to the polysemy of the term discourse, must be taken into account.

DISCOURSE, SPOKEN DISCOURSE, WRITTEN DISCOURSE, TEXT, AND CONVERSATION

Just as the term discourse has multiple meanings, so also what it is that discourse analysts or analysts of discourse study has other names or labels. First and obviously, there are specialties such as text analysis and conversation analysis, which employ special methods to study particular discourses or particular kinds of discourses. But there are also questions about boundaries of what phenomena actually constitute discourse. Consider the following:

01 Hello.
02 She said hello.
03 She greeted him/her/someone.

When embedded as they are, they occur in a longer exposition, each of these examples constitutes an element of discourse. If a person said “hello,” however, only her actual utterance could be treated as her discourse for analytic purposes; reports about her behavior are not her behavior. It follows from this that we should be careful about accepting as instances of discourse utterances like that reported in 02—particularly as reported utterances increase in length.3

Relatedly, there are questions of boundaries of all varieties of discourse. In the case of conversational discourse, for example, the following are important questions: What is conversation? What are the boundaries of conversations? All talk is speech, but is all talk conversation? Is a service encounter in which a person says “One of those” and is handed something a conversation? Does delivery of an unanswered “Thanks” constitute conversation? Must conversation consist of talk? The answer to this last question is obviously, “No,” but what makes unspoken communication conversation?

3I find particularly off-putting research where investigators report that, for example, “as soon as I completed the five-minute interview I went off and wrote down what had been said verbatim.” Only electronic recording preserves the actually said in anything like a verbatim record; even electronic recording cannot preserve everything that happens in the course of talk.
DISCOURSE WITH A MODIFIER:
LOOKING FOR “HIDDEN” MEANINGS

There are not only questions about the boundaries of types of discourse. Some modes of analysis of discourse assume that meanings of discourse are hidden—sometimes not only from those to whom discourse is directed, but even from those who create it. Critical discourse analysis, like its companion critical linguistics, views discourse with an end to how recipients may be influenced to have one or another perception of a reported event. The work of the East Anglia group (cf. Fowler, Hodge, Kress, & Trew, 1979; Kress & Hodge, 1979) essays demonstration of how, for example, cause and responsibility can be differently attributed through employ of passive versus active verb forms. Relatedly, a host of literary critics and other writers who come to discourse study (usually of written texts) from different arenas attempt through “interruption” to discern the subtle ways in which different syntactic and lexical choices differently influence readers’ understandings of what is being written about. Some speech act theorists argue that words, themselves, can do things like promise or threaten, and perhaps thus “cause” social outcomes.

DISCOURSE—OR SOMETHING ELSE

Some theorists interested in societal change (or other specific outcomes) attribute importance to discourse, but do not look at actual text (cf. Collins, 1981a, 1981b; Bourdieu & Passeron, 1977). Others (e.g., Cicourel, 1981, 1994; Bernstein, 1971) do look at discourse and come to articulate issues of change and micro–macro relations quite differently. One might ask whether data for research on discourse/text are somehow scalable on continua of, for example, concreteness-abstract-

---

4 Silverman and Torode (1980) distinguished between “interpretation” and “interruption” of text. By the former, they mean something like processing text as a representation or reflection of reality in a cooperative effort, which allows the text recipient to see the world much as its producer does. By “interruption” they suggest that the easy flow implied by the interpretation perspective is halted (interrupted) and different possible meanings (some political, many socially consequential) wrested from the text. For example, interruption may identify different voices or, as in Labov and Fanshel (1977), attempt to identify the really said and meant hidden in a string of text.

5 Although the conceptual perspectives referenced by notions such as speech act, performative, illocutionary act, and perlocutionary effect are deeply interimplicated with the foci of this chapter (i.e., genre, register, and context), there is no space here to discuss those implications. For a start on this rich and valuable perspective, see Austin (1962), Bach and Harnish (1979), Cole and Morgan (1975), Harnish (1992), Pratt (1981), Searle (1969, 1979). I drew heavily on this conceptual apparatus in my comprehensive discourse analytic study of an evaluation (1988, especially Chapter 8).
ness (Bernstein, 1971) and empiricity (i.e., actual talk or writing as contrasted to glosses), or whether the “research” consists simply of talk about talk. Relatedly, we can ask whether the focus is on discourse processes or rather careful measurement of processes somehow related to discourse, as in the case of input variables believed to be related to differences in mode, amount, or speed of comprehension.

ARTICULATION OF FINDINGS WITH SOCIAL AND PSYCHOLOGICAL THEORY

Many scholars find discourse as a phenomenon intrinsically interesting, and some specialties, notably conversation analysis (CA), have focused on conversation as an object of study. My personal bias is that discourse studies that illuminate general considerations of a sociological nature (e.g., larger social structural and processual issues) are most valuable. I have found conflict talk to be a rich resource for the study of processes of social conflict in the traditions of Simmel and of such contemporaries as Bailey (1983), Coser (1956), and Dahrendorf (1959; see also Grimshaw 1982, 1990, 1992). A CA study of conflict talk would look at features of examples such as sequencing, interruption, amplitude, and tune, and possibly at how willingness for termination is signaled. It would not, however, investigate the mechanics of how conflict density increases or decreases or the employ of spurious anger (Bailey, 1983).

There are many studies of “discourse processes” that essay refinement of understanding of how discourse works by using made-up examples of discourse in experiments on memory, recognition, or perception. This research reports uncertain results and has little if anything to say about how those results relate to either psychological theory or what individuals do in their everyday lives. There are also,

---

6There are of course theoretical questions of interest primarily for psychological and psychological social psychological students of individuals and individual behaviors, ranging from issues of self-concept to cognition to sharing of perceptions or understandings. I believe that discourse studies are also valuable to the extent that they illuminate these questions and are clearly articulated with them. My own disciplinary background nudges me in the direction of questions about how society works. In the view of Harvey Sacks, conversation analysis does address sociology’s core theoretical concerns (see Grimshaw, 2001).

7The authors of one recent study pondered the failure of their experimental variable to produce expected results and wrote, “It is always possible that the scenarios constructed for this research were too brief and contrived, in some important way unlike natural language use. . . . Although this cannot be totally ruled out, it seems a less than satisfying explanation, however, because one would expect that (sic!) context to have had at least a small effect in the predicted direction.” In other words, it didn’t work—so the design must be valid?
however, thousands of studies that use either actual discourse or notions about discourse in examination of central questions about social behavior. Space considerations allow me to mention a few representative or favorite items in the pages that follow.

For example, Collins (1981a; see Grimshaw [1987] for a gloss) proposed that social/institutional change on the macrolevel occurs as a result of countless microinteractions to which individual participants bring resources, including both their personal histories and a variety of competencies, including language competence, and leave with a different set of resources. Collins did not document his argument with actual discourse. Boden (1994) and Czarniawska (1997, 1999) used oral and written texts of quite different varieties ranging from transient records such as memos and voice mail to published policy statements and narratives of various sorts, and demonstrated his argument more or less in passing (see Grimshaw [2000] for glosses of relevant dimensions of Boden and Czarniawska). Among the many other arenas in which discourse data have been employed to ask questions about social regularities are studies of, illustratively and nonexhaustively: (a) boundary establishment, maintenance, and challenge (Erickson & Shultz, 1982; Grimshaw, 1994; T. Labov, 1980; Urban, 1989, 1991, 1996); (b) collective behavior, social movements, the peace and antinuclear war movements (Chilton, 1985; Mehan & Wills, 1988; Mehan, Nathanson, & Skelly, 1990; Urban, 1988); (c) development and related issues of language and discourse, including bilingualism, literacy, standardization, and uniformization; (d) identity and related matters (Anderson, 1983; Fishman, 1999); (e) negotiation (Firth, 1995; Fisher & Ury, 1981; Grimshaw, 1989); (f) social conflict (see supra.); (g) social control (Atkinson & Drew, 1979; Bernstein, 1971; Brown & Levinson, 1978; Fowler et al., 1979; Goody, 1978); (h) socialization (Bernstein, 1974, 1975; Corsaro, 1985; Ervin-Tripp & Mitchell-Kernan, 1977; M. Goodwin, 1990; Slobin, 1985a, 1985b); (i) sociological variables of power and solidarity (P. Brown & Levinson, 1978; R. Brown & Gilman, 1983).

There are a myriad of other arenas in which language/discourse plays critical roles. To keep the list of references manageable, I have generally cited only works that contain explicit attention to written or spoken text. Thus, I have neglected: (a) a number of critical expository treatments of relevant social processes (early treatments of language and development by Fishman, Ferguson, and Das Gupta or by Das Gupta alone on India, for example, or contrastively, more recent attention to language contraction and decline by Dorian and associates), (b) a number of valuable case studies on the macrolevel (I think here particularly of McRae, 1983, 1986, 1997), and (c) a number of quantitative studies of language variation and shift (e.g., Labov and his associates, Sankoff, or Veltman). I have also slighted rich materials of a more humanist bent on topics such as genre and register.
Discourse psychologists study a range of materials including both oral texts collected in classrooms or tutoring sessions, and so on, and a wide range of written texts including fiction, news stories, and encyclopedia articles to investigate processes of learning and comprehension (Goldman, 1997; Graesser, Mills, & Zwaan, 1977). Other investigators from a range of disciplinary and professional bases have looked at discourse and learning in the classroom (Cicourel et al., 1974; Mehan, 1979), the administration of the law in courtrooms (Conley & O’Barr, 1990; O’Barr, 1982; Philips, 1998), the specialized talk produced in medical settings by both patients and physicians (Cicourel, 1974, 1982; Shuy, 1976) and in therapeutic interaction (Ochs & Capps, 1995; Labov & Fanshel, 1977; Schefflen, 1973), and in numerous other settings including physicists’ labs (Ochs, Gonzales, & Jacobs, 1996) and airport ground operations control centers (C. Goodwin, 1996). These and many hundreds of other studies employ discourse materials to either explicitly or by demonstration address central questions of social theory. As a bonus, many have been valuable in formulating social policies in their institutional locales.

**HYMES’ SPEAKING HEURISTIC AS STARTING POINT**

Although Hymes (1972, 1974) acknowledged the influence of taxonomic schemes of predecessors (he specifically listed more than a dozen; I think he particularly liked conceptual directions suggested by Burke [1969/1945] and Jakobson [1953, 1960]) in helping organize his thinking about how to look at language in use in social contexts, I have always found his SPEAKING heuristic not only helpful in my own work (e.g., Grimshaw, 1989), but also particularly accessible to and useful for students. This heuristic is presented shortly; in my view, all students of discourse are well served by keeping in mind all the various elements of the heuristic in their investigations. They should do this even if their work focuses primarily on, for example, instrumentalities or what Hymes has called *act characteristics*. Little reflection is required to realize that the tremendous variation possible on the several elements makes possible generation of an infinite number of speech situations, and that texts that sound or look much alike may actually be different in Ends of both sorts (i.e., in what interactants seek to accomplish and actual outcomes). So, too, texts that sound or look
dissimilar may actually be similar in both varieties of End. Much of the remainder of this chapter is devoted to explication of the notion and importance of context (comprised in substantial part by S, P, E, and N of the heuristic); register, which involves features of A, K, and I); and genre (G).

**CONTEXT**

All discourse, spoken or written, occurs in social contexts. As the world changes and societies and languages within the world also change, contexts for language in social use change, but unevenly and in different ways. In the 1960s, a graduate student at my university was expelled for using the word *fuck* while confronting a university trustee. Twenty years later, students of both genders reported that they used the term (males still more frequently than females) while at the university (but not in class), but that they would not use it at home in front of grandparents and perhaps not in front of parents. In the 1990s, young women used the term more casually than males had a few decades earlier. Although there may still be some intergenerational taboos, I have the impression that the grandparents themselves increasingly use the term, perhaps sometimes shocking their children and grandchildren. There are variations in use and in response, of course, but anyone who has watched contemporary American films for 10 years or more knows that norms about taboos in public use have changed. Movement across contexts where there are different normative expectations regarding appropriate lexical choices could be mildly unsettling; there are circumstances under which failures of awareness regarding local norms or inability to produce local forms could be quite costly.
Differences in appropriateness norms may be considerably more portentous than shifts in widespread societal perceptions of what constitutes vulgarity.\textsuperscript{9} There are contexts of discourse in which participants’ worldviews are so different as to make texts that would appear to be transparently meaningful mutually inaccessible. For example, consider differences in the discursive worlds of pious and zealous fundamentalists of different religious faiths prepared to kill in the name of their deities—and the discursive worlds of those with different beliefs. Consider the damper on spontaneity in discourse production that results from living in police states where what is said or written may be monitored by the Gestapo, the NKVD, or even one’s own children. Older readers may recall the effects on both private and public discourse of the United States’s milder McCarthyism. Consider, finally, how differences in conditions in known worlds may constrain perceptions and text construction (spoken or written). What does it mean to say or write “I am elated” or “I am depressed” in different times—for example, Renaissance Florence, London during the Blitz, postrevolutionary days in Moscow, depression America, the New York Stock Exchange in the late 1990s, Washington, DC during the early days of the Kennedy administration?\textsuperscript{10}

Although normative constraints on code or intracode selection may shift and governments with different ideologies and ruling styles may come and go, some kinds of contexts persist. Many Ends of both varieties are much the same now as they were centuries or millennia ago. Weaker parties hope to get things from stronger parties—and talk and/or write letters to the stronger parties aimed at getting gifts, forgiveness of debt, positions for self or others, avoidance of punishment, or future favors. A charming recent study of patronage-seeking letters in the Italian Renaissance (McLean, 1998) reveals that considerations of relations of power and affect, and of the utility of the desired goal, operated then as they do today. Other letters from medieval times into the last century show that, although themes invoked vary with goals and participants, discourse has remained similar (Kermode & Kermode, 1995; Moriarty, 1989; for a rich fictional example, see Barth, 1979). In like manner, although dress, language, and particularities of goals and complaints vary, it appears that arguments and disputes constitute discourse situations (contexts) much like those of earlier societies (Steinberg, 1999). This section identifies (a) features of context that

\textsuperscript{9}Some observers of course see the change just described as a portent and indicator of a decline of “American values.”

\textsuperscript{10}Not everyone experiences collective “uppers” or “downers” in the same way of course. See Merton and Kitt’s (1950) insightful piece on “relative deprivation.”
are differently subject to change, and (b) some characteristics of contexts that are variable (e.g., commitment to outcomes, concern about perceptions of self by other participants or audiences, nature and intensity of affect directed to other participants, etc.). I start, however, with two distinctions relevant to all contexts.

Text and Situation, Immediate and Remote, High and Low Salience

Although Halliday credited Firth and Malinowski for early contemporary articulations of the analytic importance of contexts of situation and culture, many researchers currently engaged in the practice of discourse analysis initially discovered a distinction between embedding text and other contexts in Halliday’s work. Generally, context of text means the embedding talk (or written text) in which words, clauses, sentences, utterances, speeches, and so on (and their written analogs) are found and which, independently of contexts of situation, might reasonably lead a hearer, reader, or analyst to assign different meaning to that which is embedded. By context of situation, Halliday meant what sociologists from Thomas to Goffman have called (with varying degrees of caution and specificity of definition) simply situation, what Hymes (supra.) referred to as setting or scene, and what Halliday sometimes identified as extratextual environment.

Halliday (with Hasan, 1976) observed that coreference (as one of many devices for cohesion) in text can extend over long sequences. They wrote, “We find in everyday conversation elements turning up which presuppose earlier passages from which they are separated by many minutes and even hours of speaking time; and writers exploit this potential by making cohesive ties across very long stretches of text” (p. 294). The likelihood that remote contexts of text and situation will be recalled and influence current interpretations of ongoing interaction depends on interaction of actual remoteness, salience of the events when they originally occurred, and perceived relevance for the ongoing. I vividly recall something a grandfather told me in a conversation more than 60 years ago; people report reading something that

11 Because text is the principal focus of Halliday’s attention, “context of text” is defined implicitly. See sources cited in Grimshaw et al. (1994), from which this discussion of context borrows heavily. Readers should find Halliday and Hasan (1976) particularly helpful. Anthropological linguists, sociolinguists, and others with social science habits and biases tend to take the importance of “culture” for granted in much the same way that Halliday assumes the centrality of “text.”

12 In addition, retrospective reinterpretations of text and situation experienced earlier.
changed their lives. Such events can profoundly influence interpretations of newly experienced discourse in ways that are simply inaccessible to interlocutors or analysts. The situation is further complicated, of course, by subtle changes in “meanings” of earlier experiences and revisions of recollections.

Context: The Mundane

I noted earlier that college undergraduates are aware of rules about discourse and contexts; they report that they talk to each other differently in class and at home—and in front of parents. Many of them report having had explicit instruction about where and to whom what may be said and in what ways. A favorite story of mine came from a student who reported being in church while very young, whispering, and then having her knee painfully squeezed; the squeeze was accompanied by a verbal warning from her mother, “Wait ’til I get you home!” At least older readers of this chapter can recall the injunctions, “Little pitchers have big ears,” “Little children should be seen and not heard,” and “Not in front of the children!” These are, of course, all instances of contexts of situation; the reason that both kinds of context must be attended by participants and analysts is that what a particular piece of text “means” can vary with differences in either. To take another obvious example, the phrase “That just isn’t so!” is interpreted differently when heard in a kitchen, court, or classroom (or on a stage). Still further differences in interpretation occur when the phrase is preceded by “And then you/she/I said” or “Give me an example of a declarative.”

Both Halliday (1994) and Hasan (1994) were co-investigators in a collaborative study of extended naturally occurring discourse of a dissertation defense. Halliday (1994) gave only modest attention to context of situation. The field of discourse having been defined as serious, academic, and so forth, he did not feel it necessary to ask Hasan’s question about when context and discourse have changed sufficiently that redefinition must be taken into interpretive account.

The field of discourse/context of situation for many discourse analyses is that of quasivoluntary participation in experiments on production, perception, interpretation, recall, and so on of discourse variables. Caution is required if experimenters move to generalize findings to other discourse situations. An assumption of at least some experimental research is that subjects are more or less interchangeable. Contrarily, it can be argued that at least in some circumstances partici-

13 Grimshaw et al. (1994) contains eight independent studies of the same record of 12 minutes of a dissertation defense.
pant characteristics can be critical indeed in defining contexts of situation. Consider classrooms before and after arrival of teachers, workplaces with and without supervisors, press conferences before and after arrival of the “star.” The critical dimension here is power. Power relations are also important in ensuring that experimenters are successful in obtaining subjects. The capacity to command presence is not the same, however, as that to compel attention. Consider the following aspects of context of situation and how they may heighten/enhance or reduce attentiveness of speakers and hearers alike (to say nothing of audiences):

1. increasing danger (of being “caught,” injured, or insulted) focuses attention;
2. real or scripted dramatic tension focuses attention on performance (not on surround);
3. competing events may distract, reduce focus;
4. preoccupation may distract, reduce focus;
5. hypermonitoring\textsuperscript{14} distracts, reduces focus;
6. individual and/or collective grief/pain can distract, reduce focus;
7. indifference to “official” focus can cause inattentiveness.

One may reasonably ask, how are differences in focused attention part of context of situation—and how do they affect outcomes (experimental or otherwise)? One should ask, is the mundane, whether in naturally occurring interaction or in a rigorously defined experiment, as uncomplicated as we sometimes like to think?

**The More Obviously Complicated: How Much Less Mundane?**

Likely influence/constraint of some features of context for some interaction/discourse are transparently obvious. Changes in venue, participants, or topics can collateral change not only what is said, but also what is meant by what is said and in what gets interactionally accomplished. The appearance of a bereaved, known-to-be-terminally ill, or cuckolded party can indeed put a damper on the ongoing. Far greater complexities can be generated, however, by the fact that, although knowledges and understandings are massively shared, interlocutors

\textsuperscript{14}Hypermonitoring and other hyperinvolvement occurs when an interactant is watching so closely for some behavior (own or others’) that she misses/overlooks her own potentially damaging behavior or relevant behavior of another (Grimshaw, 1989).
can maintain conversation even though they are sometimes far from certain that they understand either what is said or what was intended by that said. Moreover, sometimes they are confident that they do understand the ongoing, but are wrong. Clark and his associates (Clark, 1992)\textsuperscript{15} helped us see how interlocutors can plumb the outlines and degree of common grounds. It is usually the case, however, that analysts cannot know the extent and consequentiality of the unshared.

Cicourel (see especially 1994) has argued with increasing forcefulness for three decades that cointeractants take much more into account in making talk sensible than “simply” the actual words—\textit{and} the kinesic accompaniments, \textit{and} the physical setting, \textit{and} the history of social relationships among the participants. In his study of the aforementioned dissertation defense, Cicourel (1994) argued that familiarity with social exchange theory, conventions of significance testing in statistics, and the candidate’s doctoral paper are important elements in context. Cicourel plans to demonstrate that analysts are less successful to the extent that they do not determine what elements of contexts are employed by interactants in their sense making so that they, the analysts, may attend to those same elements in \textit{their} interpretive labors.

Cointeractants (including, of course, correspondents) will better understand one another to the extent that their talk (or written messages) invokes shared knowledge. Analysts are more successful in interpreting the ongoing to the extent that \textit{they} “share knowledge” with participants in the discourse under study. Insistence on shared knowledge can quickly turn into a \textit{reductio ad absurdum}, however, if pushed too far; neither Cicourel nor anyone else is prepared to argue that defenses can be studied only by those who have read candidates’ dissertations (let alone those expert on the area of work in which research was done) nor that only heart surgeons can talk or write intelligently about open-heart procedures. If that were to be the case, no analyses could ever be done because there are not many people who have shared the same experiences, read the same books, had the same arguments, and so on. Much more important, there are two reasons that such matching of cointeractants, conversational participants, and analysts who study them is not necessary. The first is that there is some knowledge that is, for all intents and purposes, shared by all social members (including analysts), and this truly shared knowledge in-

\textsuperscript{15}Clark also does experiments. He is sensitive to \textit{ceteris paribus} and other considerations.
cludes what Cicourel (1974) has called interpretive procedures. The second is that, as Burke (1994) and Cook-Gumperz and Gumperz (1994) argued in their studies of the defense (and elsewhere), conversational coparticipants continuously and redundantly provide the contextual and other information necessary to make sense (and, specifically, the sense they intend) out of what they are saying.

Discursive, “Scientific,” Humanistic, and Personal Approaches to Context

As individual members of societies, most of us recognize and know, generally, how to behave in the social contexts in which we find ourselves. If recognition is delayed, we have ways that ordinarily work to make our surround sensible (Cicourel’s [1974] “interpretive procedures,” for example). To the extent that we are not able to make sense out of situations in which we find ourselves, we ordinarily increase our efforts. Continued nonsuccess can generate stress, often substantial stress. Fortunately, what we are used to doing usually works. Skilled writers of fiction, drama, biography, and autobiography, and often of other genres from jokes to history to poetry, invoke the same shared knowledge, including that of interpretive procedures, to lead us to understanding subjects’ situations and likely reactions to those situations. Such authors employ their authorial control, however, to make contexts deceiving, opaque, or in other ways deeply problematic for participants. It is often the unexpected, but not obviously impossible, that makes drama dramatic and fiction compelling. Of course, authors can make the impossible possible by generating various sorts of fantastic worlds in which at least some procedures and behaviors seem to continue to work. All this underlines the obvious: We all know a tremendous amount about an extraordinary variety of contexts, including many we will never personally experience and others that can never occur again.16 One of the things we do not know about contexts, however, is precisely how they influence behavior, either in terms of socially relevant variables or precise mechanisms. For present purposes, the principal contributors to such understandings have largely been sociolinguists and ethnographers of communication and discourse psychologists.

---

16Steiner (1992), a humanist sensitive to an extraordinary range of issues relating to uses of language in social contexts, credits unnamed linguists with invention of the term preinformation for dealing with “the problem of necessary and sufficient context, with the amount of prior material required to understand a given message-unit” (p. 11).
The massive increase in recent decades of interest in the subject matter of this handbook has been accompanied by more and increasingly sophisticated attention by discourse analysts and discourse psychologists to the contexts of text and situation (or arena, circumstances, scene, setting, or surround) in which text is generated. Space considerations allow mention of only three volumes chosen to represent the range of discursive (but nonetheless rigorous) and scientific approaches.

Readers interested in notions about conceptualization will be richly rewarded by close reading of Goodwin and Duranti’s (1992) introduction to their co-edited volume (Duranti & Goodwin, 1992). They began by underlining the importance for the meaning and influence of context of (a) participants’ perspectives, (b) ongoing activity as a basis for contingent relevancies, and (c) multiple mutabilities of context. Believing that knowing alternative perspectives enhances context-sensitive investigations, they then provided brief examinations of (a) ethnographic and philosophical precursors, (b) Soviet approaches, (c) work in human interaction including special attention to cybernetic and dramaturgic metaphors, (d) ethnography of speaking, and (e) ethnomethodology, conversation analysis, cognitive sociology, and Foucault. There follow 14 chapters that explore a variety of methodological, substantive, and theoretical issues occurring in work that is comparative on temporal, cultural, and intrasocietal (institutional) dimensions.

Also discursive in approach, but considerably more narrowly focused, is the delightful report by Silverstein and Urban (1996) and colleagues from widely differing backgrounds on decentering—examination of different meanings of “same” texts in different contexts. Some of what decenterers do is much like the interruption of Silverman and Torode (1980). A deeply interesting dimension is added, however, by scrutiny of a “same” text in different manifestations (e.g., an original oral rendition, electronic recording, phonetic transcription, transcription in original language and translation). Consider further these different renditions incorporated in one or both languages into oral scholarly presentations, scholarly and popular papers, presentation to the original informant or performer, or whatever. The meaning of such a text, as perceived by both emitters and audiences, is influenced by a host of variables including contexts of text and situation. Contemplate how different history and how different our world would have been, absent certain discourses or with different readings/interpretations of discourses that occur. Consider how carefully lawyers study contracts looking for possible variant interpretations—or diplomats and the military, treaties, or critics prose and poetry. Consider further again the complexities added when contracts or treaties are intended
to regulate behaviors of parties of different languages/cultures, critics to assess productions in translation, or, still further, when notes employed in generating the several varieties of texts are available for use in the search for meaning(s).

A quite different approach to matters of context can be found in the careful experimental work of discourse psycholinguist Herbert Clark and his associates (1992). Clark introduced his volume with invocation of the importance of context(s), shared meanings, audience design, and concomitant differences in comprehension by addressees and overhearers, coordination of meanings, and so on. Rather than careful ethnographic observation of cases, however, what Clark had in mind was a carefully defined and empirically demonstrable/testable/measurable conceptualization. Clark and his collaborators have essayed this with a series of carefully designed, cumulative, and elaborative experiments in which precisely defined dimensions with often quite familiar designations (mutual knowledge, mutual beliefs, mutual assumptions, mutual attitudes) are examined in interaction with such variables as community membership and copresence (physical, linguistic, and indirect) in context (defined as “information that is available to a particular person for interaction with a particular process on a particular occasion” [p. 65]). He suggested that participants bring to interaction diaries (with more recent events more accessible), dynamic dictionaries, and encyclopedias, which they employ in interaction with partners sharing greater or lesser proportions of those resources. The experiments are imaginative and plausible. Differentiation in performance success is generated through introduction of constraints of time, complexity, and ignorance. Outcome variables are often measured quite simply. For example, Clark found that the number of words required for identification from among 12 figures declined as more and more of the figures were eliminated.

Advances in the past half century in knowledge of what context is and what its effects are have been truly remarkable. It may be that before long analysts will need to consider the possibility of points at which additional contextual detail ceases to provide proportionate increments of new knowledge/understandings of discourse processes.

**GENRE: SELDOM LABELED BUT CONTINUALLY INVOKED**

With his combination of historical awareness and sensitivity to where study of language in use in its social contexts was and should be going.

17On diplomatic negotiation, see Smith’s (1989) intriguing examination of USA-USSR negotiations, and on negotiation by high-ranking military officers, see Grimshaw (1992).
Charles Ferguson (1996b/1985), in a paper that began with a learned disquisition on Aristotle on the topic of genre, wrote:

Two powerful tools of analysis and understanding available to the student of human language are the analysis of types of discourse and the analysis of how language varies depending on the occasions of its use. The former, the study of discourse types, is what is traditionally called “genre analysis.” The latter, the study of language variation by use, is referred to by some as “register analysis.” (p. 167)

These are the definitions that inform my thinking about genre and register. I note in the following some sometimes confusing alternative usages.

The discourse that is the topic of this volume is not the “discourse” referenced in most everyday talk (see supra. on definitions). The term context appears most frequently embedded in expressions where speakers, often public figures, claim that a quotation of something they said or wrote has been “taken out of context.” The term genre, in contrast, has been a staple conceptual notion in contemporary intellectual discourse, public and private, spoken and written. Although it has never had the range of colloquial meanings that genre enjoys in French (its language of origin), the term has had its definitional vicissitudes (Chapman, 1992; Steen, 1999). In its broadest sense, the term simply means “kind” or “sort” and is used quite comfortably in identifying/locating types of music and other performance arts, painting and other plastic arts, film, discourses written and spoken, and even cuisines.

The term’s colloquial meaning is then much the same as that understood by students of discourse both in the humanities (including literary critics) and in the social sciences. That sense is essentially that employed by Hymes in his SPEAKING paradigm (supra., p. 33)—that is, varieties of discourse such as “conversation, lecture, curse, prayer” and so on (we see later that to identify a discourse variety as a sermon on the basis of participant identity, location of event, and topic alone may make things seem less complicated than they may actually be). There are a few genres, such as “the dozens” in speech and the novel in written text, that are more or less specific to oral or written text, however, most genres appear in both varieties.

Some Dimensions of Definition
and Some Emblematic Typologies

It would be a daunting task to attempt classification of either spoken or written varieties of discourse alone; to attempt to generate an exhaustive classification/taxonomy for “all” varieties of discourse would be to
assume a life’s work, which could in all likelihood not be completed. Consider, for example, just a few of the many distinctions that might be incorporated in classification (and some of which raise questions about what criteria are for inclusion as discourse): (a) attributes such as domain, medium, content, form, function, type, and language; (b) types of discourse such as the classes narrative, argumentation, description, and exposition; (c) contemporary or earlier periods (often by century); (d) provenance by country, region, class, gender, and other more esoteric characteristics such as degree of authoritarian-ness of government and other agencies of social control; and (e) intended audience divided along such lines as age, class, education, and gender, but also by more esoteric considerations of political party, religion, or scientific specialty and even possession of such varied personal attributes as ambition, cognitive capacities, humor, openness to new experience, sympathy, and the full panoply of individual past experiences. This list of relevant distinctions among varieties of discourse could be extended; it is hard to think of social science variables that do not influence in patterned ways how people talk or write to one another and how people interpret and act on that spoken or written discourse.

The possibilities just listed constitute no more than a partial inventory and every difference mentioned is subject to elaboration. Consider the matter of superordinates of genres, genres, and subgenres— that is, levels of distinctions or types within types. Some writing (and some oral material) is (not always consensually) labeled literature. Literature is often regarded as constituted by fiction and perhaps nonfiction possessing certain qualities. Within fiction there are genres of the novel, poetry, drama. Detective stories, historical “epics,” pornographic potboilers, romances, spy tales, and westerns are all subgenres of the novel; all of them can in turn be divided into additional recognizable subtypes. There are similar complexities in the organization of the genres of nonfiction ranging, illustratively, across, autobiography, biography, current events, cook books, do-it-yourself, exposés, research monographs, self-help treatises, sports, and textbooks. For Barthes and the structuralists, “genre is a set of conventions representing the shared expectations of writer and reader, and changing from one period or another” (cited in Chapman, 1992).

Such a definition only appears to be straightforward. One difficulty is that, as we have just seen, there are for all practical intents and pur-

18Editor Graesser suggests that such an effort could prove useful, however, and recommends Biber (1988; Biber, Conrad, & Reppen, 1998).
19I draw here on Steen (1999) and Chapman (1992). Many other sources would serve equally well, and many additional dimensions could be identified.
poses an infinite variety of genres. Although writers, readers, speak-
ers, and hearers may share some sets of conventions representing
shared expectations, there are others that are shared not at all. Let me
just hint at some possible complexities. Consider, first, a situation in
which behavioral responses would seem to suggest that conventions
are shared. Since the advent of TV, there have been a series of black
stand-up comics who do shows for audiences that range from largely
White to largely Black; Richard Pryor was a pioneer in the genre (if
such it is) in the 1970s, Eddie Murphy starred in the role in the 1980s,
and Chris Rock flourished in the 1990s and into the current century.
Among conventions of this genre are profuse use of obscene language
(but not blasphemy), deprecation of Blacks, commentary on inter-
gender relations among Blacks, condescension toward Whites, and
talk about the unwisdom of racism. Audiences laugh heartily. Black
men and women laugh most heartily at different times. Do Whites and
Blacks similarly laugh more heartily at different junctures? Are some
Whites so bicultural that they emotionally respond in the same way as
Blacks to conventions of the genre? Have they become sufficiently in-
tellectually assimilated that they recognize the points at which laughter
(of various sorts) is appropriate? How do they know how to behave
when a comic transgresses the boundaries of what, even for such audi-
ences, is deemed acceptable?20 Do they know? What is this genre? Is it
“minority standup”? If so, what sorts of changes in conventions should
we expect when Black women do it, or other ethnic minorities, or ho-
mosexuals, or whomever? What sorts of changes in conventions of the
genre have occurred over the years? Do they reflect changes in race re-
lations? In conceptions of what is funny?

It may be that just as conventions of genre change over time, shared
expectations of writers, readers, hearers, and speakers also change as
pools of producers and consumers of genres become increasingly het-
terogeneous. A recent number of the New York Times Magazine (May
7, 2000; see especially the short piece by Gish Jen and survey results,
passim.) focused on changes in the “American mind” on such varied
matters as intimacy, morality, identity, religion, worries, and anx-
ieties. Americans apparently are of different minds about these mat-
ters. How, we might ask, do people who are so different share enough
to communicate successfully? People do, of course, communicate suc-
cessfully at least to some (largely uncomprehended and unmeasured)
extent—all the more reason to scrutinize differences among partici-
pants as a critical element in context.

20For example, by jokes about icons such as John F. Kennedy or Martin Luther King.
Some Vaguenesses and Boundary Uncertainties

In broadest outline, scholars would tend to agree on what they mean by genre. There are, in practice, some problematic areas for both interaction and scholarly research. For example, what are the boundaries among alibis, bald-faced lies, evasive euphemisms, exaggerations, fabrications, falsehoods, fibs, public relations pufferies, romantic “lines,” stories, tall tales—or hyperbolic teasing? What about stretching the truth? White lies? Are boundaries located in the texts or rather in the assumed motivations of those who produce the discourse? Does committing the untrue to writing somehow change its status? In what ways? Participants in interaction can be quite unsure about what kind of talk they are dealing with. Although people often behave in highly predictable ways so that we know we can expect unlikely stories from some and poker bluffing from others, the fact that almost everything said is subject to being denied can make decisions about the nature of discourse chancy in the extreme.

Similar boundary problems are endemic. We have not only the question of what, exactly, constitutes a genre; I discuss one possible answer to this question—employing quantitative differentiation of genres—later. There are additional questions about at what point the identification of subgenres becomes unproductive. If there are systematic differences and they are not genres nor genre related, what are they? There may be clues in the key of Hymes’ heuristic. For example, what about sincere and insincere productions of “same” stories? How are boundaries between genres marked? What are we to make of the television program categories such as “comedy/drama”? If arguments and disputes are sociolinguistic genres, how are they distinguished from “discussions” (in one direction) and “fights” (in another)? There are different cultural and even family criteria for distinguishing between friendly bickering and constant carping.

Elements of Variation in Subgenres: Componential/Distinctive Feature Analysis

In his classical review and commentary on genres in the American Black speech community, Abrahams (1972; a published version [1974] omits some detail; see also Kochman, 1972; Labov, 1972a) identified as basic a distinction between “house talk, especially ‘around moms,’ and street talk”; the “minority standup” mentioned earlier exemplifies “street

---

21 Note, for example, the American difficulty in dealing with Italian discussione (Corsaro & Rizzo, 1990).
talk.” Abrahams observed that street talk is primarily a male activity that does occur away from the home; that it encompasses a wide range of special ways of talking, much of it aggressive and competitive in nature; and that talk and competencies in talk by community members are a major focus of talk. Among special ways of talking at different levels of genre specialization 30 years ago, Abrahams mentioned: signifying, sounding, charging, cracking, rapping, mother-rapping, playing the dozens, mounting, harping, talking shit, running something down, running a game, hipness, putting down, low-rating, copping a plea, tomming, putting on, loud-talking, louding, marking, jiving, shucking, hooawing, talking hooaw shit, bragging, boasting, and capping. Most of these ways of talking can be heard in the performances of black standup comics, some of them are becoming familiar in school settings, and some have become part of the repertoires of some Whites.

Are these different ways of talking genres? Subgenres? If they are, how should they be taken into account in the study of discourse? How may they be systematically differentiated? I return to these questions momentarily. Let me first ask similar questions about a cluster of speech acts. Consider a collection of words that I treat as related to the speech act verb scold: admonish, ass chew, bawl out, berate, blister, call down, castigate, censure, chastise, chew (out), chide, correct, criticize, dress down, excoriate, expostulate, harangue, lecture, objurgate, ream (out), rebuke, remonstrate, reprehend, reproach, reprove, revile, roast, scold, tell off, tongue lash, upbraid, vilify, vituperate. The behaviors associated with the listed speech act verb labels have in common that they are a response to a delict by the recipient of the behavior. Table 2.1 shows, however, that they can vary with regard to other features, including anger, formality, concern for the recipient, and so on. In the illustration, two of the discourse types—those called chew out and roast—share all the features listed; determination of whether they are true synonyms would require both study of actual usage and quantitative investigation of linguistic characteristics. It is clearly possible to distinguish discourse types that may be quite similar in some ways. Whether it is appropriate to label the discourse types identified as subgenres of some sort may be a more difficult question for some. My own sense is that they are. For example, has

---

22 The list is by no means exhaustive; many readers have favorites of their own. As we all know, important degrees of intensity can be signaled by inclusion of modifiers such as gently, mildly, strongly, or severely.

23 In a draft of his 1972 paper, Abrahams employed a variety of componential analysis to distinguish among Black street talk discourse types; the material was not included in the published version apparently because Abrahams found too much temporal and geographical variation.
### TABLE 2.1
Hypothetical Componential Analysis of Some Behavior-Modifying Discourse

<table>
<thead>
<tr>
<th>Discourse Type Element/Feature</th>
<th>Chew Out</th>
<th>Correct</th>
<th>Dress Down</th>
<th>Lecture</th>
<th>Reprimand</th>
<th>Revile</th>
<th>Roast</th>
<th>Vilify</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FRS presence required</strong></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>O</td>
<td>+</td>
<td>O</td>
</tr>
<tr>
<td>Delict by FRS</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>O</td>
<td>+</td>
<td>O</td>
</tr>
<tr>
<td>Correctable error by FRS</td>
<td>O</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>—</td>
<td>O</td>
<td>—</td>
</tr>
<tr>
<td>Asymmetrical power, hierarchy</td>
<td>+</td>
<td>O</td>
<td>+</td>
<td>O</td>
<td>+</td>
<td>O</td>
<td>+</td>
<td>O</td>
</tr>
<tr>
<td>Dominance strategy/assertion</td>
<td>+</td>
<td>—</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>O</td>
<td>+</td>
<td>O</td>
</tr>
<tr>
<td>FRS's character defective</td>
<td>+</td>
<td>—</td>
<td>O</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Concern for FRS feelings</td>
<td>—</td>
<td>+</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Spontaneous, informal</td>
<td>+</td>
<td>O</td>
<td>O</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Humor</td>
<td>O</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Blasphemy, obscenity, vulgarity</td>
<td>+</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>O</td>
<td>+</td>
</tr>
<tr>
<td>Anger, aggression toward FRS</td>
<td>+</td>
<td>—</td>
<td>O</td>
<td>—</td>
<td>O</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

*FRS = Focus/recipient/subject, + = Feature present, — = Feature absent, O = Feature may or may not be present.*
anyone seen a movie about the military or in which the military plays an important role where no character was ever reprimanded or chewed out? Moreover, there are clichés in many of those same movies about tyrannical noncoms who ass chew and upbraid their men to “get them ready for combat,” thereby saving their lives.

Some Local Consequences of Genre Differentiation: Classroom, Clinic, Counseling Office (and Interview), and Courtroom

Another question about differences in genres or other sorts of discourse types is whether they have any effects on people’s lives. The answer is, of course, “Yes!” There is no need to provide an elaborate demonstration of the accuracy of this assessment (see Grimshaw [1979] for an introduction). I do no more here than reference four well-documented arenas where types of discourse definitely affect outcomes. Suttles (1968) observed problems of communicative nonsuccess and conflict among four minority communities (Blacks, Italians, Mexicans, Puerto Ricans), all of which lived in and on the street in the near west side Chicago neighborhood they shared in much the same way as Abrahams describes urban Blacks as living. If such groups, whose domestic ecology and discourse styles are in constant contact and in fact quite similar, have problems in communication, how much greater must be the problems of communication of any of these minorities with middle-class Anglos and the societal institutions that the Anglos generally control. There are, of course, more differences in oral discourse production than those related to genre. I turn shortly to differences in register and how those differences can affect interactional accomplishment. Initially, however, let me identify a variety of differences in discourse that lead to communicative nonsuccess and other unhappy outcomes in a sampling of institutional settings.

The different varieties of Black street talk (genres, whatever) identified by Abrahams and listed before have some features that White children and adolescents in the United States have historically been socialized to avoid (e.g., confrontational attitude, loud voice, casual use of vulgar language, self-assertion).24 White male readers of this chapter born before World War II would rarely if ever have said anything like, “I

24I emphasize historically because increasingly White male children are exposed to discourse with such features in school and on athletic fields (and on TV), admire the talk as “tough,” and attempt to emulate it. Given other changes that are occurring, it may be that I should delete the gender modifier.
am the greatest!” Nor when we were in school would we have inter-
rupted fellow classmates who had been awarded the floor by a teacher
and loudly challenged them. During the last quarter of the 20th cen-
tury, I regularly taught introductory level courses on language as a so-
cial problem. In the 1970s, I regularly had students read Abrahams’
(1970) Positively Black and in the 1980s and beyond Kochman’s
(1981) Black and White Styles in Conflict. White students, particu-
larly those from nonurban areas, would frequently come and thank me
for having them read the books, saying that they explained things that
had long been hard for them to understand—for example, what ap-
peared to them to be unwarranted aggressiveness and hostility of
Black fellow students (encountered for the first time when they arrived
at the university as entering students).

Interviews are a genre with a great variety of subgenres familiar to
readers of this chapter. Some are highly scripted, whereas others are
not easily distinguishable from everyday conversation. The New
Shorter OED (Brown, 1993) lists four varieties of interviews—namely,
medical and other professional consultations, those of journalists with
newsmakers, those between employers and potential employees, and
police interrogations. Even this modest subset of interview types al-
lows us to discern a number of dimensions of variation. For example,
(a) What is the range of topics that can be legitimately pursued? (b)
What are the power relations among participants, and what power
does the interviewer have to coerce answers or even participation? Are
threats to an interviewee’s family acceptable? Physical torture? (c) Do
the interviews have societal support? (d) What constituencies are ac-
corded rights to see interview findings? (e) To what extent are answers
to the kinds of questions just listed formalized in “rules” (e.g., highly so
in the courtroom and much less so in intrafamilial interviewing)? (f) To
what extent is humor permissible? Are parodies of interviews inter-
vews? (g) Do participants expect that something will be learned from
interviews (e.g., consider the predictability of both questions and
answers in postgame sports interviews). If not, what motivates partici-
pation? Is what goes on in “talk shows” interviewing? Wedding ceremo-
nies? Bar mitzvahs? Exchanges of questions and answers in presiden-
tial debates?

If some people who come to college counseling sessions (Erickson &
Shultz, 1982), job interviews, medical intake interviews (Cicourel,
1982; Shuy, 1976), or police interrogation rooms speak more like the

25By the late 1960s, however, the routine first words of a teenage son who lived with
us on entering the house were, “No applause, please, just money.”
“gatekeepers” who interview them and others less like them, it should not be surprising that those with most similar ways of talking are, ceteris paribus, also more likely to be satisfied with outcomes. This is true not necessarily because of favoritism, but minimally because there is more likely to be successful communication. Interestingly, Erickson and Shultz reported that differences in outcomes in counseling interviews occur not just between Blacks and Whites, but between members of ethnic groups in encounters with gatekeepers of same or other ethnicity. Moreover, the differences reflect not just ways of talking, but shared knowledge or its absence of topical areas as well.

As noted, some important differences in types of discourse are not differences in familiarity with genre types, but rather differences in competence in language varieties. It is true, first, that there are class and regional dialects that have higher prestige than others and that speakers of higher prestige variants are, ceteris paribus, more attractive as potential employees, patients, or clients. Anglo speakers of middle-class variants of Standard American English are more likely than speakers of most other varieties of American English to be familiar with the special codes of legalese or medical practice. In the case of law, participants in disputes even in small claims court are more successful to the degree that they invoke notions of “rules” rather than “relationships” (Conley & O’Barr, 1990). Thirty or 40 years ago, doctors and some of their patients had different lexicons for diseases, disorders, and complaints (cf. Shuy, 1976). In the intervening years, doctors have been learning both to understand how patients of different class backgrounds talk about their illnesses and to make themselves better understood when talking to patients.26

Varieties of Discourse: Ends₁ and Ends₂

Blacks, particularly young males, learn street talk because it is the oral discourse mode used in their speech community, and competence in it displays coolness, toughness, and other positively valued attributes. Counseling interviews are expected to help student clients to optimize their educational opportunities. The goal of “doctor talk” is to identify the medical problems of patients with an end to relieving suffering and curing disease through careful specification and clear explanation of alternatives of cause and cure. The purpose of legalese is to disambiguate as fully as possible the discourse of law. For the uninitiated who in-

26 Although not appropriate for citation here, there is an extensive literature on medical communication and efforts to train new doctors for more effective communication with patients. See, illustratively, work by Aaron Cicourel and Richard Frankel.
teract with Black youth or student counselees of different ethnicity than their counselors, with medical personnel, or with court officers of all sorts, the outcomes may be very different. Whites exposed to Black street talk may be intimidated; they may interpret what is going on in ways different than their interlocutors have intended (it is also the case, of course, that some Blacks may be aware of intimidation effects and exploit them to their own advantage). Counselees of different ethnicities may be excluded from desirable opportunities or be guided into inappropriate programs. Patients may receive incorrect treatment. Lay participants in the judicial system may be denied justice, be swindled, or make unfair decisions about defendants. Unfamiliar discourses may be impenetrable because of differences in contexts (including lack of shared knowledge), genres, registers, or other language variation; problems of interpretation are there regardless of whether recognized by interactants and analysts.

Discursive, “Scientific,” Humanistic, and Personal Approaches to Genre

Printed programs for solemn events such as church services, commencements, funerals and memorial services, dedications, and awardings of social recognition of one sort or another often list genres of talk (acceptance speeches, benedictions, prize essays, sermons, testimonials), oral readings (commendations, constitutional articles, enabling statutes, gospels), and collective talk (litanies, pledges, psalms), which attendees can be expected to hear or participate in during the course of the event. In more quotidian interaction, newcomers are sometimes informed that jokes are being told, “old scores being cut up,” or, in Harvey Sacks’ often cited therapy group data, “we were in an automobile discussion.” But we may also be told, “He’s just having one of his rants.” What is relevant for our current considerations is that all competent societal members recognize some genres (although not always by name), and that some members, by virtue of education, occupation, stage of life, or whatever, recognize additional varieties.

I have alluded to the long rich tradition of attention to matters of genre in criticism, literary studies, philosophy, and rhetoric. I cannot review that literature here; it is cited in writing about genre in other domains, including both discursive and “scientific” studies discussed immediately herein and in literatures referenced throughout this chapter. Similarly, there is little need to review representatives of the two principal discursive approaches to genre. The first, constituted by what may usefully be called theoretical perspectives, includes four subtypes: (a) reviews of approaches historically, followed by specifica-
tion of an author’s perspective based on that earlier work27 (such treatments range from Ferguson’s, [1996b] humanistic/linguistic/philosophical approach to the more methodologically oriented review, which can be found in the early pages of Biber [1995]; (b) explication and development of concepts, heuristics, relationships, and perhaps proto-paradigms oriented to some broader theoretical goal, such as Hymes’ (1974) “integrated theory of sociolinguistic description”; and (c) theoretical development as a by-product of substantive work in which investigators discover that available definitions and characterizations fail to identify dimensions of differentiation that emerge from the speech varieties or literatures they are examining. Recent instances here include Silverstein’s (1996) extension of Sapir’s work on Wishram myths by noting that the Wishram texts include both tales about fantastic beings engaged in both quotidian and fantastic behaviors and ordinary humans engaged in fantastic doings; Crapanzano’s (1996/1992) focus on issues of freedom and flexibility within genres, which leads, among other things, to possibilities of genre as litigable28 and to his identification of the genre of medicostructural libertinage; Haviland’s (1996) admonition that investigators must explore the connections between locally constituted genres and consciously fashioned texts and the implications of production of coherence via writing things “down”; and because I have no room to extend this list of examples, Bauman’s (1992) demonstration of multiple embeddings of reported speech, dialogue of genres, importance of who tells what, and, nicely, the notion of genre leakage. Bauman correctly observed that genre is a classificatory concept on which agreement diminishes as refinement increases; (d) expository texts that include introduction of notions of genre to familiars for purposes ranging from literary analysis to pragmatics to the history of English. Here, among many excellent introductions to linguistics or sociolinguistics, I find Traugott and Pratt’s (1980) attention to matters of genre exemplary.

The second broad variety of discursive approaches to genre in the nonhumanist realm consists of the thousands of studies of specific kinds of talk and writing that appear in journals and anthologies, monographs, and papers presented at conferences, congresses, and

---

27Biber (1995) also listed four major categories of research on register/genre: (a) synchronic descriptions of a single register, (b) diachronic descriptions tracing the evolution of a single register, (c) synchronic descriptions of the patterns of variation among multiple registers, and (d) diachronic descriptions tracing changes in the patterns of variation among multiple registers. His categorization is apt for his quantitative enterprise. It is a big pie.

28For instance, as in attempts by the heirs of Margaret Mitchell to block publication of the parody, *The Wind Done Gone*. 
annual meetings, and in fields and specialties as diverse as anthropology, education (and other “professional” fields), ethnography of communication, ethnology, gender studies, folklore, linguistic anthropology, negotiations, political science, social psychology, sociology, and those on specific languages. Quality varies widely. Readers have their own favorites.

I do not mean to imply, by my metaphorical labeling of my final category of treatments of genre as “scientific” to in any way suggest that those already mentioned are not careful, disciplined, and rigorous (those predisposed to believe otherwise should look, for example, at Friedrich [2001] on lyric epiphany). I mean rather that investigators whose work falls into this category use methods that can be sufficiently and specifically described that others can replicate their research and check their findings (although they may differently interpret those findings). Two different varieties of quantitative research can be identified here, the factor analytic textual analyses of Biber and his collaborators, sometimes called computational or corpus linguistics, and the experimental approach of the discourse psychologists.

Biber’s approach to categorical sorting of written and spoken texts of different varieties is to first determine quantitatively which texts are alike and then to identify what the texts that are most alike do pragmatically. Initially his factor analytic studies were of genre; he and his colleagues now study register, seen as a more inclusive domain. Early studies examined features of written reportage, editorials, biographies, government reports, fiction of various sorts, spoken conversations, telephone conversations, public “conversation,” and broadcasts from two major computerized corpora. Through the years, the reservoir of texts was continuously enlarged. Early studies looked at lexical variation and such features as passives, nominalizations, pronouns, contractions, and presence of the past tense; findings were displayed on single and then multidimensional plots. Later studies employed larger corpora (including corpora in other languages) and looked at more and more complex syntactic relations, and relations among syntactic relations, with an end to determining patterned co-occurrence of linguistic features in samples of text. These factor analyses convincingly demonstrate differentiation of genres by the measures employed.

One charming early study (Biber & Finegan, 1989) looked at three dimensions (informational vs. involved production, elaborated vs. situation-dependent reference, and abstract vs. nonabstract style) in fiction, essays, and personal letters over a four-century period and found not only that the genres differ on these dimensions, but also that overall characteristics of linguistic production (at least in the three genres examined) varied over time. Biber and Finegan wrote, “17th-century
texts are relatively oral; 18th-century texts become more literate in style, and later texts then gradually shift to more oral styles. By the modern period, the three genres are usually considerably more oral than their 17th-century counterparts” (pp. 498–499). Similar drift is doubtless currently underway. These techniques allow us to investigate rather than simply hypothesize about contemporary events such as language convergence, lexical diffusion, and the drifting apart of various “Englishes.”

I mention only one more approach to matters of genre. Biber and his colleagues addressed the question of how texts from recognizably or traditionally variant sorts (genres) can be differentiated through quantitative study of intratextual linguistic patterning. Discourse psychologists who also employ quantitative methods focus instead on how consumers of text comprehend what they read or hear, for example, or on how consumers use differently patterned texts to accomplish their social ends or how the ends can be identified. In a 1997 article, Goldman reviewed studies of cognitive processes by examining processing of texts. Goldman and colleagues (Bloome et al., 1999) have been interested in reconceptualizing reading as intertextual practice, with goals ranging from assessing the authority of various texts to examining texts (textbooks) as a genre, to investigation of how readers assume authority for making intertextual connections. Goldman and Bisanz (2002) sought to see how design of science texts could facilitate comprehension and incorporation of prior knowledge and to distinguish among categories of scientific texts (subgenres perhaps) based on such variant intended dominant functions as: (a) communication among scientists, (b) formal science education, and (c) popularization. They noted, further, the central importance in science socialization of learning to understand and generate the genres of one’s field.

29 There are other enchantingly complex approaches to matters of the patterned variation generally labeled genre in this chapter. Students of the standardization of English from late medieval to early modern times have found the emergence of genre at the intersection of such variables as dialect competition and spread (and decline), provincialism and cosmopolitanism, prestige and stigma, formal and informal, and, particularly, spoken, written, and printed language. Regional language variety differences interacted with regional intellectual specialization on religious or scientific topics in ways that resulted in differences in linguistic patterns being associated with emergent “genres” fostered by differently targeted audiences of readers. Once begun, further differentiation resulted from separation of intellectual networks in the capital in interaction with growing prescriptivism and concern with “proper” language. See Wright (2000) for an introduction to this suggestive literature.

30 For a quite different approach to scientific English, see Halliday’s (1988) detailed historical/comparative linguistic analysis of the register.
In a 1993 paper with Kreuz, Graesser looked at questions in a large corpus of letters written to advice columnists; as they looked for varying degrees of criticality in van der Meij’s (1987) assumptions behind sincere, information-seeking questions, their approach showed similarities to that of Biber and associates. In their 1997 article, Graesser, Millis, and Zwaan reviewed the growing literature on discourse processing and such phenomena as understanding of referring expressions, connection of statements in text, and encoding of knowledge-based inferences to identify levels of representation employed in constructing meanings in the course of reading text. In a 1998 collaboration, Graesser and associates (Graesser, Kassler, Kreuz, & McLain-Allen) looked at how reader attributes such as literary expertise, reading skill, and working memory span affect readers’ processing and response to text with assumptions that deviate from received perceptions of the world (e.g., citizens knowing exactly when the world will end or time flowing backward instead of forward).

When the several varieties of approaches to genre and other classification of variation in text are mobilized together, the resulting analytic description is richly complex, theoretically instructive, and a foundation for socially valuable practice.

**REGISTER: STILL ANOTHER DIMENSION OF VARIATION**

For situation-specific use, the British term *register* has gained acceptance (Hymes, 1974, p. 59)

“Registers,” for example, are not chosen only because a situation demands them: they may be chosen to define a situation, or to discover its definition by others (as when the choice can be taken in two different ways, depending on the relationship). (Hymes, 1974, p. 112)

The question is not what peculiarities of vocabulary, or grammar, or pronunciation, can be directly accounted for by reference to the situation. It is *which* kinds of situational factor determine which kinds of selection in the linguistic system. The notion of register is thus a form of prediction: given that we know the situation, the social context of language use, we can predict a great deal about the language that will occur, with reasonable probability of being right. The important theoretical question then is: what do we need to know about the social context in order to make such predictions? (Halliday, 1974, p. 33)

... the categories of ‘field of discourse,’ ‘mode of discourse,’ and ‘tenor of discourse’ are not themselves kinds of varieties of language. They are the backdrop, the features of the context of situation which determine the kind of language used. In other words, they determine what is often re-
ferred to as the register: that is, the types of meaning that are selected, and their expression in grammar and vocabulary. And they determine the register collectively, not piecemeal. (Halliday, 1974, p. 50; on text as defined jointly by register and cohesion, see Halliday & Hasan, 1976)

Two powerful tools of analysis and understanding available to the student of human language are the analysis of types of discourse and the analysis of how language varies depending on the occasions of its use. The former, the study of discourse types, is what is traditionally called “genre analysis.” The latter, the study of language variation by use, is referred to by some as “register analysis.” (Ferguson, 1996b, p. 167)

Finally, Biber (1995) wrote:

In my own previous studies, I have used the term genre as a general cover term, similar to my use of register in the present book. . . . In Biber (1988: 68), I describe genres as “text categorizations made on the basis of external criteria relating to author/speaker purpose” and the text categories readily distinguished by mature speakers of a language. . . . These categories are defined primarily on the basis of external format. . . . In practical terms, these categories are adopted because of their widespread use in computerized language corpora. The use of the term register in the present book corresponds closely to genre in these earlier studies.

In contrast, . . . text types are defined such that the texts within each type are maximally similar with respect to their linguistic characteristics (lexical, morphological, and syntactic), while the types are maximally distinct with respect to their linguistic characteristics. After the types are identified on formal grounds, they can be interpreted functionally in terms of the purpose, production circumstances, and other situational characteristics shared by the texts in each. (pp. 9–10)

Clearly there is no consensus among linguists and sociolinguists about criterial differences among various samples of language text, such that some samples can be labeled as representing different genres and others as representing different registers. Yet it is a common belief among linguists and sociolinguists that it is difficult to change one’s language or dialect, but that most speakers in every language can quite comfortably switch among registers. I find the genre–register distinction necessary to account for the fact that, although there is a kind of language in use that we agree can be labeled as a genre of professional/scholarly articles, my editors’ complaint that some of my writing in the article was “a bit informal” shows us that there is another “kind” of language difference—that kind of difference has been labeled register.
In my view (and usage), a register is a language variety associated with social use, such as formal or informal, sacred or secular, occupational or familial, male or female, young or old (or, more precisely, infant, child, adolescent, adult, elder), and so on. Although some registers are clearly learned (e.g., “linguist talk” or “police officer talk”), others seem to be acquired simply by living in a speech community. How many of us have heard any adult member of a speech community who, confronted with a baby, does not make a shift in the direction of baby talk whatever his or her feelings about babies? Baby talk is, of course, a language variety taught to babies. Similarly, how many of us in interaction with dependent elderly can successfully talk with them as “grown-ups”?

Although the term register, like those of context and genre, is polysemic, considerable coherence has been given to the concept by work of Michael Halliday (see especially Halliday and Hasan, 1976). As the prior examples suggest, situation is the principal determinant of register selection. In his work, however, Halliday adapted three general notions of field, mode, and tenor to particularize situation. Halliday’s own words provide the most economical gloss:

The FIELD is the total event, in which the text is functioning, together with the purposive activity of the speaker or writer; it thus includes the subject-matter as one element in it. The MODE is the function of the text in the event, including therefore both the channel taken by the language—spoken or written, extempore or prepared—and its genre, or rhetorical mode, as narrative, didactic, persuasive, ‘phatic communication’ and so on. The TENOR refers to the type of role interaction, the set of relevant social relations, permanent and temporary, among the participants involved. Field, mode, and tenor collectively define the context of situation of a text. (Halliday & Hasan, 1976, p. 22)

He continued by observing that, “the linguistic features which are typically associated with a configuration of situational features—with particular values of the field, mode, and tenor—constitute a REGISTER” (Halliday & Hasan, 1976, p. 22; italics added).

---

31 Some scholars use style and register almost interchangeably. For one thoughtful attempt to distinguish style and register by suggesting that register is what is found at the intersection of register and situation, see Peng (1987). Others see style as individual and idiosyncratic. As has so often been the case, Ferguson’s (1977) distinction clarifies:

On the one hand, register variation is distinct from regional and social dialect differentiation and, on the other hand, it is distinct from idiosyncratic and stylistic variation. A register in a given language and given speech community is defined by the uses for which it is appropriate and by a set of structural features which differentiate it from . . . other registers. (p. 212)
Again different sorts of criteria for identification of registers may lead to confusion in classification. For example, a formal–informal distinction can occur within an occupational register, and there may be variants within formal and informal. Ferguson (1983) asserted that “sports announcer” register was recognizable even when actual words were obscured; he believed that baseball and football announcers could be distinguished when text was not available. There are presumably formal and informal registers within the medical register or in that of the military. Yet the informal registers of regimental headquarters, barracks and foxholes, and, in the latter, in combat or not, are quite different—just as is the medical register in the classroom, on rounds, in conversations with patients and their families, in the operating theater, and in press conferences. In the medical case, consider in which situations different parties, including physicians, might choose to reference abdomens, bellies, guts, insides, stomachs, tummies, and tum-tums. Exactly what subvariety of a register is employed and how utterances (or written texts) are to be interpreted depends, of course, on variation in field, tenor, and mode.

Baby-Talk (BT) Register: A Special Case

Because of Ferguson’s (1964, 1971) interest, we know quite a bit about this widely distributed cross-language register. It appears to differ from other registers in that, although acquisition of all language varieties involves some explicit instruction, BT is both recipient-designed and often explicit (e.g., “Can you say doggie?”). Although registers for talking to other dependents, such as the sick and the old, are recipient-designed, I have not heard the old or sick being given explicit instruction in their use. Coupland, Coupland, and Giles (1991) referred to what they call overaccommodation, but not to attempts to train elders to use BT-like registers.

A moment’s reflection reminds the reader of some characteristic features of BT: raised pitch, high redundancy, putative simplification often through diminutivization (although neither bow-wow nor doggie is actually simpler than dog), special lexical items for food, euphemisms for bodily functions, games, and so on. BT in other languages manifests similar features, although features are by no means universal. Again, although adults comfortably switch from one register to another, small children cannot suddenly talk like adults. That said, competence in BT does not seem to erode from nonuse as happens with both mother-tongue languages and other registers. Older children sometimes employ BT to elicit favored treatment, tease younger children for acting like babies, or, at some juncture, talk to lovers.
The Instability of the Concrete: Change in Age Registers

There do not appear to be many (if any) diachronic studies of BT, either in English or any other language. I have the strong impression, however, that with the expected exception of lexical items, American BT registers remain much the same today as they were 50 or even 100 years ago. In contrast, I would hazard that there are greater changes in the registers associated with age groups such as preadolescence, adolescence, young adulthood, mature adulthood, and elder status both for cohorts as they move through their years and for those in the same age group over historical time. Having lived through shared experiences such as the World Wars, the Great Depression, the cold war and its associated fears, and technological, sexual, and other “revolutions” or not results in important differences in both language resources and worldviews.

In one of his studies, an investigation of selecting senses from among those already known as contrasted to “creating” senses, Clark (1992) employed as a sample sentence, “After Joe listened to the tape of the interview, he did a Richard Nixon to a portion of it.” To the Stanford students who participated in Clark’s experiment, the events of the end of Nixon’s presidency were still freshly in mind; it is not clear how long their peers will remember the clue. It seems doubtful that many undergraduates today could interpret the reference. How much more uncertain must be readings of a statement, “My relations with my father are/were bad/complicated/good” depending on age and gender of interlocutors. The biological relation remains constant; the social relationship may change massively. Just what would such an utterance from a contemporary grandfather to his grandson or granddaughter mean? In the first part of the last century, middle-class fathers were in charge; at the end of the century, they often were “equal partners.” What additional uncertainties are added by changing norms about talking about family matters? By differences in political and religious ideologies? Are residual generational differences enough to outweigh

32 Change is slow even in the lexicon. Note the persistence of “choo-choo.”
33 The following colloquy overheard on an 86th crosstown bus was reported in the “Dear Diary” column of the New York Times (June 24, 2001):

Grandmother: “So what are you doing after school tomorrow?”
Grandson, about 10 years old: “I think I’ll go home.”
Grandmother: “Why don’t you arrange a play date?”
Grandson: “We don’t call them play dates any more.”
Grandmother: “What do you call them?”
Grandson: “We call them ‘do something.’ ”
Plus ça change.
other considerations? How closely does anything that we hear or read match what its producers thought they were conveying or hoped to? What are tests for finding out to what extent it matters?

**Borderline Cases: Registers or Genres?**

My confidence that there is a difference between genre and register notwithstanding, exactly where the line between the two sorts of language varieties is located is not always self-evident. Consider, for example, the following sociolinguistic phenomena.

**INSTRUMENTALITIES and Speech Acts: Importuning Letters.** In the prior discussion of sorting out differences in behaviors related to scold, I showed that this could be done through variations of contrastive feature or componential analysis. That is, we can see that a set of behaviors varies from related behaviors by virtue of the possession or nonpossession of one of many attributes. INSTRUMENTALITIES are verbal behaviors employed to get target others to (a) act in ways they would otherwise not, or (b) not act in ways they otherwise would (Grimshaw, 1980, 1981, 1989). The INSTRUMENTALITIES are sorted on axes of power and affectual relations between interactants and the desirability and costs of an outcome—that is, by elements of the situation. Not surprisingly, McLean's (1998) study of favor seeking in the Renaissance found that much effort in patronage seeking was applied to attempts to construct frames of relations of power and affect and assessments of utility most likely to enhance prospects for outcomes favorable to initiators. We are rich in archives of letters from at least the past five or six centuries; it should not be difficult to discover some of the sorts of changes, if any, that have occurred in practices of using language to get others to do things.

**Distancing Genres.** Another variety of discourse, actually more often written than spoken but not always so, is not a recognizable genre or dialect, but also not clearly a register (or any other commonly identified language variety). I have in mind here an array of distancing modes of discourse nonexhaustively including: argot, bureaucratese, cant, circumlocution, computerese, double entendre, double meaning, doublespeak, double talk, dysphemism, euphemism, gobbledygook, "hidden transcripts" (Scott, 1990; cf. Abrahams, 1972; Kochman, 1981, on "shucking" and "jiving"), jargon, legalese, mumbo jumbo, Newspeak, private language, riddles, and so on. Discourses may be overt or covert, published, intended for overhearing or other discovery, encrypted in such a way that cryptanalysis can easily reveal messages in-
tended to be seen as “secret.” Texts may be intended to variously: conceal, confuse, deceive, distort, facilitate technical communication (attractive nuisance, parallel processing, ischemic cardiomyopathy) among, frighten, generate solidarity among, magically coerce, mislead, mystify (collateral damage, friendly fire, incontinent ordinance, pacification centers, surgical strikes), persuade, predispose, prejudice, propagandize, or sway those who are exposed to them—or perhaps to simply make them suspicious. Analysts, like interactants, need to be constantly alert to the fact that discourses are not always what they appear to be.

“Gossip,” “Schmoozing,” and “Keeping up Professionally.” Another set of discourses that often contain subtexts, hidden agendas, veiled criticisms, sub rosa evaluations, and so on includes the discourse types in this section’s heading—along with scandal, Schadenfreude, and certain expressions of false sympathy. Gluckman (1963) is a classic with a rich literature review (see also Besnier, 1996; M. Goodwin, 1990; Watson-Gegeo & White, 1990). In his thoughtful introduction to this topic, Gluckman (1963) wrote of gossip among professionals:

The more exclusive the group, the greater will be the amount of gossip in it. There are three forms of social group to test this hypothesis. The one is the professional group, like lawyers or anthropologists, whose gossip is built into technical discussion so tightly that the outsider cannot always detect the slight personal knockdown which is concealed in a technical recital, or the technical sneer which is contained in a personal gibe. This is, therefore, the most irritating kind of group to crash into, because one has no clue to the undercurrents, no apparatus for taking soundings. And this is why old practitioners of a subject can so easily put a comparative newcomer into his place, can make him feel a neophyte. They have only to hint in a technical argument at some personal fact about the person who advanced the theory discussed, to make the eager young student feel how callow he is. Again, the more highly organized the profession, the more effective is the role of gossip here. (p. 309; this is, of course, what Cicourel was saying about contexts in the dissertation defense materials)

Different discourses of these sorts vary on such dimensions as presence or absence of persons or institutions on whom discourse is focused, degrees of maliciousness, purveyors’ ends, the extent to which simple denial or documented refutation can repair damaged reputa-

34To get an idea of the increase in sophistication in cryptography and cryptanalysis over the past 40 years or so (particularly since the advent of large-scale use of computers), compare Kahn (1967) and Singh (1999).
tions, and so on. Are the differences, and they are familiar to most readers of this chapter, in genre, register, or something else? Again caution should be the watchword for audiences for, potential participants in, and analysts of gossip discourses.

We all have rich repertoires of registers. Register must be taken into account in discourse analysis. Linguistic and pragmatic specification of register as a language variety remains an elusive goal.

**Code, Diglossia, Markedness, Register, Repertoire, Switching, Variety.** Although there are no documented instances of what might be called “monoregisterial” speech communities, the diversity of community repertoires and, within communities, of individual repertoires varies widely. For example, consider the varietal richness that Fox (1974) described for some 100,000 members of the Rotinese speech community. Almost every member of this island population spoke depeä lote (Rotinese language) in one or another of about 20 distinguishable regional dialects. All of the dialects had two major registers—dedeäk or “talk”—which included recognizable variants of conversation, tales, true tales (history), riddles, dream interpretations, and a stylized form of mockery; and bini, a highly stylized variety of compositions in parallel verse about past events produced by lords and elders in commentary on litigation or in the initial stages of bridewealth negotiations. Some individuals know neighboring related but apparently mutually unintelligible languages such as Ndaonese and Helong. Three decades ago, increasing numbers were able to use Malay in one or another of its three varieties of biblical Malay, basa kupang (a language of the marketplace apparently picked up in nearby Timor), and modern basa Indonesia; it is probably safe to guess that most Rotinese are now fluent in the latter. When Fox was doing his fieldwork, the Rotinese enjoyed chatting about dialect differences and were deeply concerned to avoid mixing speech inappropriately. As Fox noted, however, the coexistence of varieties can generate new complexities and “mixing” takes many meanings. He wrote:

The three varieties of Malay are, however, sufficiently similar that the furtherance of one has consequences for the others. On conclusion of the national literacy campaign in the late 1950s, when the island was certified as literate in Indonesian, there was mass conversion to Christianity. Malay had served as a check to conversion and when, by decree, all Rotinese became Malay speakers this obstacle was removed. Significantly Biblical Malay is itself as a formal ritual language indispensable for Christian rituals. The parallelism that pervades the Old Testament accords well with Rotinese ideas of a ritual language. A church service consists of reading from the Bible with translations in Rotinese, Malay
songs, and long sermons, often in Malay, with long paraphrases in Rotinese, or interspersed Malay and Rotinese, or even a cacophony of two simultaneous sermons, with one preacher speaking Malay, the other translating in Rotinese. (pp. 69–70)\(^3\)

Readers of Geertz (1960) will recall that further west in the Indonesian archipelago in Java a situation existed in which speakers of clearly identifiable varieties were distributed by such dimensions as class, education, and urban–rural residence, but in which “the choice of linguistic forms as well as speech style is in every case partly determined by the relative status (or familiarity) of the conversers” (p. 248). In other words, *dialect choice* is determined by social setting, topic, interactional history of participants, and presence or absence of an audience or audiences (i.e., it is registerial).

West again and north, we find in India a billion people speaking 14 official languages and dozens of other local and world languages;\(^3\) probably hundreds of regional and class dialects; occupational, age, gender, sacred, and secular registers; and individual repertoires ranging from control of a single dialect with a modest array of registers to virtuoso polylingualism with fluency in many registers within languages. Here *varietal choice* is determined partly by register-like considerations, but also by repertoires available to participating interactants. Farther north again, about half of all Iranians are Farsi/Persian mother-tongue speakers, and the remainder are unevenly distributed among 60 or so additional languages. Speakers of many of the other mother-tongues control Persian as a second language. What is interesting here is not what varietal choices are made, but rather the ways in which apparently minor differences in lexicon, morphology, phonology, and syntax *within* Persian are employed to communicate intricate differences in social meaning and, particularly, to attempt to accomplish different social ends (Beeman, 1986). In still another manifestation of intraspeech community diversity, Blom and Gumperz (1972) reported on another situation still farther north, in Norway, in which a single linguistic system is considered by speech community members to contain two distinct codes or varieties. Blom and Gumperz examined choice of variety across settings, situations, and events and introduced their now familiar and powerfully useful distinction between *situational switching* and *metaphorical switching*.

\(^3\)Some readers will be familiar with the spectacle of Jimmy Swaggart collaboratively preaching with a Spanish-speaking associate who almost simultaneously translates Swaggart’s version of the word to huge audiences in Latin American stadia.

\(^3\)Seventeen of the 100 languages with the largest number of mother-tongue speakers are indigenous to India.
Ferguson’s (1996a) *diglossia* referred, of course, to instances where different codes (Arabic and local vernaculars in largely Muslim countries) are routinely associated with different arenas of language in use.

Varietal diversity within societies is determined by many considerations. Among the more important are external and internal geographical isolation and accessibility, population size, composition (homogeneity or heterogeneity) and distribution, social structural characteristics such as division of labor and stratification, intervariety contacts, language change, language ideology and language policy, and the presence or absence of writing and printing. Selection of one or another variety available within the repertoire of a speech community, whether for learning or for use in social interaction, occurs for a number of different reasons including, nonexhaustively, identity construction and maintenance, solidarity, self-improvement, status acquisition, and devoutness. Haeri’s (1997; see also appended commentary) observation that more is involved in the value of different codes than which variety in a society is “official” language and whether the official language is used by a dominant class is indisputable. He wrote, “... power and language are related to each other, but the relation is not always linear or free of contradictions that are a product of the tensions between various sources of power.” So it is in all situations where there are different codes, varieties, languages, dialects, registers, or different ways of talking, however differentiated, classified, and labeled.

**A Final Dimension of Registral Difference: Written and Other Media Registers.** My theoretically unmotivated sampling of instances of intrasocietal diversity in ways of talk (and writing) is intended to do no more than remind readers that there is diversity and complexity and to highlight the dangers of attempting any sort of discourse analysis without explicit attention to what is being done pragmatically when code switching occurs. The following homely example may illustrate my point.

Toward the end of a summer about 30 years ago, I received a letter from the president of a professional organization (of which I was not a member) inviting/requesting me to address the group’s annual meeting. Not long afterward, I wrote a letter declining the invitation primarily on the grounds of prior obligations. A few weeks later, I met a colleague from my own institution at a faculty committee meeting; at the end of the meeting, he said he wanted to talk with me. I said, “Fine,

37There are other considerations. In her work on Africa, Myers-Scotton (1993) foregrounded matters of *markedness* and *performance*.

38Description of this incident and some of the following discussion appeared in Grimshaw (1973).
how about right now?” He insisted, however, that he wanted to talk with me by telephone. A few days later, he called and asked me to reconsider the invitation on the joint grounds that (a) I had things to say that were important for members of the organization, and (b) I had an implicit moral obligation because of the organization’s goals. I expressed doubts about the first ground, but said I would reconsider if I was still wanted after the colleague looked at some recent work in the area of shared interest; I sent that work to him by campus mail. A few weeks later, I received a letter (again in campus mail) from my local colleague reiterating the earlier arguments and stating that the papers were directly on topic. A few days later, he called again and I agreed to participate in the program. Another few days later, I had a phone call from the president to confirm my agreement. Some days later, I received another letter from the president with formal confirmation of dates and other arrangements. I gave the paper. I hope my audience found my message (which was partly about how I ended up talking to the organization) useful.

A moment’s reflection will allow readers to recover rough guidelines for requesting behavior for a broad spectrum of middle-class Americans. Generally, middle-class professionals find it easier to refuse written requests than those made on the telephone and the latter easier to turn down than those made in face-to-face interaction. But other considerations such as (a) embarrassment about making the request, or (b) concern about exposure to kinesic or other cues that the target sees the request as too demanding can reverse this order. Intermediaries may further complicate ordering. Requests may be made ambiguous to allow a target to decline without appearing to rebuff. Note further that condolences may be delivered in person or in mail (perhaps optimally in both media), under specifically defined circumstances by telephone, but not yet sent by e-mail (this may well change in the near future).

Humans may have had speech as long as 2 million years; they certainly have done so for several hundreds of thousands of years. It is reasonable to expect that differentiation of registers in spoken language accelerated as our ancestors became more human. With the appearance of writing about 5,000 ago, however, a critical new differentiating feature became available. In the five millennia since, a number of sets of developments have been taking place with events sometimes more or less driven by their own momentum and at others resulting at

---

39Lawler (2001) reported that writing was in use in possibly three independent locations at least five millennia ago and that dates of precursors of writing systems are now sometimes dated from 4,000 to 9,000 years ago.
least in part from developments in other evolutionary streams. First, writing has evolved both as new writing technologies have developed (e.g., pens and parchment or brushes and paper replacing styluses and clay) and as efficiencies of writing have been invented (e.g., pictograms, logosyllabaries, syllabaries, alphabets, cursive writing, and new orthographies, or punctuation\textsuperscript{40}). Second, from the beginning, technologies became increasingly sophisticated and efficient, permitting both ever greater range and speed of distant communication and, through the accumulation of written records, expansion of individual and collective memory. In the last half of the last millennium (actually from 1455, when Gutenberg printed his first bible\textsuperscript{41}), the pace of technological innovation accelerated with the appearance of inventions ranging from the semaphore, telegraph, wireless, and transatlantic cable to typewriters, carbon paper, mimeograph, ditto, and multilith to hot type linotype presses. In the last half of the 20th century, the pace continued to increase with the appearance of computers, the Internet, e-mail, computer typesetting, photocopiers, fax, word processors, ink jet and then laser printers, scanners, and combinations of new hardware and software technologies in TalkBank and other such programs.\textsuperscript{42} A third set of developments has to do with expansion, or at least change, in human cognitive capacities, which may be associated with literacy (Goody, 1987). Literates are able to communicate with themselves with reminder notes, datebooks, and journals and employ lists and other memory aids learned for written communication but adaptable to oral or internal communication to store

\textsuperscript{40}On the history of writing, see Daniels and Bright (1996) and Hockett’s (1997) helpful review of that volume. On an inevitable evolution of writing systems to alphabets, see Havelock (1982). The later introduction of punctuation permitted incorporation of dimensions of meaning not available in early writing (e.g., attribution, disjuncture, distancing, emphasis, prosody, segmentation, semantic shading, subordination, tempo, and topic shift). On the revolutionary impact of punctuation see, particularly, Morrison (1987). Like Goody, Morrison uses his topic as a platform from which to examine complex issues of the interaction of textual organization with verification and argument in institutionalization of knowledge.

There is no space in this short survey to review the growing literature specifically addressing written registers and writing as register. For an initial sampling, see articles in Biber and Finegan (1994) and Ghadessy (1988, 1993).

\textsuperscript{41}Printing, also with movable type (characters), was independently invented in China at about the same time.

\textsuperscript{42}Daniels (1996), in his useful short history of “analog and digital writing,” commented, “it is clear that the alphabet drives the technology and the technology serves the alphabet” (p. 889).

It will be remembered that other new technologies were and are being developed to facilitate distant communication and expand individual and collective memory. Note only radio, telephones (and today’s ubiquitous cell variety), cinema, TV, and digital imagery.
information both in written records and memories. What, we might ask, are the cognitive effects of copying something by writing with a pen or pencil or on a typewriter as contrasted to making a photocopy of some sort?\textsuperscript{43} What are the long-term consequences for cognitive organization of practices in writing that originated in concerns about efficiency and economy, when messages were sent via semaphore or Morse, for example, but continue in such economies as the compressed language of article abstracts, encyclopedia entries, instructions, lists, program notes, summaries, synopses,\textsuperscript{44} and so on (see, in addition to Goody, 1987; Sinclair, 1988; Thorne, 1988)? The interaction among technological advance, considerations of economies of effort, and an emergent register are perhaps most visible currently in e-mail, where spelling changes, inventions of new modes of punctuation and paralinguistic signaling, changes in norms of appropriateness, code words of various sorts, and other differences from “traditional” writing appear with each new cohort of computer users; not many readers of this chapter have been able to keep up.

\textit{Ceteris paribus}, consider the differential effects in terms of certainty/nonambiguity, documentability, emphaticness, and perlocutionary effect(s) of spoken contrasted to written utterances in domains of advertising/promotion, alienation (of property), bureaucratization, caution enhancement/hyperbole dampening, clarification, commitment, comparison, confirmation, control, documentation, making explicit, rationalization (in the Weberian sense), reflection, remembrance/recall, standardization, streamlining/redundancy reduction, and systematization. In each of these domains, and in the societal institutions in which they are severally and variously embedded, the same words can do different work when spoken or written: think, again \textit{ceteris paribus}, of the three English words “I love you.” I think that most native speakers of English, at least, would agree that, whatever the level of the utterer’s actual commitment, for an adult to put those words in writing in a communication to a potential partner is a more “serious” matter than saying them. Relatedly, although some aging males may still feel that their handshake should suffice as a guarantee, contemporaries increasingly want signatures. The importance of writing in documenting commitments and the truth or falsity of claims is evidenced by the proliferation and use of certificates, contracts, deeds, depositions, diplomas, proofs of membership, mani-

\textsuperscript{43}None of this should be read as claiming that the cognitive capacities of literates are\textit{superior}. Think only of what most literate people fail to see while walking outdoors.

\textsuperscript{44}I particularly like those of movies in TV programs. There are, of course, economies in speech ranging from argots, cants, and jargons to straightforward ellipsis among familiars.
fests, registrations (birth, death, marriage, product purchase), testaments, titles, transcripts, and collections of written educational, judicial, medical, and military interactions in dockets, dossiers, files, folders, record-jackets, and, ultimately, archives in contemporary life in modern literate societies.

The fourth set of developments is generated by the fact that changes of the sort just identified have occurred not just in human cognitive processing and interpersonal relations, but in the organization of society and society’s constituent institutions. That theme is the basis for Goody’s (1986) pithy, *The Logic of Writing and the Organization of Society*. In this book, Goody asked how religion, the economy, the state and bureaucracy, and the law differ in societies with and without writing. In what follows, I can do no more than offer a sampling of some of his observations without repeating the particulars of his argument or documentation:

*Religion*: What difference if religion’s “word” is spoken or written? (a) differences in degrees of commitment, (b) only religions with written literatures are religions of conversion (note that a first move of Christian missionaries is to make the Bible available in the local language. I return briefly to this matter at the end of this section), (c) writing permits abstraction, generalization, and universalizing of the holy word, (d) differential access to sacred writings by the literate gives them control over nonreaders (cf. Schieffelin 1996), and (e) sacred writings, commentaries, and even secular records can serve as a basis for dissent;

---

45 In recent years, computer capabilities have permitted the appearance of individualized birth announcements. Innovation in weddings frequently involves reading of non-traditional text prepared by the couple marrying. At the same time, there are trends toward standardization or disappearance of genres of writing, such as thank-you notes, bread-and-butter letters, and, particularly, notes and letters of condolence, which were widespread in this country in the first half of the last century. Sympathy cards are now commonplace in many countries; “*thanks for sympathy*” cards are spreading. There are sympathy cards sold in stores; they can be constructed on the Internet and sent via e-mail. There are cards for grieving pet owners and for people whose loved ones have taken their own lives. Decorated condolence telegrams thrive in Japan.

The interaction of literacy and bereavement management is a complex one. What is going on sociologically when an acquaintance (aware of being seen) crosses the street rather than engage in interaction with a bereaved person, but then writes a note expressing sympathy and explaining an inability to manage emotionally difficult situations? What is going on when we painfully construct a condolence letter on our word processor and then copy it by hand before sending it? What is more appropriate, a word-processed letter or a sympathy card? Appropriate from whom?
Economy: Writing is important in exchange and management; concepts of money, private property, account, shareholding, and credit are facilitated by writing and written records. A central claim here is that writing and written records are a prerequisite in any modern sense of science;

State and bureaucracy: The obvious question here is how do regimes with and without writing differ? There are again a number of interesting directions for exploration: (a) distant communication, (b) archiving of information, (c) numbers permit control of time, (d) copies, (e) treaties, (f) explicitness, (g) participation and dissent, and (h) assignment of responsibility;

Law: Law in societies with literacy is written, customary law was not. Litigation occurs without writing (Frake, 1980/1969); writing allows systematic appeals, facilitates objectification of outcomes, and furthers the development of professionalization of law practitioners.

No more than a cursory skimming of some of the topics in Goody’s rich book, coupled with reflection on familiar historical changes and some retrospective consideration of invention and both planned and serendipitous usages of new communicative technologies in our own lifetimes, will remind readers of the profound impact of writing, literacy, and other developments in changing both society and its institutions and the nature and quality of interaction among members of society. I close this section with a brief review of a case that represents how critical literacy and, through literacy, access to foundational literatures or important secrets could be.

Making the Bible Available in Vernacular, Secrecy, Censorship. Although torture and execution as instruments of state policy continue in our times, it seems unlikely that it would occur to most contemporary consumers of written text that sacred and secular leaders in 16th-century England would have employed both as part of a campaign to prevent the publication of an English translation of the Bible (Bobrick, 2001; McGrath, 2001). That such efforts at censorship occurred in England and elsewhere, particularly in the Germany of Luther’s emergence as a figure of major influence, is testimony to awareness on the part of sophisticated elites of the revolutionary power of subversive written texts. Such efforts at suppression stem from the same concerns that generated laws against teaching slaves in the antebellum South in the United States to read or forbade people in territories controlled by authoritarian states to listen to “free” radio broadcasts or to read leaflets.
dropped from the sky. Similar motivations historically have discouraged female education or lead today, even in democratic states, to attempts to keep secret activities of governing elites.

Those in Britain who supported access to sacred writing through its translation into English believed both that laity able to read the Bible could participate in informed decisions about their own routes to salvation and that restricting access to church functionaries literate in Latin supported a monopoly quite inappropriate in matters of the spirit. Such a perspective was not new in the island kingdom. Lerer (1991) wrote:

Writing to his bishops at the close of the ninth century, Alfred the Great lamented the decline of learning in his country and proposed a program of official translations and scholarly study which he hoped would set his England back on the path of intellectual inquiry. . . . Alfred equated a knowledge of texts with the right governance of self and state. (p. 1)

Even before Alfred, Bede and others had translated parts of the Bible into the vernaculars then current in Britain. Because English did become current, according to Lerer, its widespread use in administration and literature was well established some three or four centuries earlier than similar use of vernaculars on the continent. Yet disputes about the use of English in Bible translation and political discussion available to the literate population (Zaret, 1985, 2000) continued vigorously as the language evolved. Moreover, although “the liberty of unlicensed printing”46 is today accepted in principle throughout most of the English-speaking world, arguments about the meaning of what happened in the struggle for a Bible in English still appear today on thousands of Web sites on the Internet. Some readers recall frustrations in the mid-20th century with obscurities of the gobbledegook of government bureaucracies and euphemistic “nuclear language” (Cohn, 1987) or talk about the Vietnam War; today governments press for legislation to control “leaks.”

When considering evolving languages, changing social contexts, different views of appropriate goals and uses of literacy, different ideologies about salvation, and different technologies of communication, it is not surprising that there are differences in genres/registers.

46Published versions of John Milton’s (1644) Areopagitica (A Speech for the Liberty of Unlicensed Printing to the Parliament of England) abound. His concern and involvement was in no way unique; strong counterviews were held by other intellectuals of the period. Compare, for example, the views of John Dryden and John Locke as cited in Bobrick (2001).
There are rich literatures on methodological matters in discourse studies. There are handbooks on data collection, transcription, and analysis and further publications on interpretations, schemas, theories of talk, and theories of interaction. Although I have thoughts on many of these matters, I use my remaining space to develop implications of assumptions about "same" and "different" meanings, "same" utterances with different goals and different utterances with same outcomes, the representativeness of populations whose discourse or processing of discourse is chosen for analysis, and the typicality of texts/discourse events chosen for analysis.

Language as Social Resource: Some Dangers in Interpretation

I suspect that most students of language are continually surprised by some new evidence of the versatility, subtlety, and potency of language as a social resource. This richness becomes problematic hermeneutically, however, when multiple ends (some quite possibly not visible to some hearers/readers) can be pursued with a single utterance or other bit of text, or when same ends can be pursued by different utterances or bits of texts. Complexities are compounded by different subcultural norms and/or individual styles. Within cultural groups, "same" behaviors may have different meanings when produced by different individuals (e.g., silence may signal tactful withdrawal from stressful interaction, but it can also signal smoldering anger preceding violence). There are similar differences across cultural/linguistic boundaries. Declination of second helpings may be polite until some adequate number of multiple invitations have been proffered in one setting while continued invitations may be read as intrusive elsewhere.

"Different" Goals for "Same" Utterances: "Different" Utterances for "Same" Goals. One variety of spoken discourse that is easy to collect in public places in much of the world is controlling talk directed by older siblings to younger. Such talk is often ostensibly intended to keep the younger sibling from doing something that might be injurious to him or herself, others, or property. Directives or instructions may take the form of imperatives or of positionals or personals in Bernstein’s (1975) sense. Thus, “Stop!”, “Stop before you hurt yourself!”, or “I told you to stop!” It is obvious to those who witness such displays, particularly if we can also see the risk, that it may indeed be the case that the older sibling wants to protect the younger from some negative outcome. Yet there may be other ends at varying levels of awareness on the
part of the issuer of the warning. There are further contingencies. Suppose that 10-year-old Anne’s younger brother, Allen, is playing moderately boisterously with a piece of family equipment that could possibly be broken. Consider the following nonexhaustive listing of possible ends and just a few of many contexting conditions.

First, why might she tell him to stop: (a) to protect the boy, the equipment, or both from damage or injury; (b) to establish herself in a position of authority in her relation with her brother; (c) to display herself in the view of onlookers as in authority; or (d) to display herself in the view of onlookers as being responsibly considerate? Among the for-all-practical-purposes endless list of contextual contingencies, note: (a) how much actual danger of damage or injury exists? (b) how likely is it that Allen will stop without being told to? (c) who is in the audience? (d) what is the current state of affectual relations between Anne and Allen? Perhaps somewhat less immediately obvious, except for parents and siblings, (e) how many times has Anne told Allen before? and (f) how has Allen responded? It seems likely that Anne and Allen and others present will differently interpret the meaning of the discourse that occurs. Experience tells us that parents may independently discover other meanings. How then can anyone talk about the meaning?47

Representativeness of Populations; Typicality of Discourse. Having demonstrated once again how complex interaction is, even between two squabbling siblings, I can now acknowledge the typicality of such interactions and discourse in American society and probably human society generally. The moves made by brother and sister are repeated millions of times everyday; we need to discover how to separate the hermeneutically elusive from the statistically countable and determine what sorts of generalizations can be made about what gets accomplished in talk. Similarly, although the doctoral dissertation defense examined by the Multiple Analysis Project and the therapy session investigated by Labov and Fanshel (1977) are each individual instances of a phenomenon that happens thousands of times every year, there are ways in which they can be characterized as “typical” instances. The question should be asked, however, whether “typical” cases are somehow “representative.” An event can be typical in that it follows the same stages and involves the same numbers of people as a statistical average

47 Comprehensive Discourse Analysis (Labov & Fanshel, 1977; Grimshaw, 1989, 1994) brings a wide range of data to bear on issues of what is said, intended, and done by an utterance or longer stretch of discourse. A considerable number of interpretation questions remain unresolved.
of all such events. However, if a therapy session ended with the slaying of the patient by the therapist or a defense when CIA or FBI agents burst into the room and arrest a committee member, could we still claim that the event was typical?48

Less exotically, how does a particular group get chosen for investigation? What makes it representative? What makes it comparable with other groups or categories? How are comparison groups chosen? If the kind of attributes noted earlier affect perceptions and attitudes, can the composition of even a group of sophomore psychology student volunteers influence experimental outcomes? Is there a variable likelihood that college undergraduate classes will contain non-native speakers? That certain regional variants will be under- or overrepresented? Do college students “represent” any population other than college students? Is gender the only discriminating variable among students? Do differences in parental social class determine/influence children’s behavior? Can groups be made comparable by matching aggregate statistics (means of age, income, education, etc.)? Is it possible that there are sociolinguistic or other behaviors that occur equivalently across groups/categories? In my view, the answer to this last question is “yes” because the sociological variables (relations of power and affect, and considerations of utility) are everywhere relevant for interaction.

The issue of generalizability and its associated problems of validity and (or perhaps versus) reliability has been discussed by Labov (1972a, 1972b) under the rubric of the “Saussurian paradox” and by Meehl (1954) as that of “clinical versus statistical prediction.” The “Saussurian paradox,” according to Labov, is: what is shared by speakers of a language can be learned by studying the language of a single individual, but that individual variations in use can be discovered only by studying talk in social contexts—that is, conversational discourse. The clinical-statistical dilemma (here I paraphrase Meehl), in contrast, is that clinical studies (of which hermeneutic micro-analyses are one type) can generate valid understanding, in depth, of individual cases, but do not permit generalization of that understanding, whereas statistical studies generate reliable and valid findings on relationships that are generalizable, but that may not be understood and are often not profound. Life is complex. So, too, is the task of discourse analysis.

48 One reader comments that such instances might nonetheless be typical of a particular subclass of events. That defenses can produce surprises is known to most academics: some instances are noted in Grimshaw (1989).
ACKNOWLEDGMENTS

The first draft of this chapter was written and submitted while I was trying to shake off the effects of medication for open-heart surgery. Many months later, when I received editorial commentary and suggestions, some of what I putatively had written was unfamiliar. My usual practice with major writing responsibilities has been to seek commentary and correction; I probably intended to ask for such collegial help in this instance. I have found no correspondence and simply do not remember. My apologies to all who may have helped me and had their advice apparently ignored. My apologies to disappointed readers who look in vain for neglected topics or literatures. I am happy to accept the responsibility for the chapter’s shortcomings or, if there is something you like, for its virtues. I am deeply indebted to Catherine Walsh for perseverance and accomplishment in dealing with often tedious editorial problems.

REFERENCES

2. CONTEXT, GENRE, AND DISCOURSE


Cicourel, A. V. (1981). Notes on the integration of micro- and macro-levels of analysis. In K. Knorr-Cetina & A. V. Cicourel (Eds.), *Advances in social theory and methodol-


A. C. Graesser (Eds.), The psychology of science text comprehension. Mahwah, NJ: Lawrence Erlbaum Associates.


McGrath, A. (2001). *In the beginning: The story of the King James Bible and how it changed a nation, a language, and a culture*. New York: Doubleday.


Comprehending text is part of the daily routine for most individuals 6 years of age and older. We read texts for a variety of reasons. We want to be informed about the state of the world, learn about new domains, escape into fictional worlds, and perform certain actions (e.g., fill out a tax form). The skill to accomplish all these things, based on perceptual processes by which our eyes fixate black marks on white and then jump to the next set of marks, is surely one of the most remarkable accomplishments of our species. Text comprehension researchers are slowly but steadily uncovering the highly complex set of cognitive mechanisms that underlie the skill to comprehend text. In doing so, they have developed an impressive range of theories and tools to test them. This chapter provides an overview of this work. The question of how people convert the proverbial black marks on white to “stories in their heads” is, in and of itself, a fascinating one. In our attempts to address this question, we are learning a great deal more about what makes us human. After all, the ability to understand language is shared by no other animal. However, text comprehension research also has a distinct practical relevance. For example, text comprehension research provides useful information on how documents (e.g., manuals, questionnaires, drug prescriptions, educational texts) can be written so they convey their information with maximum clarity.
A fundamental assumption of this chapter is that text comprehension is an instance of complex information processing, and that it therefore complies with general principles of cognition. In this section, we use the influential construction-integration theory (Kintsch, 1988, 1998; see also Kintsch, 1974; Kintsch & van Dijk, 1978) as an illustration of a theoretical framework for text comprehension. Although individual components of construction integration remain controversial, many of the principles that it encompasses are generally accepted in the field.

The theory assumes that comprehenders process text one chunk (a clause or sentence) at a time. The processing of each chunk includes a construction phase and an integration phase. During construction, lower level processes such as orthographic analysis, word retrieval, and grammatical parsing analyze the current text chunk into idea units called propositions (Carroll, 1978; Clark & Clark, 1977; Kintsch, 1974; Townsend & Bever, 1982). For example, the propositions derived from the input sentence *Alice unlocked the wooden door with the key* are represented as P1 (UNLOCK, AGENT:ALICE, PATIENT:DOOR, INSTRUMENT:KEY), P2 (WOODEN, DOOR). Each proposition includes one predicate that, using the present notation, is written first (e.g., UNLOCK). The proposition also includes one or more concepts called arguments, each of which fulfills a distinct semantic function. The psychological validity of the proposition has been supported by numerous demonstrations: Text reading time increases systematically with the number of (a) propositions, holding constant the number of words (Kintsch & Keenan, 1973); and (b) different arguments in the text, holding constant the number of propositions (Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975). This analysis also suggests that arguments appearing in the same proposition are more strongly connected in memory than are concepts from different propositions (Ratcliff & McKoon, 1978; Weisberg, 1969).

During construction, the propositions that are expressed directly in the text are organized into a coherence network. This network is proposed to include several other types of ideas. Contained in the network are: (a) close associates of text ideas (e.g., both SPY and INSECT are associated with BUG even in the spying context); (b) inferences that contribute to the coherence of the text; and (c) text generalizations (e.g., *the student sprinkled chalk dust on the chair* might be generalized to PRANK).

During the subsequent integration phase of construction integration, activation accumulates in the network propositions that are most
highly interconnected with one another. In this process, there is a de-
activation of contextually inappropriate concepts, such as the INSECT
association to BUG in the spying context. This distribution of activa-
tion has been analyzed in the form of mathematical algorithms (e.g.,
Kintsch, 1988; Rumelhart & McClelland, 1986).

These language and discourse processes continually interact with
different memory systems. The processes of the integration phase
modify the original coherence network to produce an enduring long-
term memory representation of the text (Kintsch & Welsch, 1991). At
the end of each construction-integration cycle, a small number of
highly active elements of the current clause are held over in a limited
capacity working memory (Baddeley, 1986) for further processing
(Fletcher, 1981; Kintsch & van Dijk, 1978). It is by virtue of this carry-
over that a coherent text representation may be constructed.

The long-term memory encoding of the text consists of multiple lev-
els of representation, the surface structure, a mental representation of
the actual wording of the text, the textbase, a mental representation of
the explicitly stated semantic information in the text, and the situation
model, a mental representation of the state of affairs denoted in a text.
These levels of representation are inspected later. However, we first ex-
amine the methods typically used in text comprehension research.

**TEXT COMPREHENSION RESEARCH METHODS**

The methods typically used in text comprehension can roughly be di-
vided into two categories. The first category taps the comprehension
process as it unfolds. These methods are called online methods. The
second category of methods, historically older than the first, focuses
on the results of the comprehension process, the mental representa-
tions stored in the comprehender’s long-term memory. These meth-
ods are characterized here as memory methods.

**Online Methods**

The online methods can be subdivided into four types: processing-
load measures, activation measures, information-content meas-
ures, and brain-activity measures.

**Processing-Load Measures**

Processing-load measures are used to make inferences about the
relative amounts of cognitive resources needed to process linguistic in-
formation. This idea relies on the old assumption (e.g., Donders,
1868/1969) that processes that require more attentional or memory resources take longer than processes that require fewer resources. There are several ways in which processing times can be measured. Perhaps the most widely used method in text comprehension research is the *self-paced reading task*, in which people see segments of a text, typically clauses or sentences, on a computer screen one at a time. They advance through the text by pressing a key on the keyboard or a mouse button and the computer measures their reading times. Although this method is not as precise as the methods to be discussed next, it is precise enough to tap many of the processes in which discourse psychologists are interested.

The *moving-window task* allows more precise measurements. In this task, texts are displayed in their entirety (or one page at a time) on the computer screen, but with all the letters changed into dashes or slashes. By pressing a button or key, the reader makes a word visible and, by pressing the button again, the word reverts back to dashes and slashes and the next word appears. It is as if the reader is passing a small window across the masked text. In this fashion, reading times can be obtained for each word in the text while the reader still has an idea of the overall structure of the text.

The two preceding tasks share the characteristic that they rely on keypresses for the measurement of reading times. Thus, the resulting reading times reflect not only the time to read the sentence, but also the time needed to switch attention from comprehension to the key press and the time needed to execute the motor response. Furthermore, there is the possibility that the response times are not dictated by the amount of cognitive effort involved, but by the tendency to tap the keys in a rhythmic fashion. The *eyetracking method* does not have these disadvantages (see Rayner & Sereno, 1994, for a useful discussion). With this method, texts are presented on a computer screen and a participant’s eye movements are registered by small cameras as he or she processes the text. Thus, there are no key presses involved. Contrary to most people’s intuitions, reading is not a series of smooth horizontal sweeps of the eyes across a page. Instead, the eye pauses to take in information and then jumps to the next word (some short and highly predictable words such as *the* are often skipped). The pauses are known as *fixations* and the jumps as *saccades*. A third type of eye movement is known as *regressions*. Regressions occur when the eye moves back to an earlier word in the sentence. Along with long fixations, regressions are viewed as indexes of processing difficulty. Thus, eyetracking allows more precision than do the other methods. However, often this precision is not necessary; given that the method is relatively expensive and labor intensive (the eyetracker has to be calibrated for each indi-
individual participant), many researchers prefer to use the key-press methods.

**Activation Measures**

Various measures are used to measure the availability of information to readers as they comprehend a text. These measures allow the researcher an insight into the mental representation as it unfolds in the comprehender’s mind. The three most commonly used methods to assess activation are: lexical decision, naming, and probe recognition.

**Lexical Decisions.** In traditional lexical decision experiments (e.g., Meyer & Schvaneveldt, 1971), people are presented with a string of letters and have to determine whether it is a word. A central finding is that the lexical decision for a word like *doctor* is faster when it is immediately preceded by a close associate like *nurse* than by the unrelated word *bread* or by a row of Xs. This implies that *nurse* activates or *primes* its close associate, *doctor*. In applications of this procedure to text comprehension research, one or more words from the text are used as primes, and lexical-decision latencies are used to make inferences about the degree to which concepts are activated. Thus, one could present the opening sentence of *Alice in Wonderland*:

1. Alice was beginning to get very tired of sitting by her sister on the bank and having nothing to do.

followed by the lexical decision item *bored*. The prediction would then be that, if people made the inference that Alice was bored, lexical decision times to this word should be shorter than those to a control word. This is because *bored* was activated during the reading of the sentence, thus facilitating the lexical decision when the actual word appeared on the screen.

**Naming.** Similarly, one can assess activation simply by having participants name words. The assumption here is that priming reduces naming latencies. It has been argued that naming has the advantage over lexical decision that it does not involve a yes–no decision component (Potts, Keenan, & Golding, 1988). Another advantage is that the procedure of the naming test requires fewer items than lexical decision, which requires that about half of the test items be nonwords.

**Probe Recognition.** A third way to assess activation is by using recognition probes. In this case, a word is presented and the participant’s task is to indicate whether it appeared in the text read up to that point.
Probe-recognition shares with lexical decision the fact that it involves a decision component and thus requires more items than does naming. Furthermore, it could be argued that probe recognition focuses the comprehender’s attention on the surface structure of the text. Nonetheless, many experiments have shown that the procedure is sensitive to situation-model levels of processing. Gordon, Hendrick, and Foster (2000) recently identified a new potential drawback of the probe-recognition task. In some experimental designs, participants seem to use a strategy in which they keep track of the words that might be probed. However, this problem is likely to be confined to situations in which the pool of to-be-recognized words is small.

**Information-Content Measures**

Activation measures allow the researcher to draw inferences about specific aspects of content and structure of the mental representation that is being constructed during comprehension. Information-content measures, in contrast, provide much more extensive information. However, the question is to what extent they reflect processes that are really going on during online comprehension or to what extent they reflect task demands.

One instance of information content measures is the think-aloud protocol. Think-aloud protocols have been used extensively in research on memory and expertise. In text-comprehension applications, participants are typically presented with one sentence or clause at a time and are instructed to comment on their understanding of the sentence or clause in the context of the text they are reading. The think-aloud method has been argued to provide an accurate window on cognitive processes provided the task is well defined (Ericsson & Simon, 1993). Although this is often the case in problem-solving experiments, it is often not the case in language-comprehension experiments. For that reason, think-aloud protocols are most often used as exploratory methods in conjunction with activation measures, with the latter being used as hypothesis-testing devices. A good example of this is the study of Trabasso and Suh (1993). Trabasso and Magliano (1996) developed a detailed procedure for analyzing think-aloud protocols (see also Magliano & Graesser, 1991).

Another way to elicit information is by using a question-answering procedure in which participants are asked specific questions about aspects of a text (e.g., why did X happen?). Graesser and Clark (1985) used this procedure to uncover the range of information that might be activated during online comprehension of a particular text segment. It is also possible to elicit information about the contents of compre-
henders’ mental representations by asking them to complete a text fragment. Presumably, some linguistic devices invite different continuations than others. For example, Gernsbacher and Schroyer (1989) showed how a cataphoric pronoun such as *this* (e.g., the egg in *On the beach she found this egg*) prompted participants to maintain its referent in their continuations to a greater extent than the indefinite pronoun *an* (as in *On the beach she found an egg*).

Content measures may provide useful insights into text comprehension. However, they are most profitably used as exploratory devices in conjunction with hypothesis-testing tools for tapping online processes, such as processing-load and activation measures. This is because of the potential vulnerability of content measures to strategic processes on the part of the participants.

**Brain Activity Measures**

A recent development in the study of discourse comprehension is the use of brain-activation measures. The method of measuring event-related potentials (ERP), by attaching electrodes to a participant’s scalp, has been used for quite some time in the study of language processes. However, until recently, most of its applications were at the levels of the word or sentence. A pattern of electrical activity of particular relevance to language comprehension is known as the N400 (e.g., Kutas, Federmeier, & Sereno, 1999; Kutas & van Petten, 1994), which is a shift in negative activity occurring about 400 ms after the presentation of a stimulus. The N400 is thought to reflect difficulty in semantic integration. For example, in the sentence *He dug the hole with a pizza, pizza* will yield an N400 with a larger amplitude than if the last word were *shovel*. Recently, van Berkum, Hagoort, and Brown (1999) demonstrated that contradictions in discourse yield robust N400 effects, which suggests that ERPs might be a useful tool in the study of discourse comprehension. ERPs have several advantages over reaction-time methods, the main one being their temporal resolution. Whereas the shortest response latencies in reaction time experiments are between 350 and 400 ms (in naming tasks), ERPs show effects at shorter latencies (e.g., although N400 reaches its peak at about 400 ms, the onset is much earlier). ERPs provide some information about where in the brain processes are localized, although the spatial resolution of ERPs is rather low compared with brain-imaging methods. A drawback of ERPs is that, to obtain stable patterns within participants, large numbers of items (at least 30) are needed per condition.

Brain imaging methods such as positron-emission tomography (PET) and functional magnetic resonance imaging (fMRI) allow the re-
searcher to localize cognitive processes in the brain. As such these methods hold great promise with respect to uncovering the neural substrates of discourse comprehension. Neuroimaging methods have been used in several recent studies of text comprehension (Carpenter, Just, Keller, Eddy, & Thulborn, 1999; Fletcher et al., 1995; Maguire, Frith, & Morris, 1999; Robertson et al., 2000). Among other things, these studies converge on the rather surprising finding that areas in the right hemisphere (and not the traditional language areas in the left hemisphere) are involved in the integration of information across sentences and in the construction of spatial representations from text. Although neuroimaging research is clearly still in its infancy and various methodological hurdles have to be overcome, it is likely that neuroimaging studies of text comprehension will soon begin to constrain theorizing about language comprehension.

Memory Measures

Memory measures provide information on how the mental representation constructed during comprehension is stored in and retrieved from long-term memory. The method of free recall is one of the first methods used to study discourse comprehension (Bartlett, 1932). However, questions have arisen regarding the extent to which recall provides a window on the long-term memory representation laid down as a result of the comprehension process (e.g., Corbett & Dosher, 1978). As Bartlett noted, recall is a constructive process. Thus, what is or is not in the recall protocol does not necessarily reflect what was constructed during comprehension. For example, people may remember information and then decide to edit it out of their recall protocol because they think it does not fit very well. However, having to recall information from texts (oral or print) is a daily activity for most people. Therefore, studying recall from text is an important topic in and of itself. Memory for text is discussed in a later section.

In cued recall, the participant is presented with part of a discourse and is asked to provide a missing part, such as the second sentence of a pair. Such tasks have yielded important insights into how information is integrated in long-term memory (e.g., Myers, Shinjo, & Duffy, 1987). As discussed later, recognition (“Did you see this sentence before?”) and verification (“Is this statement true with respect to what you just read?”) are used to assess the relative strengths of different levels of representation.

When measures are not timed, it is always possible that they are contaminated by strategic processes. Ratcliff and McKoon (1978) developed the method of primed recognition, which is not sensitive or is
much less sensitive to strategic processes. In this procedure, sets of statements pertaining to texts are shown, one at a time, to participants. Their job is to indicate as quickly and accurately as possible whether each sentence occurred in the text. Each set of items contains a prime-target pair, sentences from the text that the researchers hypothesize are more related in one condition than in one or more others. The recognition latencies are thought to reflect the strength of the long-term memory link between the nodes coding for the statements or events denoted by the prime-target pair (e.g., Zwaan, 1996).

**MULTILEVEL REPRESENTATIONS**

Various theorists have argued that, during the comprehension of texts, readers construct a mental representation of the text as well as the situations described in the text. For example, van Dijk and Kintsch (1983) proposed that readers construct mental representations of (a) the text’s surface structure, (b) the semantic meaning explicitly conveyed by the text or textbase, and (c) the situation described in the text, the situation model. The first two levels—surface structure and textbase—are sometimes collapsed. For example, Johnson-Laird’s (1983, 1996) propositional representation appears to be an amalgam of the surface structure and textbase. Researchers have not agreed on a single representational format for situation models. For example, Kintsch (1998) viewed situation models largely as propositional representations, but allowed mental images as well. Johnson-Laird (1996) viewed mental models as nonpropositional, but also as different from mental images. Barsalou (1999) proposed an entirely different view, in which the building blocks of mental representation are not abstract and amodal, but analog representations called perceptual symbols that are the result of perceptual activity in the brain. Thus, at the moment, situation models are less characterized by their structure than by their content. Much of the effort in research on situation models has gone into delineating which aspects of described situations are encoded in situation models and which are not.

**SURFACE AND TEXTBASE REPRESENTATIONS**

Empirical evidence suggests that readers typically maintain a brief memory record of the wording of a text (J. Anderson, 1974; Bransford, Barclay, & Franks, 1972; Clark & Sengul, 1979; Gernsbacher, 1985; Graesser & Mandler, 1975; Jarvella, 1971a, 1971b; Kintsch & Bates, 1977; Sachs, 1967). Estimates are that, under most conditions, decay
of the surface representation can be measured in seconds. However, in some cases, surface representations are maintained over long periods of time. In a series of experiments that ruled out various alternative explanations, Murphy and Shapiro (1994) showed that a critical factor in verbatim memory is the pragmatic context of a sentence. If a sentence has high interactive value (e.g., if it is an insult or a joke), the comprehender’s attention is focused on its wording, leading to better encoding and thus to a better surface memory (see also Kintsch & Bates, 1977). Similarly, Zwaan (1993, 1994) showed that surface memory can be affected by comprehenders’ expectations about the genre of a text they are reading. When participants thought they were reading excerpts from novels, they exhibited better surface memory than when they thought they were reading newspaper articles. In both cases, the same texts were used, ruling out wording as a factor. However, this does not mean that wording plays no role. For example, poetry and songs provide constraints on fit, other than semantics—namely, meter and rhyme. Consequently, people exhibit excellent surface memory for songs and poems (Rubin, 1995).

The textbase is thought of as the explicitly stated meaning of the text. Kintsch and van Dijk (1978) developed an influential model of textbase construction, whose function it was to predict recall from text. In this model, texts were hand parsed into propositions, and the propositions were arranged into a network according to several principles. For example, a connection between two propositions was made only if they co-occurred in a working memory buffer. Mechanisms such as the leading-edge rule specified that it is predominantly the most recent and important propositions that are held over in the working memory. With appropriate assumptions about the size of the buffer, the Kintsch and van Dijk model proved accurate at predicting recall from text. However, van Dijk and Kintsch (1983) noted that their earlier model failed to capture the most important aspect of comprehension: the construction of a situation model.

**SITUATION MODEL REPRESENTATIONS**

Why is it necessary to posit situation models? Why is it not sufficient to assume that comprehenders construct a textbase? The following example from Sanford and Garrod (1998) suggests an answer.

2. Harry put the wallpaper on the table. Then he put his mug of coffee on the paper.

It is rather straightforward to integrate these sentences. They call for a spatial arrangement in which the paper is on top of the table, the mug
is on top of the paper, and the coffee is inside the mug. However, consider the following sentence pair, which differs from (2a) by only one word:

3. Harry put the wallpaper on the wall. Then he put his mug of coffee on the paper.

Most readers balk at the second sentence. Surely it is impossible to put a mug of coffee on a vertical surface. If one simply assumes that comprehenders construct propositional textbases, then no such effect would be expected. The sentence pairs are equivalent in their propositional structure and connections.

Various experiments have shown that equivalent propositional structures lead to different behavioral responses. Many of these studies were reviewed by Zwaan and Radvansky (1998). There is evidence that comprehenders keep track of at least five situational dimensions during comprehension: time, space, characters, causation, and motivation. Zwaan and Radvansky (see also Zwaan, Magliano, & Graesser, 1995) proposed that the building blocks of situation models are mental representations of single events. These event representations are integrated during comprehension based on their overlap on each of the situational dimensions. Thus, an event that occurred at the same time and place as the previous event will, all other things being equal, be more easily integratable into the evolving situation model than an event that takes place at a different time or place. By the same token, two events that overlap on multiple situational dimensions are more strongly connected in the comprehender’s long-term memory representation than two events connected on only a single dimension. It has been shown that readers simultaneously monitor multiple situational dimensions during comprehension and that this is reflected in their long-term memory representations (Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995; Zwaan, Radvansky, Hilliard, & Curiel, 1998). However, in most research on situation models, the focus is on single dimensions.

Spatial Situations

The dimension that received the most attention in early research on situation models was space. A decade prior to the coining of the terms mental model and situation model, Bransford, Barclay, and Franks (1972) demonstrated empirically that the spatial structure of the described situation can have a powerful effect on the comprehender’s memory. In this study, the participants listened to sentences such as 4
and 6. Afterward they were presented with sentences such as 5 and 7 in a surprise recognition test.

4. Three turtles rested on a floating log, and a fish swam beneath them.
5. Three turtles rested on a floating log, and a fish swam beneath it.
6. Three turtles rested beside a floating log, and a fish swam beneath them.
7. Three turtles rested beside a floating log, and a fish swam beneath it.

People who had heard 4 frequently incorrectly recognized 5, whereas people who had heard 6 rarely incorrectly recognized 7. This discrepancy cannot be explained by differential changes at the surface structure level of the test items. The only surface structure difference between the members of the pairs 4–5 and 6–7 is that the pronoun them has been replaced with it. However, the pairs differ with respect to the spatial layout they describe. Sentences 4 and 5 describe essentially the same situation: The turtles are on top of the log and the log is above the fish. Sentences 6 and 7, in contrast, describe different spatial situations. Sentence 6 has the fish beneath the turtles but not the log, whereas 7 has the fish beneath the log but not beneath the turtles. Thus, 4 and 5 are being confused because they describe the same situation. In contrast, 6 and 7 are less likely to be confused because they describe different situations.

Many more recent experiments have examined the role of spatial representations in text comprehension (see Zwaan & Radvansky, 1998, for a review). The most accurate characterization at this moment is that readers spontaneously construct spatial representations of some sort, but these spatial representations are often not detailed unless at least one of the following holds: (a) The reader has detailed prior knowledge about the spatial layout of the environment in which a story takes place, or (b) the reader is instructed, constrained, or intrinsically motivated to construct a detailed spatial representation.

Temporal Situations

More recently, researchers have started investigating the time dimension (A. Anderson, Garrod, & Sanford, 1983; Bestgen & Vonk, 2000; Carreiras, Carriedo, Alonso, & Fernandez, 1997; Mandler, 1986; Münte, Schiltz, & Kutas, 1998; Zwaan, 1996; Zwaan, Madden, & Whitten, 2000). Time is the situational dimension that is the most explicitly coded in language. For example, in languages such as English and French, a tense morpheme is attached to the main verb of each sentence, specifying the temporal position of the described event rela-
COHERENCE AND INFERENCE

To be comprehensible, texts must be coherent: Readers must be able to identify relations among the text ideas. The issue of coherence pervades the study of many central phenomena of text comprehension as becomes clear in this section.

The given-new analysis is a central principle of text coherence (Clark & Haviland, 1977; Haviland & Clark, 1974). Most sentences convey both given and new information. For example, the grammatical construction of *What Alice painted were the roses* conveys that the given information is that Alice painted something. The new information is that roses were painted. To fully grasp a sentence, a given-new strategy must be executed: The reader must distinguish the given and new information underlying a sentence, identify—in memory—a referent for the given idea, and link the new information to that referent.

The given and new ideas that underlie sentences are differentiated by many text characteristics (Clark & Clark, 1977). The grammatical structure of *What Alice painted were the roses* performs this function. Likewise, in *A dormouse drank the tea*, the definite article *the* designates *tea* as given information, whereas the indefinite article *a* signifies *dormouse* as new information. To be pragmatically cooperative, the writer must, by means of various devices, distinguish given and new sentence information in a manner that best coincides with his or her beliefs about the reader's knowledge and the previous discourse history.

Coherence and Coreference

In identifying the referent of the given information, the reader establishes that the given idea and its referent corefer to the same entity in the world. Coreference can be signaled by many linguistic devices and semantic relations. Consider the sentence, *The dormouse drank the coffee*. This sentence has many possible continuations, such as *The coffee . . ., The java . . .*, and *The beverage . . .*. These continuations are, respectively, recognized as coreferential with *the coffee* by virtue
of lexical identity (*coffee*-coffee), synonymy, and category relations. Each continuation (e.g., *The beverage* . . .) is a definite noun phrase—that is, the nouns are modified by the definite article *the*. These noun phrases function as *anaphors*—expressions that refer back to an entity or constituent previously denoted in the text.

**Anaphoric Resolution**

The prototypical anaphor is the pronoun. In *The dormouse drank the tea. It was tasty*, *it* constitutes a definite pronoun that is coreferential with *tea*. Extensive research concerning the role of definite pronouns has imparted many general principles of anaphoric resolution. A central finding is that the surface, textbase, and situational representations of texts individually and interactively influence pronoun resolution.

In their *surface* expression, pronouns vary in gender (*she*–*he*), number (*she*–*they*), and person (*she*–*I*). As a result, pronouns may (sentence 8) or may not (sentence 9) unambiguously signify their referents.

8. Sally rewarded Ron because *he* was on time.
9. Tom rewarded Ron because *he* was on time.

Sentences like 8 with pronouns that, by virtue of their syntactic characteristics, have a single referent and take less time to read than those with multiple candidate referents (Ehrlich, 1980; Frederiksen, 1981; Springston, 1975). Likewise, the pronoun is understood more quickly when the text provides only one referent (Caramazza, Grober, Garvey, & Yates, 1977; Vonk, 1985).

If anaphoric resolution were a strictly accurate process, then the anaphor would access only its correct referent. In that event, the pronoun *he* in 10 would access only KEVIN:

10. Gary gave Kevin a lot of money and *he* spent it foolishly.

In a test of this hypothesis (Corbett, 1984), people read sentences such as 10 and 11, and then had to indicate whether a test name, such as *Gary*, appeared in the sentence.

11. Gary gave Kevin a lot of money and Kevin spent it foolishly.

People needed less time to recognize *Gary* after sentence 10 than 11, which suggested that the pronoun *he* accessed GARY as well as KEVIN.
A general finding in this realm is that if a text provides several candidate referents, the pronouns temporarily access all of them (Corbett & Chang, 1983; McKoon & Ratcliff, 1980).

Another surface characteristic that influences pronoun resolution is the physical distance in the text between the pronoun and its referent. In sentence 12c, the only possible referent for the pronoun it is BOOK, but the pronoun and its antecedent are separated by intervening sentence b.

12. a. Yesterday I met a woman who had written a book on viruses.
   b. She had studied them for years and years.
   c. It was selling very well.

People take longer to read sentence 12c in a sequence such as 12a–c than when book appears in the sentence immediately preceding 12c (Lesgold, Roth, & Curtis, 1979; see also McKoon & Ratcliff, 1980; O’Brien, Duffy, & Myers, 1986). This outcome is likely a joint function of (a) the deletion of BOOK from working memory on reading 12b, and (b) the change of focus away from BOOK in sentence b.

Semantic factors interact with surface variables in pronoun resolution as suggested by 13.

13. Clinton confessed to Archie because he wanted forgiveness.

Although Archie is the most recent noun that agrees syntactically with he, the implicit causality of the verb confessed guides the resolution of he to Clinton—that is, it is usually the characteristics of the confessor that prompt confessing. In a study designed to probe these issues, participants read sentences with because clauses that were consistent (sentence 13) or inconsistent (sentence 14) with the implicit causality of a preceding verb (Caramazza et al., 1977).

14. Clinton confessed to Archie because he offered forgiveness.

The participant had to judge whether a pronoun in the because clause referred to the first or second noun in the preceding clause. Consistent with their analysis, Caramazza et al. found that pronoun resolution took longer in the inconsistent sentences than consistent sentences (see also Au, 1986; Ehrlich, 1980; Matthews & Chodorow, 1988; Springston, 1975). Another semantic relation that influences anaphoric resolution is the category link between text elements (Corbett, 1984; Garrod & Sanford, 1977). For example, reading times for sentences containing a definite anaphor denoting a category (e.g., the vehi-
cle) are longer when the antecedent is an atypical member of the category (e.g., a tank) than when it is a typical member (e.g., a bus).

Finally, pronoun resolution is influenced by the situation model of a text. As discussed earlier, the different dimensions of the situation model, such as time, spatial location, and character, may interact in complex ways. For example, after the shift from one situational episode to another, it is difficult to resolve a pronoun that refers to a minor narrative character mentioned in the former episode. However, resolving a pronoun to the main character who was last mentioned in the previous episode does not pose difficulty (A. Anderson et al., 1983). In another study that addresses narrative situation models, people read passages in alternate versions, such as ones that included either 15b or 15b (Morrow, 1985).

15. a. Paul caught the flu and was feeling pretty awful. He told his eldest son Ben to keep the house quiet. He got up from bed to go to the bathroom, irritated by the noise. ( . . . )
   b. That noisy Ben was messing up the kitchen.
   c. Ben was wondering when his father would feel better as he ate in the kitchen.
   c. The floor felt cold on his feet.

The readers generally interpreted the pronoun his in c to refer to the main character (Paul) even in a version of the passage that mentioned the son (Ben) more recently (e.g., sentence 15b). Only when the perspective was changed to that of Ben (15b) was Ben preferred as the referent of his (see also Malt, 1985).

The relative contributions of the surface, textbase, and situational representations to anaphoric resolution have been subjected to empirical scrutiny. Garnham, Oakhill, Ehrlich, and Carreiras (1995) showed that when definite anaphors appear in text shortly after their antecedent, surface (e.g., gender agreement) and situational representations make independent contributions to resolution. This contradicted a linguistic hypothesis that stated that pronoun anaphors are resolved exclusively on the basis of a referential situation model (Sag & Hankamer, 1984). Garnham et al. also reported that, with more intervening text, the contribution of the surface level quickly diminished.

In other instances, the situation model appears to predominate. Consider sequences 16 and 17:

16. I was really frightened by a Doberman. (They are dangerous beasts./ It is a dangerous beast.)
17. Last night we went to hear a new jazz band. (They/it) played for nearly five hours.
The *they* versions of 16 and 17 are ostensibly ungrammatical because the pronouns disagree in number with their antecedents. However, Gernsbacher (1991) proposed that the first sentence of each sequence supports a situation model that can function as a referent for the plural *they*. In 16, a *Doberman* signifies the category of all Dobermans, and in 17, *band* consists of a number of musicians. In fact, people needed less time to read and assigned higher naturalness ratings to the *they* versions of 16 and 17 than the *it* versions (Oakhill, Garnham, Gernsbacher, & Cain, 1992).

**Toward a Model of Anaphoric Resolution.** The present treatment might suggest the following simple analysis of anaphoric resolution: “Anaphors constitute retrieval cues for their referents. Anaphoric resolution is guided by a variety of syntactic, semantic, and discourse relations between anaphor and antecedent.” However, even aside from the aforementioned contributions of situation models to anaphoric resolution, other anaphoric phenomena reveal this analysis to be a considerable oversimplification. First, if anaphors function mainly as retrieval cues for their antecedents, then pronouns should be readily replaceable with corresponding noun phrases. As a result, sentences 18 and 19 should be equally felicitous.

18. Does Bob think that his performance will go well?
19. Does Bob think that Bob’s performance will go well?

It is apparent, however, that sentence 19 is either ungrammatical or refers to two different Bobs (Halliday & Hasan, 1976).

Second, the use of definite noun phrase and pronoun anaphors is interwoven with issues of discourse focus. In this regard, Vonk, Hustinx, and Simons (1992) showed that when people are asked to continue a story with a given word, pronouns prompt people to maintain the current text focus, whereas noun phrases (character names in the Vonk et al. study) initiate a topic shift. Conversely, comic strips that depicted topic continuity favored participants’ descriptions beginning with a pronoun, whereas those depicting topic shifts were described with names (Vonk et al., 1992, Experiment 2).

Furthermore, pronouns are preferred to definite descriptions for referents that form the current text focus (Gordon & Scearce, 1995). In a revealing study of this phenomenon, Almor (1999, Experiment 1) examined sequences like 20 and 21:

20. It was the robin that ate the apple. The bird seemed very satisfied.
21. What the robin ate was the apple. The bird seemed very satisfied.
The grammatical constructions (called *clefts*) of 20 and 21, respectively, designate *robin* and *apple* as the new (or *focused*) elements. People needed less time to read the anaphoric noun phrase *the bird* in 20 than 21. More important, Almor (1999, Experiment 3) observed the opposite pattern when the anaphor repeated the expression of its antecedent:

22. It was the bird that ate the apple. The bird seemed very satisfied.

Earlier, sequence 19 illustrated that the repetition of noun phrases can be grammatically unsuitable. Together these results show that the suitability of noun phrase repetition depends on the current condition of focus. Almor presented these observations in the framework of the proposal that anaphoric phenomena are regulated by the processing cost of identifying the referent and computing the new information signaled by the anaphor. This offers a useful step toward a unified model of anaphoric reference, and hence a deeper understanding of text coherence.

**Inference Processes**

Almost every facet of comprehension is at least partly inferential: Complex inferences are needed to identify the intended meaning of ambiguous words (*The farmer filled the PEN*), the grammatical analysis of sentences (*The old man the boats*), and the thrust of ordinary comments (*Do you have a watch?*). As a result, inference processing has been a major focus of investigation for the past 30 years.

In this section, inferences in the textbase and situation model representations are scrutinized. If from the sentence *Alice unlocked the door* the reader inferred that Alice used a key, the resulting textbase representation might be (UNLOCK, AGENT:ALICE, PATIENT:DOOR, INSTRUMENT:KEY). In contrast, understanding *The lightning struck, The hut collapsed* might require an inference in the causal situation model about the relationship between the two events.

**Bridging Inferences**

Of central importance are bridging inferences, which link the current clause to the preceding text. Consider sequence 23:

23. a. The pitcher threw to first base.
   b. The ball sailed into the field.
In keeping with the given-new analysis, understanding 23b requires identifying the referent of the given information, *the ball*. Sentence 23a does not mention a ball, but the reader can draw on world knowledge to determine that the ball was the object that the pitcher threw. Measures of reading time indicate that the processes of bridging a definite noun phrase (*the ball*) to the prior text are more consuming of cognitive resources than resolving anaphors based on identical phrasing, synonyms, and pronouns (Haviland & Clark, 1974; Lesgold et al., 1979).

Similarly, understanding sequence 24 depends on detecting a causal relation between its two sentences (Black & Bern, 1981).

24. The boy walked over to the refrigerator, bumping a bowl he had left on the table. Suddenly, it fell off the edge and broke.

By identifying the links between the current and prior text, bridging inferences preserve text coherence. In their absence, ordinary text would appear as disjointed as *Alice unlocked the door. The ball sailed into the field*. Evidence derived from numerous measures indicates that bridging inferences routinely accompany comprehension. First, Black and Bern (1981) reported that causally related sentences are more effective recall cues for one another than are similar temporally related sentences (e.g., obtained by replacing *bumping* with *seeing* in sequence 24). They concluded that readers inferentially integrate the sentences of 24.

Second, reading time has been used to probe the bridging of the following sequences:

25. Tony’s friend suddenly pushed him into a pond. He walked home, soaking wet, to change his clothes.

26. Tony met his friend near a pond in the park. He walked home, soaking wet, to change his clothes.

According to rating data, the sentences of 25 are causally closer than those of 26. Reading time for the second sentence of such sequences systematically increased with causal distance (Keenan, Baillet, & Brown, 1984; Myers et al., 1987). This outcome was interpreted to reflect the readers’ execution of bridging inference processes.

Other investigations of bridging inference have used stimulus materials such as the following:

27. The spy threw the report in the fire. The ashes floated up the chimney.

28. The spy threw the report in the fire. Then he called the airline.
Both 27 and 28 permit the inference that the spy burned the report, but only in 27 is this inference needed to bridge the ashes to the first sentence. People need more time to answer Did the spy burn a report? after sequence 28 than 27. Furthermore, answer time for 27 is indistinguishable from that observed when the first sentence explicitly states that the spy burned the report (Singer & Ferreira, 1983). Likewise, naming time for a word representing the central inference (BURN) is similar in an explicit condition and after the bridging sequence 27, but longer after sequence 28 (Potts et al., 1988). These results favor the conclusion that bridging inferences reliably accompany comprehension. The status of inferences such as those underlying sequence 28 are examined later in the section on Elaborative Inferences.

Theoretical Interpretations. The fundamental conclusion that bridging inferences routinely accompany comprehension is relatively uncontroversial. However, the development and comparison of two contemporary theories of comprehension and inference—memory-based text processing and constructionism—have focused on more subtle aspects of the computation of bridging inferences. The central feature of the memory-based analysis is that the current clause, plus additional ideas carried over in working memory, function implicitly as memory cues (e.g., McKoon, Gerrig, & Greene, 1996; O'Brien, Lorch, & Myers, 1998). In this capacity, the current clause induces the activation of prior text and relevant world knowledge by means of passive resonance processes (Ratcliff, 1978) on the basis of surface form or semantic similarities. Consistent with this analysis, the current text can restore, to working memory, matching concepts from earlier in the text and ideas associated with those concepts. A strong version of the memory-based processing analysis is that comprehension computations are restricted to (a) ideas carried over in working memory, (b) current text, and (c) ideas that have resonated to the current text (Albrecht & Myers, 1995). The memory-based analysis receives support from the findings that people's access to a text idea is regulated by its recency and degree of elaboration in the text and its similarity to the contents of working memory (Albrecht & Myers, 1995; O'Brien & Albrecht, 1991; O'Brien, Albrecht, Hakala, & Rizzella, 1995).

The competing constructionist theory has as its central principle that readers engage in a search after meaning (Bartlett, 1932; Bransford et al., 1972; Graesser, Singer, & Trabasso, 1994). Two assumptions of constructionism are that readers (a) monitor coherence at all levels of text representation, and (b) routinely seek explanations for text outcomes such as physical effects and intentional actions. The thrust of these assumptions is that the computations of comprehen-
sion do not strictly proceed from lower representations, such as surface form, to higher ones, such as the situation model. In fact, situational inferences may sometimes take precedence (Barton & Sanford, 1993; Sanford & Garrod, 1998).

Constructionist theorists have proposed that causal situation models take the form of networks of text events interconnected by links of physical, motivational, and psychological causation. Consistent with this analysis, memory for text and importance ratings of text ideas are influenced by the degree of connectedness of these ideas and by their appearance on the main causal chain underlying the text (Trabasso, Secco, & van den Broek, 1984; Trabasso & Sperry, 1985). The need to inferentially link a text statement to other, causally related ideas, is associated with higher reading times for those statements (Bloom, Fletcher, van den Broek, Reitz, & Shapiro, 1990). That result converges with the finding that achieving a complete understanding of a sentence entails the validation, with reference to world knowledge, of candidate causes of a text outcome (Singer, Halldorson, Lear, & Andrusiak, 1992).

The representation of the situation models of goal and cause that underlie text has provided a field for comparing memory-based text processing and constructionism. Several findings have been invoked as support for memory-based processing. In one study, people read passages such as one describing Mary as having to make an air reservation but getting sidetracked (Albrecht & Myers, 1995). Later she is described as going to bed—an action that is inconsistent with the goal of booking a flight. However, the detection of this inconsistency, as indexed by the reading time for the second event, depended on surface overlap between the two critical clauses. In a similar vein, readers overlooked a nearby cause for a father becoming angry ("broken window") when the text provided a different, highly elaborated cause ("lost keys") earlier (Rizzella & O'Brien, 1996, Experiment 2a). In both studies, the authors concluded that the results were consistent with the memory-based explanation and inconsistent with the constructionist explanation because the absence of surface overlap (Albrecht & Myers, 1995) and presence of a competing, elaborated cause (Rizzella & O'Brien, 1996) prohibited the detection of text causes.

In another study, the description of a character eating a cheeseburger provided access to the concept VEGETARIAN even when the character had been described as no longer being a vegetarian or even never having been one (O'Brien, Rizzella, Albrecht, & Halleran, 1998; see also Gerrig & McKoon, 1998). Consistent with memory-based processing, this result indicates that the semantic relation between CHEESEBURGER and VEGETARIAN is sufficient to reinstate VEGETARIAN.
TARIAN to working memory even when it has little bearing on the present context. Gerrig and McKoon (1998) claimed that the retrieval of text ideas not germane to the current causal structure of the text (e.g., VEGETARIAN in this example) contradicted constructionism.

However, consistent with constructionism, there have been numerous demonstrations that readers detect cause and goal relations that span moderate text distances. In these studies, either the causal conditions have been compared with control conditions that are matched for the degree of surface overlap (Long, Golding, & Graesser, 1992; Singer & Halldorson, 1996; Suh & Trabasso, 1993; van den Broek & Lorch, 1993) or surface overlap between the current outcome and its distant cause was absent (Richards & Singer, 2001). In another study, the amount of intervening text did not influence the detection of the relation between a text outcome and its antecedent cause (Lutz & Radvansky, 1997). This outcome was argued to challenge a tenet of memory-based text processing.

The distinction among these and other theoretical analyses (see Graesser et al., 1994, for a review) remains controversial. However, consideration should be given to the possibility that memory-based and constructionist processing constitute complementary rather than competing analyses. The impact on comprehension of passive resonance processes has been effectively scrutinized in the framework of the memory-based analysis, but those processes are not inherently inconsistent with the constructionist theory (Graesser et al., 1994). Conversely, the detection of complex, situational relations need not be prohibited in the memory-based framework (Rizzella & O'Brien, 1996).

Elaborative Inferences

Earlier it was recounted that people correctly answer Did the spy burn the report? faster after sequence 29 than sequence 30 repeated here. This indicates that readers inferentially bridge the sentences of 29.

29. The spy threw the report in the fire. The ashes floated up the chimney.
30. The spy threw the report in the fire. Then he called the airline.

The implication that the spy burned the report is also carried by sequence 30, but text coherence does not depend on that inference. Ideas that are strongly implied by a discourse context but do not bear on coherence are called elaborative inferences.

Considerable evidence based on measures of cued recall (Corbett & Dosher, 1978), answer times (Singer, 1980), and speeded judgments
of single words (McKoon & Ratcliff, 1986; Potts et al., 1988) has indicated that elaborative inferences, perhaps counter to intuition, do not reliably accompany comprehension. These studies scrutinized inferences ranging from implied roles, such as the participation of the police when a *burglar was arrested*, to causal predictions, such as the breaking of a delicate vase when it is described as being dropped. These results have been understood in terms of the observation that every message suggests so many plausible inferences that to compute, during comprehension, all of them would overwhelm the available cognition resources: In other words, a computational explosion would result (Charniak, 1975; Rieger, 1975).

There is, however, accumulating evidence that certain elaborative inferences accompany comprehension. It is instructive to consider some prominent examples. First, people appear to encode implied semantic features that are especially relevant to text. For example, people verify *Tomatoes are red* faster than *Tomatoes are round* after reading a text that emphasizes the color of tomatoes and vice versa for one that emphasizes the shape of tomatoes (McKoon & Ratcliff, 1988).

Second, people elaborate category terms to their specific members. Consider sentence 31:

31. Julie was convinced that spring was near when she saw a cute red-breasted bird/robin in her yard.

When sentence 31 uses the category term *bird*, people generate the elaborative inference ROBIN as is evidenced by reading times for a continuation sentence (Garrod, O'Brien, Morris, & Rayner, 1990; McKoon & Ratcliff, 1989, Experiment 2; O'Brien, Shank, Myers, & Rayner, 1988), speeded recognition (McKoon & Ratcliff, 1989), and cued recall (R. Anderson, Pichert, Goetz, Schallert, Stevens, & Trollip, 1976; see also Dubois & Denis, 1988; Whitney, 1986).

Third, it was mentioned earlier that people do not reliably draw elaborative inferences about implied roles. In one study (Lucas, Tanenhaus, & Carlson, 1990), however, people heard sequences such as 32 over headphones.

32. a. There was a broom in the closet next to the kitchen.
   b. Bill swept the floor every week on Saturday.

The participants were required to make lexical decisions about implied instrument words such as *broom*, which for this sequence appeared on a screen coincident with the end of the spoken word *week*. Lexical decision time was lower for implied instruments than for con-
trol words. Thus, although an inference about a broom does not routinely accompany the comprehension of 32b, it does when the relevant concept has been made available by the text (see also McKoon & Ratcliff, 1992).

Consistent with the latter observations, readers may also encode implied semantic roles into text representations. In this regard, Mauner, Tanenhaus, and Carlson (1995) proposed that sentence 33a is grammatically more suitable than sentence 33b.

33. a. The game show wheel was spun to win a prize.
   b. The game show wheel spun to win a prize.

Rationale clauses such as to win a prize require the participation of an agent. The short passive construction The game show wheel was spun implies the involvement of an agent, but the intransitive The game show wheel spun does not. Consistent with this analysis, in people’s examination of to win a prize, reading time was longer and the number of “does not make sense” responses greater in sentence 33b than 33a.

Fourth, there is evidence that people draw elaborative inferences about text themes. For example, 1 second after encountering the last word of The townspeople were amazed to find that all the buildings had collapsed except the mint, people made faster lexical decisions about earthquake than when the same word followed a control sentence (Till, Mross, & Kintsch, 1988).

Elaborative Inferences and the Construction-Integration Analysis. The construction-integration model (Kintsch, 1988) provides a framework for understanding the circumstances in which elaborative inferences reliably accompany comprehension. Recall that the reader initially constructs a coherence network consisting of explicit text propositions and their close associates, coherence preserving inferences, and thematic generalizations. During integration, activation accumulates in those elements of the coherence network that are highly interconnected. Consider, in this context, some of the classes of elaborative inferences just examined. On reading sequence 34, the concept RED might be constructed as an associate of TOMATO.

34. The still life would require great accuracy. The painter searched many days to find the color most suited to use in the painting of the ripe tomato.
However, RED would also bear links to COLOR, the compound concept RIPE TOMATO, and other text ideas. As a result, it would remain highly activated after the integration process. ROUND might be similarly constructed after reading 34, but, in the absence of connection with many text ideas, would likely not survive integration.

Likewise, the sequence There was a broom in the closet next to the kitchen. Bill swept the floor every week on Saturday explicitly denotes BROOM. The joint presence in working memory of BROOM, SWEEP, and FLOOR might be sufficient to result in the encoding of inferential links among them.

There is experimental evidence consistent with this analysis of elaborative inference processing. Consider sentence 35:

35. After standing through the three-hour debate, the tired speaker walked over to his chair.

One-quarter second after reading sentences such as 35, Keefe and McDaniel’s (1993) participants viewed a word that represented an elaborative implication of the sentence, such as sat. Naming time for the test word was about the same when it was implied and explicitly stated in the sentence, but longer when it followed a control sentence with similar wording but quite different meaning. This outcome suggests that the readers had drawn an elaborative inference about sitting. However, when just one additional sentence intervened between sentence 35 and the test word, naming time in the inference condition resembled the control condition rather than the explicit condition. This result pattern generally coincides with the construction-integration notion that ideas which initially appear in the coherence network may not endure integration. Immediate probing might suggest that an elaborative inference has accompanied comprehension, but it might not long survive in its competition with other encoded ideas.

MEMORY FOR TEXT

To benefit from understanding a text, one generally must remember it. The quality and quantity of people’s memory for discourse provides a reflection of the mental processes of comprehension, representations that result, and changes in those representations over time. Text memory has been examined in terms of question answering, sentence recognition and verification, and free and cued recall of text. These tasks
provide complementary and converging evidence about retrieval from text.

Stages of Question Answering

A prevailing research approach to question answering has been the identification of the processing stages that contribute to successful answering. First, questions must be encoded to their propositional form. For example, What did Alice paint? would be denoted as (PAINT, AGENT:ALICE, PATIENT:?). The related yes–no question Did Alice paint the roses? would be encoded as (PAINT, AGENT:ALICE, PATIENT:ROSES?). These two questions, respectively, request (a) the identity of the painted object, and (b) the accuracy of ROSES. This notation highlights the particular importance, in question answering, of distinguishing between the given and new information (Clark & Haviland, 1977). For What did Alice paint?, it is given that Alice painted something. The question requests the identity of the new information—namely, the painted objects.

Next, the answerer must categorize the question. Questions may interrogate the semantic roles of a question statement, such as its agent, instrument, or location (e.g., Fillmore, 1968). Other questions address complex relations between the question statement and related ideas, including relations of cause (why?), reason (how?), and time (when?) (Graesser & Murachver, 1985; Lehnert, 1978; Trabasso et al., 1984). Both wh- (Who painted the roses?) and yes–no (Were the roses painted by Alice?) questions may be formulated about any of these categories (Singer, 1990). The correlation between interrogative term and question category is imperfect: For example, why can query either a reason (Why did Alice paint the roses?) or a cause (Why were the roses red?). Therefore, question categorization typically requires syntactic and semantic analysis by the answerer.

Strategies of Question Answering. Strategy selection has been proposed to form a distinct stage of question answering. To answer a question about a text, one can attempt to retrieve the question statement from memory or judge its plausibility (Camp, Lachman, & Lachman, 1980; Lehnert, 1977; Reder, 1982). For example, to answer Did Alice paint the roses with a brush?, one could retrieve that statement from one’s memory of the story or evaluate it with reference to one’s knowledge of painting.

Strategy selection is guided by factors (a) extrinsic to the question such as task instruction, and (b) intrinsic to the question such as its current activation. The current activation of a question might be due to
the recency of encountering the question statement (Reder, 1987). In one study, participants read stories and subsequently answered questions about them immediately or 2 days later under instructions either to retrieve or judge the plausibility of the questions (Reder, 1982). The influence of instruction on strategy adoption is reflected in Table 3.1. Mean correct answer time was lower for the retrievers than for the plausibility judges in immediate testing, when the verbatim details of the story were still available, but the reverse pattern was found after 2 days. However, delay not only influenced answer times, but also strategy: Answer time under the plausibility instruction was longer in immediate than delayed testing. This suggests that, contrary to their instruction, the plausibility judges were attempting to retrieve the question statements in immediate testing. In general, people are versatile in their adjustment of answering strategy. They can adjust their strategy for each question on the basis of (a) advice (Reder, 1987), and (b) whether the question represents a recent or much earlier story (Reder, 1988; Singer & Gagnon, 1997).

**Memory Search in Question Answering.** Memory search in question answering has been clarified by J. Anderson’s (1974, 1976) fanning paradigm. In this paradigm, the participants learn facts like *A teacher is in the garage*, *An architect in the park*, and *A teacher is in the church*. In testing, recognition time varies directly with the number of facts in which the concepts of the test fact participated. For example, answer time is longer for *A teacher is in the garage* than for *An architect is in the park* because teacher occurred in two facts and architect in only one. This effect is attributed to weaker concept–fact links for concepts that have participated in more facts (J. Anderson & Reder, 1999; but see Radvansky, 1999, for a different view).

An important qualification of the fan effect is that people can focus memory search on a queried category. In one study, people learned that characters like concepts in two categories: For example, the dentist likes five (specified) cities and one animal. When asked, *Does the

---

**TABLE 3.1**

Mean Correct Answer Times (in Sec) as a Function of Assigned Strategy and Test Delay (Reder, 1982)

<table>
<thead>
<tr>
<th>Strategy</th>
<th>0 min</th>
<th>20 min</th>
<th>20 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieval</td>
<td>2.51</td>
<td>2.65</td>
<td>2.85</td>
</tr>
<tr>
<td>Plausibility judgment</td>
<td>3.21</td>
<td>2.68</td>
<td>2.61</td>
</tr>
</tbody>
</table>
dentist like giraffes?, answer time is mainly determined by the number of animals that the dentist likes (McCloskey & Bigler, 1980). Similarly, people can focus their memory search on an interrogated semantic role (Singer, Parbery, & Jakobson, 1988).

Answering questions about text clearly requires focused memory search: Why did Alice paint the roses? and How did Alice paint the roses? demand different answers. Graesser (Graesser & Clark, 1985; Graesser & Murachver, 1985) proposed that each combination of text-statement and question categories is associated with distinct memory search procedures. For example, why-action questions (Why did Alice paint the roses?) require the tracing of reason links between facts in the backward direction, resulting in answers such as, “She wanted to help the gardeners.” Evaluation of this proposal revealed that a large proportion of people’s answers to questions about narratives comply with the analysis, and that people assign higher goodness-of-answer ratings to answers that conform with the analysis than those that do not (Graesser & Murachver, 1985).

In some instances, the focused search of memory may reveal that the text representation includes no information in the queried category. For example, the memory search initiated by Did Alice paint the roses with a roller? may reveal no information about the instrument that Alice used. Frequently, this circumstance permits a rapid indication that the answerer does not know the answer (Collins, Brown, & Larkin, 1980; Costermans, Lories, & Ansay, 1992; Glucksberg & McCloskey, 1981; Nelson & Narens, 1980; Singer, 1984). More generally, there is evidence that people can assess the availability of pertinent information before they retrieve it from memory (Reder, 1987).

Parallel Processes in Question Answering

The inspection of question-answering stages such as encoding, categorization, strategy selection, and memory search carries the suggestion that the stages are executed serially. Intuitively, however, it seems likely that the appearance of who at the outset of Who painted the roses? would permit the question to be categorized, and the search for an agent to begin before the question was completely encoded. In a test of a hypothesis of this sort, Robertson, Weber, and Ullman (1993) determined that reading time for sentences beginning with interrogative terms is longer than for control declarative sentences. They concluded that the memory search initiated by the interrogative word occurred in parallel with question encoding and slowed the latter process. These observations are generally consistent with proposals of parallel processing in highly interconnected representations.
Rumelhart & McClelland, 1986). Such processing is highly interactive in the sense that low- and high-level processes mutually influence one another. In the example of Robertson et al. (1993), it is presumed that there is interaction among encoding, categorization, and memory retrieval. Likewise, Reder's (1988) finding that the recency of the question statement affects strategy selection strongly suggests that memory retrieval (which provides the index of recency) and strategy selection influence one another. Contemporary question-answering theories exhibit many of the properties of connectionist computational models (Rumelhart & McClelland, 1986). According to the ACT-R model of fact retrieval (J. Anderson, 1993; J. Anderson & Reder, 1999), for example, the connection strengths between a fact and its constituent concepts are determined by a connectionist learning rule. During subsequent testing, activation is spread in parallel from all of the concepts of the test fact (e.g., A teacher is in the garage) until a fact is retrieved.

Recall and Recognition of Text

People frequently strive to remember text in circumstances in which they are not directly queried about it. The most open-ended text-retrieval task is termed free recall. Some robust phenomena were revealed in studies of students' recall of newspaper articles. Kintsch and van Dijk (1978) classified their participants' recall responses as reproductions, which corresponded to propositions expressed directly in the original article, and reconstructions, which represented sensible guesses but did not have counterpart propositions in the text. Figure 3.1 reveals that text recall becomes systematically more reconstructive with increased test delay; after 6 weeks, the two classes account for approximately the same proportion of all recall responses. Singer (1982) measured a similar pattern—one that was hardly distinguishable between reading in the laboratory and in the natural setting (i.e., before experimental recruitment). Singer also reported that the participants reproduced 26% of 124 text ideas in immediate recall and only 6% after 6 weeks.

The convergence of reproduction and reconstruction in text recall is further clarified by people's text recognition. Text recognition is influenced by the relationship between the text and the test sentence—the test sentence can express an explicit text idea, paraphrase the text, represent a reasonable inference based on the text, or constitute an inaccurate distractor. In immediate recognition, the acceptance rate for explicit items is considerably higher than for paraphrases and inferences. With increased delay, the explicit acceptance rate does not decline much, but the rates for paraphrases and inferences rises toward
the explicit rate, and even the acceptance of distractors increases somewhat (Kintsch et al., 1990; Reder, 1982). This profile is illustrated in Fig. 3.2.

This striking pattern can be explained as follows. In immediate testing, participants have reasonably detailed representations of the surface form, textbase, and situation model at their disposal. The recognition rate reflects the number of levels with which the test item is consistent. Thus, explicit items, which are consistent with all three levels, are accepted the most often, followed by paraphrased items, which do not match the surface form, but are consistent with the textbase and situation model. Finally, inferences are accepted the least often because they match neither the surface nor textbase representations and are consistent with the situation model only (Kintsch, Welsch, Schmalhofer, & Zimny, 1990; Schmalhofer & Glavanov, 1986). With increased delay, the verbatim form of the message is all but forgotten. Likewise, there is considerable loss of the idea content (textbase) of the text over time (Kintsch & van Dijk, 1978; Singer, 1982). Therefore, delayed recognition increasingly depends on the robust situation model.
As a result, the relative recognition advantages of explicit and paraphrased test items are lost.

The complex profile in text recognition meshes with the proposal, scrutinized earlier, that text comprehension results in multiple levels of representation. It also meshes with the convergence, with increased delay, of reproduction and reconstruction in text recall (see Fig. 3.1). In recall, like recognition, the participant must increasingly rely on the situational representation as the retention interval grows. Insofar as the situation model represents the blending of text information and general knowledge, many delayed recall responses are reconstructions—they do not correspond to a text proposition.

CLOSING COMMENTS

We have reviewed the psychological theories and methods of psychological text comprehension research. Text comprehension proceeds in cycles and involves the simultaneous construction of three levels of representation: surface structure, textbase, and situation model. Key among these levels is the situation model. The situation model is a mental representation of the state of affairs denoted by a text. In most cases, the purpose of text comprehension is the construction of a situa-
tion model. Several decades of research have revealed text comprehension to be a highly complex cognitive process that involves most areas of cognition: perception, working memory, long-term memory, problem solving, and imagery.

Research is currently underway to study the neural mechanisms that subserve text comprehension. Furthermore, the traditional assumption that language comprehension involves the disembodied construction of abstract propositional networks is being challenged by recent proposals and supporting evidence that language comprehension involves analog, perceptual representations that reflect how we, as humans, interact with our environments (Barsalou, 1999; Glenberg, 1997; Stanfield & Zwaan, 2001; Zwaan, Stanfield, & Yaxley, 2002). It is to be expected that these developments will make fundamental contributions to the understanding of what goes on in our mind/brains when we process the black marks on white.

REFERENCES


3. TEXT COMPREHENSION


3. TEXT COMPREHENSION


3. TEXT COMPREHENSION


3. TEXT COMPREHENSION


This page intentionally left blank
Conversation is different from other sorts of discourse. Perhaps because we all engage in it so often, it seems simple: People who speak the same language send and receive a series of messages in sequence. Actually, it isn’t nearly so simple, because those messages are jointly shaped on a moment-by-moment basis (Clark, 1992, 1996, 1997; Goodwin, 1981; Krauss, 1987; Sacks, Schegloff, & Jefferson, 1974; Schegloff, 1982). Unlike people engaged in monologue or in reading or writing text, conversationalists have the opportunity to rely on their partners in ways that structure the discourse itself.

For example, speakers rely on their partners when they solicit help in completing an utterance, as in this interchange from Wilkes-Gibbs’ (1986) corpus, where B, having trouble finding a word, relies on A’s completion:

A: Every time I get on the elevator when it comes from the fifth floor it has this funny smell.
B: I’ve been using the. . uh. .
A: stairs?
B: stairs.
A: yeah.
Speakers rely on their partners when they alter what they say midstream based on their addressee’s feedback—for example, by amplifying or clarifying their utterances when their addressees appear puzzled or ask what they mean, or by cutting themselves short when their addressees already seem to understand what they are getting at. Speakers who have direct evidence of what is in their partner’s visual field can refer to objects differently than speakers who don’t know what their partner can see, and speakers who know their partner’s cultural background or level of expertise have the opportunity to tailor their utterances more appropriately than speakers or writers who don’t know their audience.

Addressees in conversation can rely on their partners, too. They can often assume that utterances are designed with their needs in mind. They can expect speakers to take all of their responses into account: clarification questions, backchannel evidence of comprehension like “uh-huh” or paralinguistic cues of uncertainty like “um,” or looks of confusion or impatience. Consider this example from Wilkes-Gibbs (1986), in which B’s query shapes A’s continuation:

A: We really just disagree on one thing.
B: Which is . . .
A: Which is that he wants to use terms like “mind” and I do not think that they’re necessary.

Rather than being passive recipients of speakers’ utterances, addressees in conversation can be thought of as co-creators of the utterances (Krauss, 1987).

In this chapter, we first outline features of interactive spoken discourse that distinguish it from other forms of discourse and the research methods that have evolved to examine it. We then review some of the basic facts about production and comprehension in conversation, which show various ways that speakers and addressees take each other into account, or at least seem to. We present arguments from the discourse literature about what sorts of mental processes and representations are needed to explain these findings. Based on these arguments, we lay out a set of theoretical distinctions that need to be kept in mind in the empirical study of conversation. Finally, with these distinctions in mind, we review the burgeoning and controversial empirical literature that focuses on when and whether people adapt their discourse to each other. The strengths and weaknesses of these studies highlight the difficulty of studying conversation both naturally and with experimental control, and they make it clear that the issues are far more complex than they at first appear.
WHAT MAKES INTERACTIVE SPOKEN DISCOURSE UNIQUE?

Conversation has been argued to be the primary site of language use, from which all other forms of discourse, including internal speech, are in some way derivative (e.g., Bakhtin, cited in Todorov, 1984; Clark, 1996; Rommetveit, 1974; Vygotsky, 1962). Although this point is not undisputed, it seems clear that conversation differs in fundamental ways from other sorts of discourse, and thus within the field of discourse studies requires its own theoretical distinctions and methods of study.

The obvious distinguishing features are that addressees can give speakers immediate feedback and that the medium is speech rather than writing. But there are others. Under Clark and Brennan’s (1991) characterization, participants in prototypical face-to-face conversations are physically copresent in the same environment. Face-to-face conversationalists are visible to each other, and so, unlike telephone users, they can make use of each other’s head nods, eye gaze, and gestures as they interpret each other. The participants are audible to one another, and so, unlike letter writers or users of Internet chat rooms, they can make use of paralinguistic cues and the timing of each other’s utterances to help understand each other. They observe each other’s behaviors more or less instantaneously—that is, with no perceptible delay—again unlike in chat rooms or over e-mail. Unadulterated conversational speech, unlike writing or audio recordings, is evanescent and without record; utterances aren’t available for replay or further inspection, except in the participants’ memories. Conversationalists can produce and receive simultaneously, unlike two-way radio or e-mail users. Unlike writers, conversationalists form their utterances extemporaneously, and they spontaneously determine the actions they take, unlike actors performing scripted lines, survey interviewers asking scripted questions, or participants in traditional religious ceremonies. Finally, conversationalists typically speak for themselves, unlike spokespersons, simultaneous translators, actors, or authors who take a fictitious narrative stance.

A key feature of conversation is that it is far less neat than forms of discourse that allow communicators to edit or plan more carefully before they produce utterances. Consider this interchange (overlapping speech is enclosed in asterisks):

D: uh hoo the next one is um let’s see . okay it’s similar to the one I said had the periscope
M: yes
D: but it would be scrunched down more and the second diamond is out in front like for feet. does that make any sense. the head.
M: uh
D: it's a smaller figure
M: there isn't there is no is there let me tell *you*
D: *mm hm*
M: is the highest thing on there a diamond
D: yes
M: head
D: yes
M: and then right under the diamond head to the right is a triangle
D: uh huh
M: *okay*
D: *uh huh*
M: okay (discussed in Bortfeld et al., 2001, from Schober & Carstensen's corpus)

Conversational speech is filled with phrases of indeterminate syntax, breakoffs, repairs, disfluencies, and other evidence of speakers' planning, interpretation, and coordination problems. Such bits of evidence have been proposed to function as meaningful displays of speakers' mental states (Clark, 1994, 1996; Clark & Wasow, 1998; Fox Tree & Clark, 1997; Levelt, 1989). Regardless of whether speakers intend these as displays, as Clark (1994, 1996) has argued, listeners can use them to infer how confident speakers are in what they are saying (Brennan & Williams, 1995) and may even use displays such as interrupted words to infer what speakers do not mean (Brennan & Schober, 2001).

Conversation (face to face and on the telephone) differs from other forms of interactive discourse such as voice mail, instant messaging, or e-mail in its granularity, or potential for rapid turn interchange. Turn interchange seems to follow fairly systematic principles (Sacks et al., 1974), although the principles may vary in different cultures (e.g., Moerman, 1987; Reisman, 1974; Schiffrin, 1984; Tannen, 1984, 1994). In spoken conversation, speakers often tend to minimize overlap, perhaps because it is harder to hear when more than one speaker is talking. In other forms of interactive discourse (e.g., with shared electronic whiteboards where several remote communicators can write or draw at once), communication can take place in parallel (Whittaker, Brennan, & Clark, 1991).

Another way that spontaneous spoken discourse differs from other sorts of discourse is in the roles that participants can take (Bell, 1984; Clark 1992, 1997; Clark & Carlson, 1982; Goffman, 1976; McGregor,
1986; Schober & Clark, 1989; Wilkes-Gibbs & Clark, 1992). In a two-party conversation, at any one moment one party is the speaker and the other is the addressee (except when their speech overlaps, when they momentarily play both roles simultaneously); these roles switch from utterance to utterance. A third person who is part of the conversation but isn’t currently being addressed is a side participant. A listener who isn’t part of the conversation and whose presence is known to the conversationalists is a bystander, and a listener whose presence is unknown to the conversationalists is an eavesdropper. These distinctions don’t make as much sense for other forms of discourse, where participant roles need to be characterized differently. For example, even if the author of a novel could be equated with the role of speaker (although it may make more sense for the narrator to be considered the speaker, if anyone is), readers aren’t quite addressees (readers can’t provide immediate feedback of comprehension), nor are they exactly side participants, bystanders, or eavesdroppers (yet at some level readers are being addressed; see Gerrig, 1994).

A processing theory of conversation must go beyond modeling the participants’ individual processes, integrating them into a larger whole. One explicit way to do this has been developed by Clark and Schaefer (1989), whose contribution theory expands on insights from investigators who pay close attention to details of interaction (e.g., Goodwin, 1981; Jefferson, 1974, 1989; Schegloff & Sacks, 1973; Schegloff, 1984, 1987, 1988), as well as insights by philosophers of language (Austin, 1962; Grice, 1975; Searle, 1969). The idea is that any utterance by a speaker in a two-party conversation is simultaneously a presentation by an individual and the start of a contribution at the level of collective action, and should be modeled accordingly. When a speaker presents new information in a turn, this doesn’t guarantee that the addressee will understand what was said. The speaker can assume that what was said has been grounded (understood well enough for current purposes) only by getting confirmatory evidence that the addressee has accepted the utterance. Only when the speaker acknowledges that the addressee has accepted the utterance can the original presentation be assumed to have been grounded, and a contribution to have succeeded. This acknowledgment may come easily; if an addressee continues with a relevant next turn, the speaker can presume that the contribution was a success. Or it can take more effort, as when an addressee needs multiturn clarification to understand what the speaker meant.

In so doing, speakers take advantage of another feature of conversation that distinguishes it from other kinds of discourse: the ability of one partner to reduce individual processing effort by relying on the
other. Speakers in conversation have the advantage that if they have trouble thinking of a word, they can solicit their addressees to complete their utterance (Wilkes-Gibbs, 1995), as in our example at the start of this chapter. Clark and Wilkes-Gibbs (1986) proposed that speakers follow a principle of least collaborative effort. The idea is that speakers don’t merely try to minimize their own effort in producing and comprehending, but rather try to minimize the effort they expend as a pair. This becomes especially visible in situations in which speakers and addressees’ relative abilities or expertise are asymmetrical. For example, speakers who know that their partners won’t easily see things from the speaker’s perspective (e.g., native Long Islanders discussing train timetables with non-natives) will probably put the effort into producing utterances that are intelligible to the non-native (see Bortfeld & Brennan, 1997; Isaacs & Clark, 1987). In the reverse situation, nonexpert speakers might rely on their expert addressees to fill in any missing information. Even though the precise nature of collaborative effort can be difficult to define in less concrete contexts (Schober, 1998a), the notion of least collaborative effort highlights the extent to which conversation is a joint activity, requiring as much coordination between both parties as dancing, shaking hands, or making music together.

RESEARCH METHODS

The two main methods for studying conversational interaction have been analyses of transcribed corpora of conversation collected outside the laboratory and laboratory studies of conversational interaction. Corpus studies have the advantage of ecological validity; in most cases, they represent conversations between unconstrained speakers who conversed for real-life purposes not set for them by an experimenter. They have the disadvantage that corpus researchers are usually overhearsers; because there is no independent evidence about what speakers intended and addressees comprehended, researchers must make inferences about intentions and mental processes based purely on the transcript. Another methodological issue for corpus analyses is what is considered an appropriate sample size for generalizing about the phenomena under consideration. On one end of the spectrum, ethnographic studies by conversation analysts have contributed many important insights into the moment-by-moment processes of coordination between conversational partners. However, these results are typically based on just a handful of examples of a conversational phenomenon like third-position repair (speaker A repairs B’s misunderstanding of A’s original utterance) to demonstrate the existence of the
phenomenon and are less concerned with the frequency or generalizability of the phenomenon. On the other end of the spectrum, large-scale corpus analyses rely on frequency counts of the phenomena they examine (e.g., Bortfeld et al., 2001; Clark & Wasow, 1998; Fox Tree & Clark, 1997; Poesio & Vieira, 1998), which runs the risk of lumping together phenomena that aren’t strictly the same.

Laboratory studies of task-oriented conversation have the advantage of allowing researchers to assess speakers’ intentions and addressees’ comprehension independently of the conversation, through external behaviors like grasping and moving objects. They have the disadvantage of limiting themselves to task-oriented conversation, which may differ in important ways from small talk at a cocktail party, flirting, greetings with coworkers, or shooting the breeze. And even within task-oriented paradigms, conversations are shaped differently by different tasks.

Our focus in this chapter is on laboratory studies because we believe they most directly allow investigation of the mental processes and representations underlying conversation. However, we also believe that corpus analyses are important complements to laboratory results and provide useful checks of the generalizability of laboratory results.

The most frequently used method for studying conversational process in the laboratory is some variant of the **referential communication task**, which can be traced back to Piaget’s ideas in the 1920s (see Yule, 1997), but which began to be used more frequently again in the 1960s (e.g., Glucksberg et al., 1966; Krauss & Glucksberg, 1969; Krauss & Weinheimer, 1964, 1966). In the referential communication task, two people carry out a task in which one person knows information that the other needs; the task can be carried out only through conversation. The person who knows the information has variously been called sender (Krauss & Weinheimer, 1964), director (Clark & Wilkes-Gibbs, 1986), information giver (Garrod & Anderson, 1987), instructor (Lloyd, 1991), explainer (Blakar, 1984), expert (Kraut, Miller, & Siegel, 1996), and announcer (Brennan, 1995), and the other party has been called, respectively, receiver, matcher, information follower, instructee, follower, worker, or audience. The information itself varies from task to task. In some, one partner describes objects or shapes for the other, who has several objects or shapes to choose from. In others, one partner has route information on a map that the other needs—information about how to construct a model, location information on a display, or expertise on how to repair or assemble machinery. In still others, one partner may refer to actions on a videotape that the other partner has to answer questions about later.
This basic setup allows for all sorts of variations in the nature of the material to be discussed (real physical objects, photographs, objects on computer screens, maps, mazes, missing cells in a spreadsheet, actions, stories, objects with common lexicalized labels vs. difficult to describe objects), the set of alternative objects the matcher must choose from (or what Olson [1970] called the referent array), and whether the two parties’ materials or perspectives on the materials are identical. Experimenters can vary how often the director must refer and whether the matcher and director switch roles. They can manipulate whether the matcher is another experimental participant or a confederate, whether participants are experts or novices in the task domain (Isaacs & Clark, 1987), adults or children (Glucksberg et al., 1966), strangers or acquaintances (Fussell & Krauss, 1989; Schober & Carstensen, 2002), and native or non-native speakers (Bortfeld & Brennan, 1997). They can manipulate whether speakers are in the same physical space or talk over the telephone, whether they can see each other’s faces (Boyle, Anderson, & Newlands, 1994) and each others’ displays (Brennan, 1990), whether they are seated close by or at a distance (Karsenty, 1999), and whether they are side to side or at some other offset (Schober, 1993, 1995).

The benefit of this sort of task over simple observation or corpus analysis is that it provides a number of potentially objective measures. Intentions are less mysterious because they are constrained by the task, and comprehension can be measured via task performance. At a coarse grain, measures include success or failure at the task and time elapsed. At a medium grain, measures include how a speaker chooses to linearize discourse entities (the assumption is that when an entity is available early in planning, it is mentioned early in the utterance; see Bock, 1995; Levelt, 1989). At a fine grain, measures include the time course of a partner’s performance, such as the timing of hand movements in reaching for an object, computer mouse movements traversing a map on a screen, and eye movements. The assumption with such online measures is that what is currently reached for, moved toward, or gazed at is information about what is currently being processed or under consideration.

BASIC FINDINGS

Across a range of referential communication and production studies (for a lengthier review, see Krauss & Fussell, 1996), some phenomena are quite robust. One consistent finding is that when speakers re-refer to objects that they have discussed earlier, their references change in various ways. First, referring expressions become shorter but are just
as comprehensible by their addressees (Bortfeld & Brennan, 1997; Carroll, 1980; Clark & Wilkes-Gibbs, 1986; Fussell & Krauss, 1992; Isaacs & Clark, 1987; Krauss & Weinheimer, 1964, 1966; Schober & Clark, 1989). Another regularity is that speakers tend to go from using indefinite descriptions like *a lawyer* or *some houses* to definite descriptions like *the lawyer* and *those houses* on repeated referring to the same objects (Clark & Wilkes-Gibbs, 1986; Hupet & Chantraine, 1992; Wilkes-Gibbs & Clark, 1992).

One explanation for these changes in successive referring has been that people in conversation accrue common ground or achieve joint perspectives that they rely on later. Consider the representative example in Fig. 4.1 (Brennan, 2000, from a corpus collected by Stellmann & Brennan, 1993). A and B, separated by a barrier, were trying to match sets of duplicate pictures of tangrams (abstract geometric objects). The first interchange about a given tangram took many words, several turns, and much time. Figure 4.1 (Trial 1) shows a typical first interchange; A and B produced several proposals for how to conceptualize the tangram and finally settled on one of these. After finding the right tangram, they matched (on average) 11 more and changed director/matcher roles. By the next time they referred to the tangram (in Trial

**Trial 1**

A: ah boy this one ah boy all right it looks kinda like, on the right top there’s a square that looks diagonal  
B: uh huh  
A: and you have sort of another like rectangle shape, the like a triangle, angled, and on the bottom it’s ah I don’t know what that is, glass shaped  
B: all right I think I got it  
A: it’s almost like a person kind of in a weird way  
B: yeah like like a monk praying or something  
A: right yeah good great  
B: all right I got it

**Trial 2**

B: 9 is that monk praying  
A: yup

**Trial 3**

A: number 3 is the monk  
B: ok

FIG. 4.1. Two people in search of a perspective.
2), the process of referring had become remarkably efficient because they could rely on a conceptual precedent (see Brennan & Clark, 1996) about how to view and label it. Although this process is most evident with atypical or complex objects, it happens as well in discussions of more common objects such as postcards of landmarks (Isaacs & Clark, 1987), maps (Brennan, 1990; Wilkes-Gibbs, 1986), common objects (Bortfeld & Brennan, 1997; Brennan & Clark, 1996; Fussell & Krauss, 1992), or photographs of children (Schober & Carstensen, 2002).

A second robust finding in spoken language production studies is that repeated tokens of a word (representing given information) are reduced in duration compared to the word’s first mention (representing new information; Bard et al., 2000; Fowler & Housum, 1987; McAllister et al., 1994; Samuel & Troicki, 1998). And the same word uttered under less predictable circumstances will be articulated more clearly. For example, when speakers utter the word nine, they articulate the word more clearly in a sentence like “the next number is nine” than in a sentence in which the word is highly predictable, like “a stitch in time saves nine” (Lieberman, 1963).

As Fig. 4.1 demonstrates, the shortening of expressions on repeated referring is striking when speakers address a partner who is allowed to interact freely. However, this phenomenon is strongly attenuated when speakers address a silent partner, an imaginary partner, or a tape recorder (Krauss & Weinheimer, 1966; Schober, 1993). This illustrates a third fundamental finding: Speaking in dialogue differs from speaking in monologue. The words produced by speakers describing tangrams in dialogue are of shorter duration than the exact same words produced in monologues, and speakers in dialogue shorten these words on second mention in dialogue far more than they shorten words in monologue (McAllister et al., 1994). And speakers describing locations to interacting partners were more likely to do so from their own perspectives (e.g., describing a location as “on the right” when this was true from their own vantage point, but not their partner’s), whereas speakers with imaginary partners were more likely to take the partner’s perspective (Schober, 1993).

A fourth fundamental finding is that speakers produce descriptions that appear to be tailored to particular addressees. These descriptions may be based on speakers’ prior beliefs about the addressee or on judgments about the addressee’s needs in the current situation (see Krauss & Fussell, 1996, for discussion). An oft-cited experiment by Kingsbury (1968) demonstrates the power of speakers’ prior beliefs about addressees: Speakers on a Boston city street gave an addressee who had asked for directions longer and more detailed instructions
when the addressee appeared to be from out of town (based on dress and accent) than when the addressee appeared to be a local (see also Isaacs & Clark, 1987). In another study, speakers’ gender stereotypes influenced their choices of referring expressions (Fussell & Krauss, 1992).

Choices of expressions on repeated referring appear to be based on the recent dialogue history with a particular conversational partner as well, as Brennan and Clark (1996) demonstrated. In their studies, speakers who had evolved informative referring expressions (e.g., the man’s penny loafer) with a particular partner continued to use those expressions even when the environment changed to allow simpler expressions to be used, whereas with new partners, speakers referring to the same objects tended to return to the basic level terms (e.g., shoe). Speakers often say more than they need to, despite Grice’s (1975) maxim of quantity, which suggests that cooperative speakers are only as informative as they need to be. Such choices of course may be due to the speakers’ own needs; speakers may mention what is particularly salient about an object rather than only what should be needed for listeners to identify it (Ford & Olson, 1975; Mangold & Pobel, 1988) particularly if they begin describing it before observing the other objects it needs to be distinguished from (Deutsch & Pechmann, 1982; Pechmann, 1989). Yet speakers seem to rely on more than just what they themselves find salient, adding additional descriptors when referents have what Hupet, Seron, and Chantraine (1991) have called “low discriminability” (when addressees are more likely to have trouble identifying them).

The corollary to this finding is that conceptual precedents set with successive partners can lead to a surprising degree of group consensus in a short time without explicit discussion of what the conventions should be. Garrod and Doherty (1994) demonstrated that, in a group of eight pairs of speakers referring to locations in a maze, individual pairs tended to come up with fairly idiosyncratic schemes. When they subsequently switched partners to discuss similar mazes with others in the group, the schemes began to converge. Within three to five partner switches, the entire group of pairs had converged on the same scheme. This suggests that community-wide discourse conventions can arise fairly easily as local conceptual precedents are refined.

A fifth fundamental finding concerns the experience of the addressee: Listening is not the same as being addressed. In a referential communication study by Schober and Clark (1989), matchers who interacted freely with directors did better at matching tangrams than eavesdropping matchers who did the same task silently and covertly, although both kinds of matchers heard the entire conversation. Even
when another group of silent matchers did the task later by listening to a tape recording that they could stop and start at will (presumably making it easier for them to keep up with the task), they did not do as well as the matchers who could freely interact with the directors (for a convergent finding, see Kraut, Lewis, & Swezey, 1982). Addressees also comprehend speakers’ utterances differently than do side participants and bystanders (Wilkes-Gibbs & Clark, 1992). And better than no interaction at all is listening to others interact; noninteracting matchers did better at a referential communication task when they listened to recorded dialogues than when they listened to monologues (Fox Tree, 1999).

ARGUMENTS ABOUT WHEN—AND WHETHER—ADAPTATION OCCURS

These findings show that in interactive discourse speakers often make choices that benefit listeners, as when adults and older children simplify their syntax and exaggerate their prosody for younger children (Fernald & Simon, 1984; Shatz & Gelman, 1973) or when speakers vary what they mention and how they mention it depending on what their addressees can see (Boyle et al., 1994; Brennan, 1990; Lockridge & Brennan, 2002). And listeners behave in ways that suggest that they are adapting to speakers, as when they interpret utterances differently depending on who uttered them (consider how differently one might interpret the intention underlying the question “How much do you earn per year?” if it’s asked by a prospective employer, tax auditor, date, telephone market researcher, or doting relative; see Schober, 1999). But to what extent are these true adaptations, rather than simply egocentric behaviors that look like adaptations? How much adaptation is really required for successful conversational coordination, and when does the adaptation occur? Do people need elaborate mental models of their conversational partners in order to coordinate, or can they rely on less elaborate cues?

Researchers have made various arguments in the discourse literature about these questions. One argument (Clark & Marshall, 1981) arises from the proposal that the use of definite references like the sofa requires speakers to model what they mutually know. The idea is that for a speaker to talk about the sofa (as opposed to a sofa or some sofa), she must believe that her listener can determine which particular sofa she is talking about (that the listener must be currently attending to the right sofa or be able to easily figure out which sofa is the right one). Technically, this isn’t sufficient; the listener must also believe that the speaker believes that the listener knows
which particular sofa is under discussion, and the speaker must believe that the listener believes that the speaker believes that the listener knows which particular sofa is being referred to, and so forth. Of course true mutual knowledge can’t really be computed because it entails a recursive series of such beliefs about the partner’s beliefs. In practice, Clark and Marshall’s argument goes, speakers rely on assumptions that don’t require infinite computation.

These assumptions about what is mutually known are based on three main kinds of information: physical copresence, linguistic copresence, and community co-membership. Using physical copresence, speakers and listeners can assume that the objects in their immediate perceptual field, to which they could both potentially attend (and which, in fact, they may see each other attending to), are mutually known. Based on linguistic copresence, speakers and listeners can assume that what they have said to each other in current or previous conversations can be assumed to be mutually known. And based on their assumptions about community co-membership, speakers and listeners can assume that they mutually know various facts, beliefs, and assumptions that are shared within the many communities to which they both belong. At the most generic end of the spectrum, two adults who are both native English speakers can assume that a great number of words and constructions is intelligible to each other. And at the more particular end of the spectrum, two people who have lived together for years can assume that their particular shared vacations, meals, disagreements, and private jokes are mutually known and can be referred to. In between, the many other sorts of groups that people can belong to, by choice or not—families, professions, genders, age groups, cities, nations, religions, ethnicities—can provide further bases for assuming community-based mutual knowledge.

This matters for our current purposes because it proposes that whenever a speaker does something so simple as to use the definite article the, she must be relying on personal or cultural common ground with a particular audience. This requires that, in Clark and Marshall’s (1981) terms, “we carry around rather detailed models of people we know, especially of people we know well” (p. 55). The argument is that having common ground with a partner requires a mental representation of the partner. That is, as speakers shift from one conversational partner to another, they switch gears much as they would when turning from talking with an English speaker to talking with a French one—activating one set of information and deactivating the other set.

According to Clark and Marshall (1981), the level of detail of the information that comes into play depends on how much speakers and listeners know about each other. When strangers meet, they can make
only generic assumptions about the communities to which each other belongs, based on appearance (e.g., woman in a postal uniform, young girl holding a toy truck, man in an expensive suit with a stylish haircut) and the situations they happen to be in (e.g., behind the counter at the post office, at the playground, at a bon voyage party). When strangers meet, their introductions and greetings often help them quickly determine some basic facts about each other that help narrow down the communities to which they belong: where they are from, their professions, their interests. The longer people know each other, the more information they have to draw on about which communities they do share and don’t share, and this allows them to produce and interpret utterances appropriately.

A quite different possibility for explaining the apparent adaptation between conversational partners is represented by Garrod and Anderson’s (1987) proposal that speakers and listeners in conversation rely on an output–input coordination principle. The idea is that conversationalists don’t need to have detailed models of each other in order to perform at least some kinds of speech adaptation. A speaker who uses the same words her partner did, for example, needs only to adopt the most recent successful referring expression her partner produced; output is based on the most recent input, and the dialogue history before that should not matter. Output–input coordination is an extremely local and literal form of coordination.1 The speaker takes the lead role with the addressee as follower (see also Sperber & Wilson, 1986).

Another logical possibility is represented by Brennan and Clark’s (1996) notion of conceptual pacts, or temporary, flexible agreements by partners to conceptualize an object in the same way. On this view, a speaker’s initial referring expression represents a proposal for a perspective on an object, which the addressee then ratifies or revises. The initial referring expression may be marked as a proposal via hedges (e.g., “a car, sort of silvery purple colored”), which drop out on re-referring. Two partners mark having reached a conceptual pact by reusing the same or similar expressions. This view differs from output–input coordination in that a conceptual pact is not local; the more well established a pact is between two people (the more often they have used it to refer to an object, in Brennan and Clark’s experiments), the more likely it is to persist. Unlike Clark and Marshall’s view, a conceptual pact need not be based on a model of a partner’s knowledge, but may be shaped and maintained by the partner’s feedback.

---

1In fact, this sort of simple, local adaptation underlies some successful interactive conversational software programs (e.g., Brennan, 1988; Graesser, Person, Harter, & the Tutoring Research Group, 2001).
To what extent are speakers’ choices really adjustments to others as opposed to being automatic and self-centered? Brown and Dell (1987) contributed an important insight: What looks like a partner adjustment in production doesn’t necessarily reflect partner-adjusted processing. An utterance may have been produced in a way that benefited a particular addressee, but this doesn’t by any means guarantee that the speaker was taking this particular addressee into account. The speaker may have planned the utterance in a way that suited the *speaker* best—that was easiest to plan and produce. The fact that the resulting utterance was easier for the addressee to comprehend doesn’t demonstrate that the speaker’s processing was adapted to this *particular* listener’s needs. Nor does it demonstrate even that the speaker’s processing was adapted to the needs of *any* listener; the form of the utterance might reflect purely egocentric processing.

The problem, as discussed by Keysar (1997), is that whatever is mutually known is also individually known: this means that when speakers or addressees seem to be acting on the basis of mutual knowledge, it is possible that they are merely acting on the basis of their own personal knowledge. These two kinds of knowledge are difficult to tease apart unless the speaker’s and addressee’s perspectives differ in some measurable way.

Brown and Dell (1987) also made a useful distinction between *particular-listener* and *generic-listener* adaptations. Particular-listener adaptations include any adaptation to the unique informational needs of a particular listener and so, by definition, require a model of the partner. Generic-listener adaptations are the kinds of adjustments speakers do that would benefit any listener (presumably within a given language community). For example, the fact that new or unpredictable words are more clearly articulated and given or predictable words are relatively reduced benefits listeners. But speakers seem to perform this reduction even without actual addressees; this reduction could occur if the speech production system is organized in a modular fashion. For generic listener adaptations, Brown and Dell contrasted the possibility that speakers’ planning processes are encapsulated from knowledge about addressees (a “modularity” hypothesis) with the possibility that speakers actually take addressees into account while planning (a “listener-knowledge” hypothesis).

---

2 Note that Brown and Dell’s logic assumes that speakers and addressees are monolingual and speak the same language. For the many who are multilingual, the choice of what language to use in an utterance is already a partner adaptation.

3 Grosjean (1998) proposed that bilingual speakers in conversation enter either a bilingual or monolingual mode that enables them to readily produce and understand code switches in conversation or else suppress the language that is not needed. Presumably this sort of adaptation would have to be driven by listener knowledge.
Another set of arguments in the discourse literature focuses less on whether speakers and listeners adapt to each other, but when they do. Keysar and colleagues proposed a *perspective adjustment* theory (Keysar, Barr, & Horton, 1998), also called *monitoring and adjustment* (Horton & Keysar, 1996) from the perspective of the speaker or *unrestricted search* (Keysar, Barr, Balin, & Paek, 1998) from the perspective of the addressee. This theory proposes that, although speakers and addressees often ultimately make the appropriate adaptations to each other, their initial processing is always egocentric, and “common ground . . . only plays a role in monitoring” (Horton & Keysar, 1996, p. 91). On this proposal, speakers monitor whether the utterances they have already designed are appropriate from their addressees’ perspectives and adjust to their partners only as a repair when they notice infelicity. Similarly, addressees first interpret utterances from their own perspective and only later adjust their interpretations to match what speakers might have intended from their (speakers’) perspectives.

Horton and Keysar (1996) pitted their monitoring and adjustment theory against an extreme alternative, *initial design*, in which interpretation is restricted to only the information that is in common ground. Their initial design alternative takes seriously suggestions like Clark and Carlson’s (1981) that “when a listener tries to understand what a speaker means, the process he goes through can limit memory access to information that is common ground between the speaker and his addressees” (p. 328), and Clark, Schreuder and Buttrick’s (1983) proposal that “all the information the listener should ever appeal to is the speaker’s and addressee’s common ground” (p. 258). Horton and Keysar’s initial design formulation strips the original proposals of some of their nuances; for example, it leaves out that people recognize that speakers can misjudge common ground (Clark, Schreuder, & Buttrick, 1983) and that speakers may take shortcuts in planning in order to produce an utterance sooner (Clark & Wilkes-Gibbs, 1986). But it creates a clear and testable processing alternative (although see Polichak & Gerrig, 1998, for arguments on its implausibility).

In considering when partner models come into play, we can distinguish among these two extreme possibilities and a *dual process account* (Bard & Aylett, 2000) or what Hanna, Tanenhaus, Trueswell, and Novick (2002) called a probabilistic or *constraint-based* account. In both the Bard and Aylett and Hanna et al. approaches, adapting to a partner is not a two-stage process, first egocentric and then adjusted. Rather, alternative interpretations or distinct processes proceed in parallel, as Tanenhaus and colleagues (e.g., Tanenhaus,
Spivey-Knowlton, Eberhard, & Sedivy, 1995) have proposed for the language processing system more generally. Bard et al. (2000) emphasized a difference between fast automatic processes and slower controlled processes in language production. Fast automatic processes, like priming, would likely be impervious to any information about addressees and "default" to entirely egocentric processing. Controlled processes, such as updating and assessing models of a conversational partner's informational needs, would proceed simultaneously but more slowly. Under time pressure (e.g., as Horton & Keysar [1996] concluded from their study), speakers and listeners may rely on just the fast, automatic processes and thereby fail to take their partners into account. As Polichak and Gerrig (1998) pointed out, common ground (including the potential for physical copresence) was proposed by Clark and Marshall (1981) to be a source of inferences about what partners know, and so it should not be surprising if such inferences require time and effort. It would be extremely unlikely (and computationally infeasible) if interlocutors routinely pre-computed all relevant information according to whether it was in common ground (Polichak & Gerrig, 1998).

The distinction between more controlled and more automatic processing raises the likelihood that different stages of language processing should be more or less amenable to a partner's influence (see Bard & Aylett, 2000). For example, if articulation is more automatic than the word selection that precedes it (see Levelt, 1989), it might be less likely to be affected by a partner's informational needs.

THEORETICAL DISTINCTIONS

These considerations show that the question of if, when, and how conversationalists take each other into account is far more complex than it at first appears. There are a number of critical distinctions to keep in mind. First, there are several different possibilities for adjustment:

1. Specific-partner adjustments. In speaking, such adjustments would involve making choices based on what the speaker believes this specific partner will understand. In understanding, such adjustments could involve using what the addressee knows about this specific speaker to avoid any ambiguities, better recognize the speaker's intentions, and guide relevant inferences.

2. Cultural/community/group-based adjustments. In speaking, speakers make choices based on what they believe a member of a par-
ticular group (age, gender, ethnicity, club, neighborhood, family) can understand. For example, bilinguals mix words from both languages only to addressees with whom they share those languages and not to monolingual speakers. In understanding, addressees take the group context into account in interpreting specialized language (e.g., jargon, slang) and making relevant inferences.

3. **Generic partner adjustments.** Speakers make choices that the typical user of the language would find easiest to comprehend; addressees make assumptions for processing that would be warranted for any speaker of the language.⁴

4. **No adjustments (egocentric processing).** Processing is based on speakers’ or addressees’ own ease of production or comprehension. This may or may not benefit a partner, too.

Note that the adjustments that partners make may be based on information copresent to them (that the partners mutually know), or it can be based on information that one partner happens to know about the other (not in common ground).

Second, speakers and addressees might make these adjustments for any or all of the various aspects of spoken language use that psycholinguists have catalogued (see Bock, 1995; Levelt, 1989, for models of the processes involved in language production). Speakers might adjust word choice, pronunciation or other articulatory features (e.g., word duration), syntactic structure, sentence-level topic selection, or higher level discourse planning. For example, a New York native giving subway directions to a non-New Yorker could potentially select different words than she would when speaking with another New Yorker (“the number 2 train over there” rather than “the uptown IRT”). She might adjust her pronunciation for her addressee (“Forget about it” instead of “Fuggedaboudit”). She might clarify her syntax to maximize the chances of being understood amid the surrounding chaos, saying, “the train that you should take” rather than “the train you take”). She might organize her sentences to be as clear as possible to the uninitiated (“When you leave the station, go up the stairs to your left” rather than “Go to the northeast corner”) and organize her entire discourse to be as helpful as possible (“This will be complicated—are you ready? First you take the number 2 train over there to 42nd Street, and transfer to the 1 or 9” rather than “Transfer to the local at Times Square”).

Correspondingly, addressees might adjust their comprehension of speakers’ intentions for any of these aspects of language use. For ex-

---

⁴Some generic adjustments might be context specific, as when speakers speak more loudly to addressees at a distance.
ample, a father being addressed by his daughter must adapt to her nonadult pronunciation (blanky rather than blanket), word choice (doggie instead of horse), protosyntax (up! rather please pick me up), and difficult-to-interpret sequencing of information. Some of his adaptations are likely to be generic, others group level (what children his daughter’s age are likely to say), and others quite individualized (what his daughter in particular says).

Third, if adaptations occur, their time course needs to be specified. Speakers and addressees who ultimately adapt to their partners may take their partner into account from the first moments of processing or they might adjust only later when they recognize a discrepancy. Or they might process several alternatives in parallel. How they do this may vary for different aspects of language use; that is, some aspects of language use (e.g., higher level discourse plans) may be more amenable to controlled adjustment than others (e.g., motor processing during articulation). Following Gerrig's (1986) distinction between processes and products of language use, we argue that both the processes and ultimate products of partner adaptation need to be elucidated. But the processes involved in partner adaptation are likely to unfold over longer stretches of time than most language researchers usually consider. Even the simplest models that conversationalists might create of each other would have to be built up as interaction progresses. To borrow from Gibbs (1994), the first few milliseconds of processing are important, but so are later moments of integration, interpretation, and appreciation.

These three variables—type of adaptation, aspect of language processing, and time course—define a complex potential space of partner adaptation. To further complicate matters, in actual conversation, additional factors are likely to come into play. Conversationalists' knowledge of differences in perspective, ability to see things from the partner's point of view, and motivation for doing so can affect when and how they adapt to each other.

**Knowledge**

Speakers and addressees can adapt to each other only to the extent that they are aware of differences in their perspectives and informational needs. To carry out group/community-based adaptation, they need to be aware of which groups their partners belong to and understand what pieces of knowledge come with membership in those groups. To carry out individualized adaptation, they need to know the individuals’ characteristics. Sometimes the immediate physical environment supports the knowledge that conversational partners need in
order to adapt; for example, adapting to one’s partner’s point of view when describing a location is relatively straightforward if one can see one’s partner (Schober, 1993, 1995). But keeping in mind the partner’s group membership, conceptual preferences, conversational agenda, or world view is much less straightforward, as it relies less on direct or perceptual evidence and more on memory and inferences (Schober, 1998a).

In fact, the evidence is that people are far from perfect at estimating others’ knowledge. Undergraduates who were asked to rate public figures for the extent to which the figures would be known to other undergraduates overestimated the identifiability of figures they themselves knew and underestimated the identifiability of figures they didn’t themselves know (Fussell & Krauss, 1992). As Nickerson’s (1999) review shows, this tendency for people’s estimates to be biased in the direction of their own knowledge has been found in various knowledge domains (see e.g., Bromme & Nückles, 1998; Fussell & Krauss 1991, 1992; Nickerson, Baddeley, & Freeman, 1987). In some contexts, people can even be shown to be biased in assessing their own knowledge accurately, claiming to have known more in hindsight than they did earlier and generally overestimating what they know (for reviews, see e.g., Keren, 1991; Lichtenstein, Fischhoff, & Phillips, 1982).

To adjust appropriately, not only must speakers and addressees assess each other’s knowledge, but they also must know what the appropriate adjustments are and when to deploy them. If they don’t, they are liable to adjust ineffectively (see Horton & Gerrig, in press). The native speaker who speaks more loudly for the non-native addressee has appropriately assessed the partner’s group membership and is carrying out a group-based adjustment, but she has inappropriately judged what the right adjustment might be (in fact, it may be the only adjustment she is capable of making). Similarly, an addressee who knows the speaker is a child may incorrectly assume that all single-word utterances are requests, although he has correctly assessed the speaker's group membership and lack of adult syntactic ability.

Ability

Even when people are fully knowledgeable about their conversational partner’s perspective, they may vary in their momentary or chronic levels of ability to adapt to their partners, at least within particular domains. At the momentary level, conversationalists under greater cognitive load may have fewer attentional resources available to devote to the work that adapting to a partner requires (Flavell et al., 1968; Horton & Keysar, 1996). This is likely when discourse tasks are partic-
ularly difficult (Bard et al., 2000), raising the possibility that the nature of the discourse task could play a role in contradictory findings about partner adjustments.

At the chronic level, some people may be more egocentric than others—that is, they may lack general perspective-taking skills (Davis, 1983). Others may lack the expertise or ability within a particular domain of discussion that would enable them to take another’s point of view. For example, in one study (Schober, 1998b), people with low spatial (mental rotation) ability were far less likely to accurately produce or comprehend descriptions of locations from their partner’s point of view than people with high spatial ability.

**Motivation**

Even if people know their partner’s perspective and are able to adapt to it, they may not be willing to, depending on their goals. People who are not getting along are, in fact, likely to diverge on various features of their speech as conversation proceeds (Bly, 1993; Danet, 1980; Giles, Coupland, & Coupland, 1991). According to Brown and Levinson (1987), adapting to a partner’s perspective is usually more polite than requiring the partner to adapt because it imposes less on the partner. Speakers continually monitor each other to achieve the desired levels of politeness; if they don’t intend to be polite, they adapt less. This is supported by findings like Graf’s (described in Herrmann, 1988) that students describing the location of a plant’s position in a room to an addressee represented by a puppet took their partner’s perspective far more often when they were told that their addressee represented a professor than when it was supposed to be a fellow student. Presumably this kind of influence on adaptation is fairly high level and controlled.

Finally, we should note additional complexities in the notion of egocentrism. What may look like egocentric behavior could actually turn out to be the result of sophisticated consideration of one’s partner (Schober, 1998a). For example, imagine that a speaker with low spatial ability is fully aware that his partner has high spatial ability. If he chooses to describe a location from his own perspective on the assumption that his partner can easily see things from his point of view, has he really behaved egocentrically? At some level, such a speaker has considered his own and his partner’s needs and has made a judgment that reflects the balance between the two (or that minimizes the collaborative effort they expend together), which hardly seems like an egocentric judgment. Of course the speaker’s behavior alone doesn’t allow us to know whether this has happened (see Keysar, 1997), but the point is that just as what looks like a partner adjustment may not
really be one, what looks like egocentric behavior may not really be egocentric.

A related complication to the notion of egocentric processing is the necessity of distinguishing between truly not knowing or processing the partner’s point of view and constraint-based processing. In constraint-based processing, listeners use whatever information is available to them at the moment, including, for example, information related to a speaker’s goals that may serve to determine what a speaker means by an ambiguous referring expressions (see Hanna, 2001). If egocentric information is more accessible, it may win out. But this is no guarantee that the partner-based information wasn’t being processed simultaneously (see Bard & Aylett, 2000; Hanna et al., 2002).

From these various theoretical considerations, it is clear that at any one moment of production or comprehension, people may be taking many sources of information into account at different stages of language processing. Various situational features could affect how much or in what ways a particular partner model is elaborated. The extent to which conversationalists pay attention to the relevant details of the conversational partner’s knowledge is likely affected by both parties’ conversational agendas, how much they like each other, the nature and the importance or difficulty of the task at hand, and various other factors.

THE ROLE OF THE PARTNER IN LANGUAGE PROCESSING: EMPIRICAL FINDINGS

With these theoretical distinctions and caveats in mind, we next lay out relevant empirical findings about whether, how, and when speakers and addressees adjust to each other in processing. These studies have varied speakers’ and addressees’ knowledge systematically to look at online effects of partners’ knowledge on utterance production, utterance interpretation, and language use in conversation. The challenge these studies face is establishing the experimental control required to measure processing while maintaining the realistic interpersonal settings in which interactive spoken discourse occurs. The findings are inconclusive, with some studies suggesting that speakers or addressees do not take each other into account in the earliest moments of processing and others suggesting they do.

Effects of Addressee Knowledge on Speaking

An early study (Brown & Dell, 1987; Dell & Brown, 1991) focused on how speakers express typical and atypical information when the addressee has more or less ability to infer this information. The rationale
was that instruments that are easily inferrable (as knife is from stab) are often not mentioned explicitly because they represent predictable attributes in a schema, whereas atypical instruments that depart from the prototypical value of an attribute (as ice pick does in a stabbing) are tagged and highly available during utterance planning. From the speaker’s perspective, this means that an atypical instrument is more likely to be mentioned explicitly than a typical instrument, as well as more likely to be mentioned in the same syntactic clause with the main action. The interesting question is whether the pattern of mention is affected by the needs of the addressee.

Brown and Dell had speakers retell very short stories to confederate addressees, to examine how the speakers chose to mention the instruments used to perform the stories’ main actions. Storytellers saw a picture illustrating each story along with its instrument. In half of the pairs, the addressee saw a copy of the same picture; for the other half, the addressee had no picture, and the speaker was aware of this. The reasoning was that if storytellers simply did what was easiest for themselves, they should mention atypical instruments more often than typical ones regardless of whether their addressee could see a picture. If, on the other hand, storytellers adjusted to their addressees, they should be most likely to mention atypical instruments when their partners did not have pictures because there would be no way for the addressee to infer the instruments.

Brown and Dell concluded that storytellers’ choices of whether and how to mention instruments were impervious to addressees’ needs, at least in early utterance planning. That is, atypical instruments were mentioned more often than typical instruments as expected (and more often within the same clause with the target verb than typical instruments), but whether the addressee could see a picture did not seem to matter. The only reliable addressee effect was that storytellers mentioned instruments in separate clauses after the verb more often when the addressee could not see a picture than when the addressee could. Brown and Dell proposed that this represented a relatively late adjustment or repair on the part of the storytellers. They reported no evidence for any early adjustments to addressees.

This experiment has the merit that it involved an actual addressee who was introduced to the storyteller as another experimental participant, unlike in many studies of language production with no explicit audience at all. However, no detail was provided about the confederate’s responses during the storytelling task; this is of concern because the latency and content of an addressee’s feedback can affect what speakers say. For instance, in other experiments that used a storytelling task in which addressees’ feedback was disrupted by having to do
a distracting secondary task, speakers told the stories in less detail (Bavelas, Coates, & Johnson, 2000; Pasupathi, Stallworth, & Murdoch, 1998). Thus, although the use of confederate addressees can be carefully staged so that subjects do not report suspicions about their partners, it is possible that confederates may not always provide the same kind of feedback that uninformed addressees do. This caveat applies to the Brown and Dell study, in which two confederates each heard the same stories dozens of times. Under normal circumstances, when addressees believe they already know what speakers are saying, they would probably deliver acknowledgments more quickly than when the information is new to them (see Lockridge & Brennan, 2002). If the confederate’s backchannels were too prompt, the speakers may have concluded that their addressee understood the stories all too well regardless of whether they had visual copresence.

A second feature of Brown and Dell’s study that may have affected the results is that it simulated physical copresence; speakers and addressees had separate displays and were not able to easily make eye contact and monitor each other’s attention to the same display. A final caveat is that the findings of Brown and Dell’s most relevant to our questions were null findings, which are difficult to interpret. One relevant comparison did approach significance: Speakers were marginally \( p < .09 \) less likely to explicitly mention either sort of instrument (typical or atypical) when addressees had pictures than when they did not. If this effect were reliable, it could be construed as a coarse adjustment to addressees’ having pictures.

Lockridge and Brennan (2002) used Brown and Dell’s method to look for addressee-centered adjustments with genuine (naive) addressees. Contrary to what Brown and Dell found, storytellers adjusted what they said according to whether addressees could see the pictures illustrating the stories. When addressees had no way to infer the instruments—that is, when they had no pictures and the instruments were atypical—speakers were about 10% more likely to mention the instruments in the same syntactic clause with the target action verb. This small but reliable effect suggests that speakers’ choices in early utterance planning are not fully dictated by addressees’ needs, but are not impervious to those needs either.

Other studies have examined the definiteness of referring expressions as signals by speakers about whether information is new or already known. In their studies, Clark and Wilkes-Gibbs (1986; Wilkes-Gibbs & Clark, 1992) reported that speakers were most likely to mark referring expressions as new (with an indefinite article) on the first reference to an object in conversation. Furthermore, the use of indefinite referring expressions was sensitive to addressees’ participa-
tory status in previous conversation (Wilkes-Gibbs & Clark, 1992). In Wilkes-Gibbs and Clark’s study, directors and matchers did a card-matching task while observed by either bystanders (whom directors knew about but could not see) or side participants (who sat close by the director and visibly observed the task). When the side participants or bystanders were later paired with the directors to perform the same task, the directors tended to mark their initial referring expressions as new (with indefinite articles) with matchers who had been bystanders and as given (with definite articles) with matchers who had been side participants. This is a clear adjustment to addressees’ needs because the objects marked as new were new only to the addressees (in immediately previous trials with different matchers, the speakers had used definite references). Consistent evidence that speakers adjust the definiteness of referring expressions to a partner’s needs comes from Lockridge and Brennan’s (2002) experiment: Speakers were more likely to mark their references to instruments as definite when addressees had pictures or when the instruments were typical, and they were more likely to mark references as indefinite when addressees had no pictures and when instruments were atypical.

It is reasonable to ask whether adjusting the definiteness of a referring expression to an addressee’s knowledge is done early or late in planning. Wilkes-Gibbs and Clark’s (1992) and Lockridge and Brennan’s (2002) studies did not control the timing by which speakers made their choices. Both studies examined first spontaneous mentions, which does not of course rule out that speakers may have hesitated prior to a noun phrase in order to complete a monitoring and adjustment process of the sort suggested by Horton and Keysar (1996). Yet in Lockridge and Brennan’s study, the apparent adjustment to an addressee having a picture was made just as often when the first mention of the instrument was in the same clause as the main action as when the first mention was later in the utterance. On the assumption that entities available earlier are lexicalized earlier (Bock, 1995; Levelt, 1989), this suggests that information about addressees’ knowledge was available relatively early on.

In a reanalysis of corpus data presented in Bard et al. (2000), Bard and Aylett (2000) coded referring expressions for definiteness using a hierarchy inspired by Ariel (1990) and Gundel, Hedberg, and Zacharski (1993). When a speaker produced a referring expression for the second time but directed it at a new addressee, the expression was no more likely to be marked as given than it was on being directed to the first addressee. This (null) result is consistent with Clark and Wilkes-Gibbs (1986; Wilkes-Gibbs & Clark, 1992) and Lockridge and Brennan (2002), further supporting a conclusion that the marking of
given and new status using definite and indefinite articles is sensitive to the cognitive state of the addressee, rather than the cognitive state of the speaker alone.

With respect to a different sort of syntactic choice, a series of carefully designed studies by Ferreira and Dell (2000) examined whether speakers make optional syntactic choices in a way that could actually reduce ambiguity for listeners (“ambiguity-sensitive sentence production”) or whether these choices are based on what is easiest for speakers (“availability-based sentence production”). In these studies, speakers were prompted to reproduce sentences they had heard earlier in which optional complementizers were either present or omitted (e.g., that in The coach knew you . . .). When these complementizers are absent, such sentences can be temporarily ambiguous for addressees because at least for a moment they allow either a direct object interpretation of the following pronoun (e.g., The coach knew you were younger) or a subject interpretation (e.g., The coach knew you were the best tennis player in town). This ambiguity is eliminated when the pronoun has clear case marking (e.g., The coach knew me . . . vs. The coach knew I . . .). If utterance design is availability based, speakers should produce a complementizer to fill the time when the following word is not yet available enough to be articulated, but not when the following word is already activated. If utterance design is ambiguity-sensitive, speakers should be more likely to produce a complementizer when the following word is ambiguous as to case (you rather than I); without the that, the incremental utterance (e.g., The coach knew you . . .) would be temporarily ambiguous to the addressee. Availability was manipulated by which words were used as recall cues. The same paradigm was also used to examine speakers’ production of optionally reduced or full relative clauses.

In five of six experiments, Ferreira and Dell had speakers talk into a microphone while facing a screen with a visual display of a deadline within which they had to finish producing the sentence; the experimenter sat next to them. In the sixth experiment, half of the speakers did this while the other half addressed live partners who were given the task of rating the sentences for clarity. The results of all six experiments supported the availability-based (speaker-centered) hypothesis: Speakers were no more likely to use optional function words in ambiguous than unambiguous sentences. However, in the sixth experiment, speakers did produce slightly more (7%) complementizers overall when they spoke to naive addressees (significant by-items but not by-subjects) than when they produced their utterances only for the experimenter. If speakers indeed use more optional function words overall with overt addressees, this could be construed as an adjustment to
their needs, albeit not a sensitive one because the use of complementizers was still unrelated to whether the sentences were temporarily syntactically ambiguous. Of course the question remains as to whether the sentence recall task performed with a mute addressee who would later be given a rating task is sufficiently like what speakers do spontaneously with interacting addressees (as the authors themselves pointed out). It is also possible that sentences can be disambiguated using prosody, which was not measured in these experiments.

Next we turn to an innovative experiment by Horton and Keysar (1996) explicitly aimed at teasing apart speakers’ and addressees’ distinctive (or privileged) knowledge and examining the time course by which knowledge about a partner’s knowledge is used. In that experiment, subjects referred to a target object in the context of a background object that they were told an addressee either could or could not see. The addressee was a confederate seated behind a barrier. Speakers appeared to take addressees’ knowledge into account only when they were allowed to form referring expressions without time pressure. When placed under time pressure, speakers were just as likely to produce the adjective contrasting the target object (e.g., big in “the big square”) from its background object (a smaller square) when the addressee could see the background object as when the addressee could not. These findings were interpreted as supporting a fast, automatic stage of egocentric processing in utterance planning and a slower stage of monitoring and adjusting the utterance for appropriateness. Presumably these two processes could happen in parallel, although Horton and Keysar focused on serial processes of egocentric processing followed by monitoring and adjustment.

This experiment has generated some controversy (see Polichak & Gerrig, 1998; Keysar & Horton, 1998) for several reasons. One critique is that the task afforded no physical copresence between the partners. Speakers in the privileged knowledge condition were required to keep in mind that their partner could see one part of the display (the moving part) but not the other (see Polichak & Gerrig, 1998). It would not be surprising if having to keep track of which information was and was not known to a partner, in the absence of perceptual cues to visual copresence, was disrupted by pressure to speak quickly. In addition, as the target object moved off the subject’s screen onto the confederate addressee’s screen, it sometimes changed shape (the confederate’s task was to identify whether the object was the same as or different from the one the speaker had described), which also violates the physical properties of most shared visual environments. Speakers may be better able to represent and adjust to an addressee’s perspective when there is actual or potential physical copresence between them.
In contrast to Horton and Keysar's findings with adults, Nadig and Sedivy (2000) found that even 5- to 6-year-old children (typically assumed to be more egocentric than adults) take into account the common ground they share with addressees. Children had to refer to a target object within a field of three background objects that included (a) a competitor object similar to the target (e.g., a big and a small cup) that was also visible to the addressee (and so was in common ground), (b) a competitor similar to the target but invisible to the addressee (in privileged ground), and (c) an unrelated object in privileged ground. These conditions provided a similarity control not found in the Horton and Keysar experiment. Children provided informative adjectives to disambiguate the similar targets more often in (a), when common ground included a similar competitor, than in (b), when the similar competitor was in privileged ground, as well as in (c), when there was no similar competitor.

Now we move from the level of word and syntax choices to variations in articulation. A series of studies by Bard et al. (2000; Bard & Aylett, 2000) presented evidence that articulation is an egocentric process, in which intelligibility is adjusted based not on what addressees have heard previously, but on what speakers themselves have heard. They compared intelligibility of repeated words by the same speaker talking to two different addressees in succession, finding that speakers shortened repeated words on re-referring, even though their second addressee was hearing them for the first time. This suggests that speakers do not actually take their addressees' needs into account when articulating given and new information, despite widespread assumptions to the contrary (e.g., Nooteboom, 1991; Samuel & Troicki, 1998). With the same corpus, Bard and Aylett (2000) found that, although speakers adjusted the definiteness of their descriptions to new addressees, they did not adjust their articulation, which Bard and Aylett argued supports the dual-process model. Unfortunately, the corpus didn't enable an important control: repeated mention to the same partner. That is, intelligibility degraded slightly with a new partner, but it might have degraded even more with an old partner. Recall that McAllister et al. (1994) found that repeated words had shorter durations in dialogues (with an addressee present) but not in monologues, suggesting that partner adjustments may be possible at the articulatory stage after all.

Particularly interesting is the Bard et al. (2000) finding that when the first mention of a word is by one partner and the second mention is by the other, the second mention is shortened as if it had been produced by the first partner. This suggests that both the speaker's own and her partner's utterances affect the speaker's representations of given and new information.
A study by Schafer, Speer, Warren, and White (2001) examined the use of prosody in a task-oriented discourse context, finding that speakers indeed used prosody to disambiguate prepositional phrases with high versus low attachment, but that the use of this prosodic contrast was independent of whether the prepositional phrase would actually have been ambiguous to addressees in the syntactic context.

In sum, the evidence about whether and how speakers adjust to addressees is mixed, due in part to the difficulty of teasing apart the addressee’s knowledge from the speaker’s, to variations in how co-present the partners are and how realistic the tasks are, to other confounds and missing controls, and to the difficulty of interpreting null effects. It seems clear that some types of adjustments are more feasible than others. Taken at face value, the findings suggest that articulation and some syntactic choices are more automatic and less influenced by addressees’ needs than are other syntactic choices, referring expressions, and their marking as definite or indefinite.

Effects of Speaker’s Knowledge on Addressees’ Interpretation

An extensive program of experiments by Keysar and colleagues has tested whether interpretation is restricted to only that information presumed to be in common ground or whether privileged knowledge (information known only to the addressee) interferes. In the first Keysar et al. (1998) study, subjects who acted as addressees were given information on a computer screen (e.g., “Rachel delivered the sofa”) with which to answer yes or no to confederate speakers’ questions (e.g., either “Did Rachel deliver the sofa?” or “Did she deliver the sofa?”). The questions were prerecorded (but were assumed by subjects to be live) and delivered over an intercom. Subjects also had a second task in the guise of increasing their memory load; they had to remember a second statement that was presented right after the relevant information (e.g., “Marla delivered a cake”). They were told this information was irrelevant to the task and unknown by the speaker (making it privileged to the subject). The key comparison was whether subjects would be slower to answer questions containing pronouns that were ambiguous between referents in the shared and privileged sentences. Indeed, the privileged information (e.g., Marla) interfered with the information in common ground (e.g., Rachel) in the pronoun/same gender condition, yielding longer latencies and increased errors.

In the second Keysar et al. (1998) study, a live confederate served as speaker and the addressee’s gaze was tracked, replicating the interference of privileged knowledge. These results were used to rule out a restricted search hypothesis, concluding that addressees do not restrict
their search for referents to only that information presumed to be in common ground.

In both of these studies, however, the privileged information that competed with the information in common ground was not only ambiguous, but strongly favored by recency. It would have been remarkable if subjects had been able to suppress this salient information entirely; interference between similar items is pervasive in the human memory system. The question remains: If both privileged information and common ground were to start on a level playing field, would common ground tend to be consulted first?

In a study with a confederate speaker speaking from a script to naive addressees (Keysar, Barr, Balin, & Brauner, 2000), a vertical display of see-through cubbyholes held objects that were either visible to both speaker and addressee (common ground condition) along with objects occluded from the speaker (addressees’ privileged ground condition). The goal was to test a version of the *optimal design* principle—that addressees would search only common ground when the speaker made a reference that was ambiguous between an object in common ground and another in privileged ground. Addressees’ first looks were not limited to what was in common ground, but went to objects in privileged ground just as often. The authors concluded that initial processing is egocentric.

However, in this study the critical instructions were biased in favor of the objects in privileged ground over those in common ground. That is, with “put the smaller candle. . . ,” there was one large candle, a smaller candle in common ground (the speaker’s “intended” referent), and an even smaller candle in privileged ground (for discussion, see Hanna et al., 2002). The Hanna et al. (2002) study removed this bias so that the privileged object was not the most typical referent for the critical referring expression. When this bias was removed, first looks were made significantly more often (although not solely) to the object in common ground than to the one in privileged ground. This indicates that common ground does not operate only as a late filter, but can be used in the earliest moments of processing.

Apparently small children can use common ground in comprehension as well. To return to Nadig and Sedivy’s (2000) eye-tracking study, 5- to 6-year-olds were instructed to pick up a target object (e.g., the cup) when there was (a) a similar competitor also visible to the speaker that rendered the expression ambiguous (e.g., a second cup), (b) a similar competitor that was visible only to the child and thus in privileged ground, and (c) an unrelated competitor. The children’s eye movements showed significant numbers of looks to the similar competitor in (a) (where the reference to the cup was infelicitous), but no interfer-
ence from the similar competitor that was invisible to the speaker in (b). This evidence is consistent with the interpretation that common ground helps constrain the interpretation of referring expressions from the earliest moments of processing; the fact that it can do so in children this young is remarkable.

Interpreting the Findings About Speakers’ and Addressees’ Apparent Adjustments and Failures to Adjust

These empirical findings leave us with the proverbial problem of deciding whether the glass is half full or half empty. Should speakers and addressees be characterized as adapting to each other well or even adequately? Or should they be considered mired in egocentricity? The trouble is that data from the same studies can be given alternate spins, either focusing on people’s successful adaptations or on their mistakes.

Consider the data from Clark and Schaefer’s (1987) study, in which pairs of undergraduate friends were given the task of describing campus landmarks for each other in ways that would conceal their identity from overhearers. They managed to develop successful private keys about 50% of the time. But the rest of the time they leaked far more than they realized; overhearers could guess the referents fairly accurately based on descriptions the pairs thought would be private (like “This is where someone wanted to put my teddy bear” to refer to a fountain). One spin on these results is that students were moderately successful at concealing information from others, but the alternate spin is that they made a great number of errors.

Or consider Lockridge and Brennan’s (2002) finding that speakers adapt to their addressees’ needs in instrument mention. The raw percentage of adaptation, although statistically reliable, was actually fairly small; speakers were 10% more likely to mention atypical instruments early in syntactic planning when addressees lacked pictures. Again, one spin is that speakers adapted to their listeners; another is that the adaptation was relatively modest.

The truth of the matter is obviously that people can fail to adapt to each other. We have all been in situations where speakers have seemed oblivious to the knowledge or informational needs of their audiences; think of the jargon-spewing lecturer who assumes too much (or the condescending lecturer who assumes too little) or the person in the cinema who somehow can’t seem to whisper. And we have all been in situations when addressees have failed to consider the source as they interpreted speakers’ utterances, finding insult where none was intended, misinterpreting small talk as serious talk, or assuming they
understand what is beyond their level of expertise. Experimental findings on referential communication from the beginning have demonstrated that people don’t adapt to each other in every way they possible could. Glucksberg et al. (1966) found that 4-year-old speakers in a dyadic communication task produced messages that were unintelligible to others, although the messages were intelligible to the speakers. More recently, Russell and Schober (1999) found that speakers who were not informed that their partners’ task goals differed from theirs persisted in assuming that their partners shared their goals, ignoring mounting evidence in the discourse that there were problems (see also Schober, 1998b).

In fact, although people can form effective conceptual pacts that allow them to refer efficiently in the task at hand, there is no guarantee that successful referring entails truly aligned conceptualizations. Sometimes people in conversation do not detect important mismatches. In studies of how respondents in standardized surveys comprehend seemingly banal questions about facts and behaviors (Conrad & Schober, 2000; Schober & Conrad, 1997), respondents interpreted ordinary words like furniture, bedroom, and job quite differently than the survey designers intended a substantial portion of the time (see also Belson, 1981, 1986). It almost never occurred to survey respondents that their interpretations could possibly be different than the interviewers’, even with extensive instructions to that effect. Survey respondents seemed to be relying on a presumption of interpretability (Clark & Schober, 1991)—that if their own interpretation was insufficient, the interviewer would have somehow made this clear.

The moral seems to be that people’s apparent success at a communicative task doesn’t in and of itself provide evidence about how—or if—they have taken each other into account. For example, in one study, survey respondents who were asked “Have you smoked at least 100 cigarettes in your entire life?” answered “yes” or “no” without hesitation; they seemed to be communicating successfully. Although they never suspected that their notions of what counted as smoking (finishing cigarettes? taking a puff?) or cigarettes (tobacco? marijuana? cloves? cigars?) might differ from other people’s, 10% of them changed their answer to the question when provided with a uniform definition later (Suessbrick, Schober, & Conrad, 2000). Similarly, people’s communicative failure doesn’t show that they haven’t taken each other into account. For instance, when speakers in phone conversations suddenly discover that they don’t mean the same person by a proper name, it often isn’t because they have failed to assess their partner’s knowledge appropriately, but rather that they have failed to produce an appropriate cue to that knowledge (Horton & Gerrig, 2002).
Much remains to be discovered about when and how speakers and addressees adapt to each other. In particular, we know little about the mental representations involved in partner adaptation. When there is a model of a partner's characteristics or beliefs, we don't know how elaborate or coherent it is, nor what role it plays in cognition in general. We don't know exactly how detailed the representation is, how specific it is to an individual, or how often it is updated.

What is striking from the literature on how and when speakers and addressees take each other into account is how much more complex the issues are than they at first seem. Given the complexities, it seems naive to imagine that we can determine, across discourse contexts, and for a particular aspect of language use (say word selection in production), what sorts of adaptations always occur. Presumably the adaptations vary according to conversationalists' (chronic or momentary) attentional capacities, discourse goals, interest in taking their partner's perspective, and so on. It is tempting to set the research agenda in traditional terms: to examine the limits on what speakers could do. That is, if we could discover that one aspect of processing is entirely modular and can't allow initial partner adaptation, this would constrain our theory of just how much speakers can adapt. But such a research agenda is problematic, not least because it leads to a search for null findings (and it is logically impossible to show that something isn't possible). A traditional research agenda would also require researchers to generalize across too many unspecified variables, as any one experimental setting can only look at one small corner of the space of possibilities.

The evidence so far suggests that adaptation doesn't seem to be an all-or-nothing phenomenon at any level; people can be shown to adapt under some circumstances and not to adapt under others at virtually every level of language use—from higher discourse-level functions to articulation. Thus, we propose, the more fruitful research agenda is to explore the factors that affect conversationalists' adaptations in particular circumstances—the sorts of tasks, individual ability differences, discourse goals, and so on that affect when and how partners can adapt to each other.

Systematic research always requires trade-offs between naturalism and experimental control. But the methodological issues in studying interactive spoken discourse are particularly thorny because the challenge is to model people's individual mental processes while they are simultaneously involved in the collective activity of conversation. The trade-offs that experimental psychologists often make—for example,
having participants listen to prerecorded disembodied utterances so as to control the stimulus—simply won’t do for studying processing during interactive discourse because the prerecorded utterances aren’t contingent on the participants’ actions. An interacting scripted confederate may be effective for some purposes, but only if the confederate’s conversational feedback is as contingent and plausible as a native partner’s would be. The tasks prescribed by an experimenter are necessary for control, but unless these tasks contact what people do naturally during communication they may not generalize. Obviously intrusive online tasks like lexical decision are unsuitable for studying interactive discourse. The use of relatively nonintrusive head-mounted eye-tracking devices is extremely promising, but the eye-tracking method brings its own methodological challenges. Ultimately, a complete theory of when and whether partners adapt to each other in conversation requires researchers to use a wide variety of tasks that involve different kinds of goals and affordances.

An even greater challenge is to be clear about theoretical distinctions that haven’t always been made. Researchers must distinguish individual and mutual knowledge. They must be clear about what sorts of common ground and individual knowledge are involved in any particular interaction. Perceptual copresence, for example, potentially provides several kinds of information, any of which might make a difference in processing; participants can gain knowledge about their partners by seeing whether they look confused, what they appear to be looking at or pointing to, what progress they have made in a task, and so on. Experiments that set up scripted situations in which conversationalists can both see some of the same physical objects while individually seeing privileged competitor objects need to be clear about the relative salience of those objects (how typical they are, how relatively well they fit the referring expressions, how recently they were mentioned, and how they were previously mentioned). Studies examining situations where speakers have common ground based on linguistic copresence or community comembership must be clear about exactly what sorts of inferences participants must make, taking care that the experimental demands aren’t too different from real-life attentional demands.

Finally, experimenters need to recognize that laboratory settings themselves are social situations with their own logic and intentional structure (for explicit discussion, see Schwarz, 1996, 1998). To the extent that experimental participants recognize that there is an extra agenda-laden audience—the experimenter—for anything they do in the lab, the seemingly simple two-participant experiment may actually become a complex multiagent situation in which any utterance may in-
solve more than one level of intentionality. How people adapt to their partners in such circumstances may or may not reflect what they would do unobserved. The challenge for studies of interactive spoken discourse, then, is to make appropriate design trade-offs and really understand the language game in which experimental participants find themselves.

ACKNOWLEDGMENTS

We are grateful to many colleagues for helpful comments and debate concerning this chapter. This material is based on work supported by the National Science Foundation under Grants No. 0081550, 0082602, 9980013, 9711974, and 9730140. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES


Horton, W. S., & Gerrig, R. J. (in press). Speakers’ experiences and audience design: Knowing when and knowing how to adjust utterances to addressees. *Journal of Memory and Language*.


knowledge: Process models of common ground in comprehension. *Journal of Memory

Keysar, B., Barr, D. J., & Horton, W. S. (1998). The egocentric basis of language use: In-


Kingsbury, D. (1968). *Manipulating the amount of information obtained from a person
giving directions*. Unpublished honor’s thesis, Department of Social Relations, Har-
vard University.

Krauss, R. M. (1987). The role of the listener: Addressee influences on message formula-

communication. In E. T. Higgins & A. Kruglanski (Eds.), *Social psychology: Hand-
book of basic principles* (pp. 655–701). New York: Guilford.


Krauss, R., & Weinheimer, S. (1964). Changes in reference phrases as a function of fre-

Krauss, R., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the en-
coding of referents in verbal communication. *Journal of Personality and Social Psy-
chology*, 4, 343–346.


tasks: Effects on outcomes and communication. *Proceedings, ACM Conference on
Computer Supported Cooperative Work*. Cambridge, MA: ACM.

Press.

state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment
under uncertainty: Heuristics and biases* (pp. 306–334). Cambridge, England: Cam-
bidge University Press.

Lieberman, P. (1963). Some effects of semantic and grammatical context on the produc-

Lloyd, P. (1991). Strategies used to communicate route directions by telephone: A com-
parison of the performance of 7-year-olds, 10-year-olds, and adults. *Journal of Child
Language*, 18, 175–189.


McAllister, J., Potts, A., Mason, K., & Marchant, G. (1994). Word duration in monol-


In the final decade of the 20th century, the quality of classroom discourse has become a prominent focus in discussions of school reform. There are several reasons for this increased attention. According to two economists of education (Murnane & Levy, 1996), the “new basic skills” required for high-wage jobs include “the ability to communicate effectively, both orally and in writing” and “the ability to work in groups with persons of various backgrounds” (p. 32). Demographic and technological changes in society have meant that these same skills are necessary for effective participation in and maintenance of a democratic and just society. As a result, schools are now charged with creating not only individual human capital for a healthy economy, but also social capital for healthy communities. At the same time, what counts as knowledge has shifted away from inert information passively received from books and teachers toward dynamic understanding that is collaboratively constructed in discussion among students.

Education’s new emphasis on the ability to communicate requires that classroom interaction change dramatically to foster such ability. Curriculum standards now place less emphasis on facts or procedures to be learned by heart and more emphasis on strategies for learning and doing. Instead of the traditional pattern of classroom talk in which teachers ask testlike questions and students give short, testlike answers, teachers are being asked to lead discussions that stimulate and support higher order thinking. Likewise, students are
being asked to explain their reasoning, listen, learn from, and even argue with their peers. Current high-stakes tests administered by states and school districts reflect a similar shift: Instead of answering only multiple-choice questions, students must now write out extended explanations and arguments for their answers to open-ended questions. As one example of such a test, we cite the Massachusetts Comprehensive Assessment System (MCAS) from our own state (Massachusetts Department of Education, 2001). This test, which is administered at several elementary and secondary grade levels, requires students to write extended answers to open-ended questions in all subject areas of the test, including science and math. The National Assessment of Educational Progress (NAEP), which is widely recognized by educators as a valid measure of students’ achievement in school subjects nationwide, also demands that students demonstrate their thinking skills in extended written discourse (see National Center for Education Statistics, 1998, for examples from the NAEP subject area of reading). These changes mean that classroom discourse now counts as much more than just a background context for individual students’ learning. It has become an essential social process by which students accomplish complex conceptual and communicative goals.

At the same time, a change in how we think about the term *discourse* reflects its social significance. This term used to refer to any stretch of spoken or written language longer than a single sentence. Recently, however, Gee (1996) made an influential distinction between that meaning, which he called *little d discourse* and what he called *big D discourse*. Discourse with a capital “D” is

> a socially accepted association among ways of using language, other symbolic expressions, and ‘artifacts’ of thinking, feeling, believing, valuing and acting that can be used to identify oneself as a member of a socially meaningful group or “social network,” or to signal (that one is playing) a socially meaningful “role.” (Gee, 1996, p. 131)

This definition highlights what is entailed in the fluent enactment of any language pattern. Learning new ways with words entails taking on new interactional roles and the new identities they create and express. Thus, we refer to classroom discourse with a “little d,” whereas we refer to the Discourse of schooling, or academic Discourse, with a “big D.”

For children born into language communities outside the mainstream, taking on such roles and identities is crucial to mastering the Discourse of schooling. Changes in curriculum and assessment methods are happening with the expectation that they will provide equal op-
opportunities for all students. Finding ways to decrease the achievement gap among ethnic and social class groups has become a national priority. It has always been the case that formal schooling requires forms of Discourse from students that are different from the informal talk of home and street. Yet the more different the forms of schooling are from the informal discourse of students' out-of-school lives, the more attention we have to pay to helping all students learn to play their new roles. Within the past decade, educators such as Britain's Sheeran and Barnes (1991) and African American Delpit (1995) have urged teachers to be explicit with students about the ground rules for speaking and writing in these new roles. In other words, the cultural conventions of Discourse have to become part of the new curriculum.

This chapter is divided into the following six sections:

- Sharing time: traditional and nontraditional
- Lessons: traditional and nontraditional
- Variations in discourse features
- Talk with peers and computers
- Differential treatment and cultural differences
- Students' Discourse development

Throughout these sections, we include studies by teachers in their own classrooms as well as academic researchers. We also visit and revisit several methodological concerns: (a) the advantages of longitudinal analysis as a way to document changes in children's patterns of language use; (b) the necessity of identifying speakers and other participants in a discourse situation by whatever dimensions of diversity are relevant to the situation under study, and of tracking who holds the speaker's role; and (c) the importance of including some measure of students' learning separate from what can be read off the discourse. These methodological issues are important for reasons of validity and rigor, but they also speak to widespread concerns about equity. If we wish to understand how to foster students' communicative abilities to nurture the development of complex cognitive skills, we must have a way to describe the trajectory of their growth as users of language. If we want to know whether a classroom situation is encouraging democratic participation by members from a range of social classes or ethnic groups, we must document their participation according to such membership. If we want to know how opportunities for participation in talk are related to achievement for all students, we must have some index of students' achievement beyond their participation alone.
This chapter draws on a recent book by Cazden (2001), which includes extensive examples, including transcripts, of discourse forms and functions referred to in the present chapter. This chapter offers a more detailed discussion of discourse research methods.

**SHARING TIME: TRADITIONAL AND NONTRADITIONAL**

As a routine classroom activity in many primary classrooms, Sharing Time (ST) is unique in two ways. In many classrooms, it may be the only chance that children have during official classroom time to compose their own oral texts and select a topic not constrained by relevance to previous discourse. Moreover, ST may be the only time when sharing out-of-school experience is considered appropriate in the classroom. Due to these special features, researchers have studied its traditional characteristics. More recently, teacher researchers have described its variations.

**Traditional Sharing Time**

A series of studies of traditional ST in California (Michaels, 1981) and the Boston area (Michaels & Cazden, 1986) provide a picture of typical children’s narratives and typical teacher responses. There are certain microfeatures—such as a beginning that provides information about time and main characters and a rising intonation at the end of a tone group—that are common to all children’s ST narratives. At the macro-level of the genre, traditional ST narratives can be categorized into two types: topic-centered, which focus on a single object or event, and episodic, which include shifting scenes and are usually longer as a result. Whereas in informal conversations narrators hold the floor and are immune to control by other participants (Pratt, 1977), ST narrators have no such immunity. Teachers’ responses to ST narratives fall along a continuum of appreciation based on the extent to which the child and teacher share a sense of appropriate topics and appropriate ways of talking about them. This continuum extends all the way from the teacher’s appreciative “Wow!” or “That’s neat!” to a request to stop or talk only about things “that are really important.”

In the classrooms observed by Michaels and Cazden, teachers’ responses between these two extremes can be grouped at four points along the continuum from most appreciative to least. Less purely appreciative than the enthusiastic response at the conclusion of the child’s story is a response that combines appreciative evaluation with questions interjected throughout the child’s narrative. Less appreciative still is a response in which the teacher refrains from any evaluation
and, at the end, simply adds a question or neutral comment. More toward the negative end of the continuum is a teacher question that reveals her inability to follow the child’s narrative thread. Finally, the teacher may attempt to shift the child to a topic that the teacher either understands better or values more highly.

What causes children and teachers to disagree on appropriate topics or narrative styles? It is not that children are unable to accommodate the needs of an audience. In a microanalysis of children’s self-corrections, Cazden, Michaels, and Tabors (1985) showed that the children—topic-centered narrators and episodic narrators alike—demonstrate both their syntactic competence and concern for their listeners. Yet there are other plausible explanations. Episodic stories tend to be longer and more complex and, as a result, may place a greater cognitive load on the audience. Moreover, teachers may have more difficulty understanding stories about out-of-school experiences that are distant from their own lives. Because narrative styles are often culturally specific (Smitherman, 1977/1986), the White teachers in the classrooms studied by Michaels and Cazden may have had different narrative expectations than their African-American students. Finally, teachers and children may see different purposes for ST. Whereas children may tell stories to entertain or impress their peers, teachers may consider ST an opportunity for children to gain experience composing an oral text that is similar to a written composition in clarity, precision, and explicitness.

Some of the influences on teachers’ and students’ participation in ST routines point to cultural, generational, and socioeconomic factors that can contribute to misunderstandings and decreased opportunities for participation and narrative development. Researchers studying these influences need to document the identities of the teacher and student participants in a way that recognizes the salient features of their backgrounds. Whenever possible, it is important that they augment the audio transcripts with interviews.

**Nontraditional Sharing Time**

Teacher researchers have recently reported what happened when they designed and encouraged variations in the purposes and structures of traditional ST. These reports complement those by academic researchers in significant ways because of teachers’ more intimate knowledge of their children and their school lives over time and perhaps also because of the interests of the particular teachers who become teacher researchers.

At a middle-class elementary school in Maine, Stires (1998, personal communication) reported the evolution of ST traditions initiated
by her students and maintained by subsequent groups of students. Her purpose for the ST activity was the development of formal oral presentational skills. She even encouraged children to plan their scheduled presentations at home. Stiles could observe the evolution and peer transmission of this innovation over more than 1 year because in her mixed-age classroom, older students passed on emergent traditions to newcomers. Within this framework, one child-sponsored innovation was the addition of a title to each ST narrative.

Reports from two other teacher researchers focus on the Discourse development and related peer relationships of African-American students who began the school year far behind their peers in both Discourse and social skills. Gallas (1992) and Griffin (1993), members of a teacher research seminar that focused on classroom discourse issues (Gallas et al., 1996), found that simply exposing students to a variety of discourse styles is not enough to cultivate the Discourse skills of all children. Teachers must actively work to make the discourse environment of the classroom welcoming to and respectful of students from all cultural backgrounds if the classroom is to afford learning opportunities for all children.

As researchers, teachers are in an excellent position to describe and interpret children’s discourse development over time. Because they come to know children as participants in an ecologically valid (Bronfenbrenner, 1979) discourse environment and are participants in this same environment, they may be especially sensitive to the nuances of children’s development as users of language. They may also be especially sensitive to reciprocal influences between narrators and their peers because concern for the children as a classroom community tends to be inherent in the teaching role. These reports also remind us that the multimeaning term culture refers not only to the memories and expectations that all classroom participants bring to their membership in this new face-to-face community, but also to the new patterns of behavior, interpretation, and expectations that evolve gradually within it.

LESSONS: TRADITIONAL AND NONTRADITIONAL

Traditional classroom lessons show a pervasive discourse pattern of teacher initiation, student response, and teacher evaluation (IRE; Mehan, 1979) or feedback (IRF; Wells, 1993). Even the traditional ST structure discussed earlier fits this pattern: The teacher nominates a child to share (I), the nominated child tells a narrative (R), and the teacher comments on the narrative (E/F). Linguists would call IRE/F the unmarked pattern of classroom discourse. In the 1990s, there was
a visible movement to change the structure of classroom discourse for a better fit with changing educational goals. Although these newer forms of lessons are called a variety of names, such as inquiry and discourse intensive, to contrast them with the traditional IRE/F structure, we refer to them simply as nontraditional.

Traditional Lessons

The most detailed analysis of traditional lesson structure is Mehan’s (1979) study of nine elementary school lessons in a San Diego classroom. According to Mehan’s analysis, IRE/F sequences are embedded in larger discourse structures that he called topically related sets (TRS). For example, in a social studies discussion about where the children and their families—all either African American or Mexican American—had lived before San Diego, the basic IRE/F sequence included determining each student’s or family member’s birthplace, locating that place on a map of the United States and Mexico, and placing that information on the map. Discussions of some birthplaces had only that basic sequence, whereas discussions of others extended beyond the basic sequence to optional conditional sequences about relative distance between cities and the basis of student knowledge. The combination of basic plus conditional sequences constitutes a TRS. In the lessons that Mehan analyzed, the teacher’s evaluations always occur at the end of sets, although not always after each student response within them. Thus, in Mehan’s analysis, the IRE pattern alone does not suffice to describe the structure of traditional lessons; it must be considered within the larger TRS structure. Other descriptions of traditional lesson structure report only the IRE/F level, and it is unclear whether they investigated such larger structures.

The TRS coding category is important to understanding relationships among utterances in a sequence of classroom talk. Topical relationships can hold across stretches of discourse within the larger TRS structure, termed reflexive tying structures in Mehan’s (1979) analysis. For example, the teacher’s question to one child “How did you know that?” and a child’s question about the location of Baltimore are not answered until eight or more turns later. For researchers, the discovery of such larger units is a reminder that the generalization that cognitive contingencies among mental units are not limited to behaviorally adjacent actions applies to discourse as well as syntax (Chomsky, 1957) and the performance of music (Lashley, 1961). For participants, this means that what counts as a relevant contribution to the classroom discussion may be more complex than if the preceding utterance alone were used as a guide.
Mehan’s work stands out now for his methods of coding and quantifying certain discourse features (see his coding as a heuristic device). For example, among 193 teacher elicitations in his corpus of nine lessons, only 8 (4%) were what he called metaprocess questions, which ask students “to formulate the grounds of their reasoning” (p. 46). Today such questions are termed metacognitive and receive much more attention for their potential benefits to student learning.

Whereas IRE/F triads and the larger TRSs are units of the syntagmatic (sequential) aspect of discourse structure, there is also a paradigmatic aspect of selection among alternatives within each functional slot. The IRE/F structure of traditional lessons constitutes a set of functional categories for teachers’ and students’ contributions. Within these functional categories, speakers select from a range of formal options. For example, the teacher’s question, “Can you come up and find San Diego on the map?” is understood without confusion as a command to act, not as a yes–no question, although it is interrogatory in form. The question is an example of what Sinclair and Coulthard (1975) and others call an indirect directive. In the routinized structure of traditional lessons, preconditions for the interpretation of such indirect directives can be derived from the culturally understood rights and obligations of classroom participants and do not need to be invoked separately (by researchers) or figured out (by students) for the interpretation of particular instances. Lemke (1990) found examples of even more complex form-function relationships in secondary school science lessons.

Like jazz musicians who use the conventions of the musical system as a point of departure for improvisation, participants in classroom lessons act on more than the general rules for classroom interaction. Erickson (1982) demonstrated the aptness of the musical analogy by using musical notation to analyze patterns of interaction in a primary school math lesson. His analysis shows that “what is happening is adaptive variation, specific to the immediate circumstances of practical action in the moment of enactment” (p. 178). In the words of Mehan and one of his colleagues (Griffin & Mehan, 1981), classroom discourse can be characterized as “negotiated conventions—spontaneous improvisations on basic patterns of interaction” (p. 205). Within the framework of traditional lessons, the metaphor of improvisation seems especially appropriate for illustrating the emergent nature of children’s communicative competence.

In a more recent analysis of traditional lessons, Wells (1993) responded to what he considered some misconceptions about the traditional IRE/IRF lesson structure and thereby represents a significant shift from analysis only of discourse structure to the inclusion of atten-
tion to pedagogical functions. Wells argued for *Feedback* rather than *Evaluation* as the label for the third part of the triadic pattern because the former term more accurately describes the range of teaching functions fulfilled by the teacher’s contributions in this spot. For example, in the review session, the teacher used the Feedback slot to reinforce the processes of doing science by drawing out the significance of the students’ own work. Wells also argued against the common criticism of the triadically structured traditional lessons that the teacher only asks display questions to which she already knows the answer, thereby expressing more of a testing than a teaching purpose. In the Time review session, the teacher knew the answers to the questions she posed, but in this case “the purpose [was] to establish an agreed account of what they did that will serve as an instantiation of the practice of controlling the independent variables to ensure that it is a fair test” (Wells, 1993, p. 27). What Wells called an “agreed account of what they did” is an example of the kind of *common knowledge* established through classroom discussion identified by Edwards and Mercer (1987). We can even reinterpret the teacher’s questions in the map lesson analyzed by Mehan. Their purpose was not to test students’ knowledge, but to get important information onto the shared public space of the map—information that the teacher did not possess.

Wells also analyzed features of a particular disciplinary register in the planning and review phases of a sequence of elementary science lessons. He singled out the teacher’s choice of particular words and phrases as examples of discourse as specialized register. Examples of discourse features that characterize the scientific register in these lessons include the term *experiment*, the technical question “What was the variable we held constant?”, and the complex noun phrase *release height*. She also used definition by means of a synonym. As she was writing a public record of the class’ scientific investigations, the teacher gave a more scientific meaning to the relatively commonplace word *fair* when she asked her students, “fair meaning consistent, right?” Whereas no children picked up these register features in their own talk during the planning session, a technical phrase “What we kept constant . . .” and the complex nominalization *release height* did appear in their talk in the later review session.

This aspect of Wells’ research shows the importance of tracking the development over time of children’s competence—in this case, the use of a specialized science register. In the same way, Mehan (1979) followed the development of children’s ability to use the bid-for-a-turn processes for their own initiatives. If other researchers also find a time lag between teachers’ initial seeding of specialized vocabulary into class discussions and children’s take-up of this vocabulary into their
own utterances, it would imply the value of staying with a curriculum topic and providing multiple opportunities for discussing its content in specialized language. A similar developmental sequence between comprehension and production has been noted in studies of young children’s language development in situations of repeated readings and talk about a single book (Snow & Goldfield, 1981, 1983).

Nontraditional Lessons

Wells’ (1993) analysis (also discussed in O’Connor & Michaels, 1996) shows that the same triadic IRE/F structure can serve a variety of pedagogical goals, yet some curriculum goals advocated in current school reform efforts require lessons that diverge more radically from the traditional patterns. One curriculum area where nontraditional structures have been especially encouraged is mathematics. Under the influence of reforms sponsored by the National Council of Teachers of Mathematics (NCTM, 1991), classroom discourse has become the essential and dynamic medium for accomplishing fundamental learning goals. In the words of one math education researcher (who is also a classroom teacher) and her colleagues, “Classrooms should be ‘mathematical communities’ rather than collections of individual learners” (Lampert, Rittenhouse, & Crumbaugh, 1996, p. 739). In lessons governed by this philosophy, teachers still ask questions, but the subsequent student and teacher turns depart from the typical triadic IRE/F pattern.

For example, in one of Lampert’s classes (Lampert, Rittenhouse, & Crumbaugh, 1996), the teacher presented four sets of paired numbers (e.g., 8-4, 4-2, 2-1, 0-0) and asked the class to state the rule for getting from the first to the second in each pair. When one student, Ellie, asserted that eight “minus a half” is four, her peers gasped in surprise, but Lampert understood that Ellie might be expressing a beginning understanding that “a number could be both a quantity [a half] and an operator on quantities [a half of eight] . . . an understanding on the boundary between arithmetic and algebra” (Lampert et al., 1996, p. 734). She engaged Ellie in an extended dialogue to draw out her reasoning without evaluating it and then encouraged other students to take the floor and give their response. This teaching strategy reflects NCTM’s call for classroom discourse in which students “listen to, respond to, and question the teacher and one another” and “try to convince themselves and one another of the validity of particular representations, solutions, conjectures, and answers” (NCTM, 1991, p. 45).

We have found no examples of research in these nontraditional mathematics lessons where the researcher has coded and quantified
such alternative patterns. One would expect, for example, that metacognitive questions would be far more frequent than the 4% reported by Mehan, but that remains an untested hypothesis. Perhaps the methodological contrast is because this more recent work has been done primarily by researchers with strong interest in mathematics education per se, rather than by a researcher like Mehan, a sociologist, whose interest is in communicative processes. Perhaps theoretical and empirical questions about the validity of coding and counting have made the finding and comparing of strong discourse patterns a casualty of changing conventions of inquiry. However, it may be that the discourse in these nontraditional classrooms is simply too complex to be captured analytically no matter who does the research and how.

These mathematics lessons also point to the problem of correlating studies of discourse with studies of children’s learning. In cases in which the teacher’s or researcher’s goal is the development of some aspect of students’ communicative competence, indicators of such development can be derived from transcripts alone. If some students fail to talk in multiple instances of analyzed lessons, that may indicate a need for further attention. Yet when the goal is substantive learning of, say, mathematics concepts, problem solving, and computation, transcripts must be supplemented with independent measures of such learning.

In all studies, in traditional and nontraditional lessons alike, it is important to keep track of the identities of speakers along whatever dimensions of difference are significant in the particular classroom. Among analyses of nontraditional mathematics lessons, for example, Lampert did not provide this evidence, although gender can usually be inferred. Ball and Wilson (1996), however, explicitly addressed issues of participation and mathematics understanding of children in Ball’s classroom who vary in ethnicity and the dilemmas that such differences pose for the teacher’s response.

**VARIATIONS IN DISCOURSE FEATURES**

Variations in classroom discourse patterns can be categorized according to particular features. Five features have been shown to be amenable to conscious variation by the teacher and have been the focus of study by researchers:

- Speaking rights and listening responsibilities
- Teacher questions
- Teacher feedback
- Pace and sequence
- Classroom routines
Speaking Rights and Listening Responsibilities

In traditional classrooms, there is an asymmetry between the teacher’s and students’ rights to speak. Given that the right to speak to anyone at any time inheres in the teacher’s role, we report research on how students obtain the right to speak during teacher-led activities.

The most common way is by teacher nomination or, less often, student self-selection (both part of Mehan’s [1979] analysis), but there are others. For example, teachers may establish a practice by which each student speaker calls on the next speaker. Sometimes yielding of absolute control by the teacher is not enough to ensure that all students do take or get a turn. For example, Chuska (1995) described how a teacher prepared a reluctant speaker to participate in advance by letting the student know what questions the teacher would ask him and when. By gradually scaling back this kind of support, the teacher helped the reluctant student develop into a confident participant.

Researchers have studied a range of influences on students’ willingness to speak, including those that are inherent in the discussion situation as well as those that participants bring to the discussion from their previous experiences. Wollman-Bonilla (1991) took advantage of the movement of one student from one reading group to another within the same classroom to study how ability grouping affected the length and quality of that student’s contribution to discussions about literature. Social class differences in talking styles can also affect students’ participation. Hemphill (1986) showed that in informal conversations, middle-class girls used conversational overlap to make a bid for the next turn in a conversation, whereas working-class girls used overlap to show support for the speaker. Hemphill’s contrast fits Tannen’s (1984) distinction between interruption and cooperative overlap and points to the importance of conducting a close qualitative analysis before starting to count any speech act category. There is a long history of research on cultural differences in preferred talking styles, which is discussed in a later section of this chapter.

There is more to speaking rights than simply getting a turn, however; there is also the question of which topics are picked up and developed and whose ideas thereby “get the floor.” One way in which students get the floor is through uptake by another speaker (Nystrand, 1997), analogous to the quality of semantic contingency well documented in the research on first-language acquisition.

In the nontraditional math lessons, teachers place special importance on how students hear and take up the contributions of their peers. Drawing on Goffman’s (1974, 1981) now-classic studies,
O'Connor and Michaels (1996) identified a strategy called revoicing, which teachers use to incorporate students’ contributions into the subsequent discourse. Goffman noted that when one speaks for oneself, one authors an utterance as well as speaking it. Yet one can also animate another’s utterance by respeaking it, perhaps altering it in some way for one’s own purposes. Revoicing is another name for Goffman’s animation.

In a detailed analysis, O’Connor and Michaels (1996) showed how a teacher’s revoicing moves served several purposes. At a minimum, such moves repeated a student’s contribution to rebroadcast it back to the group, giving that contribution added importance and attention. Sometimes the teacher reformulated the student utterance in the process of revoicing it for the sake of greater clarity, more complex conceptualization, or a more specialized register. At the same time, O’Connors and Michaels (1996) showed that when the teacher introduced her revoicing by the phrase, “So . . . lemme see . . .” or “So you’re suggesting . . .”, she maintained the student’s “right to evaluate the correctness of the teacher’s inference” (p. 82). In addition to these potential effects on the interaction with a single student, the teacher’s revoicing can position students in relation to each other by contrasting divergent statements to encourage further discussion. In these ways, revoicings can be a strategy for building a powerful community of learners.

Teacher Questions

Teachers and researchers alike believe that teacher’s questions can have a powerful impact on students’ learning and have therefore been the subject of many studies. As Duckworth (1981) pointed out 20 years ago, the teacher’s metacognitive questions such as “What do you mean?” and “Why do you say that?” call learners’ attention to their own thinking and knowledge.

Metacognitive questions are one type of a larger category often called open questions—questions for which the teacher, ostensibly, is not seeking one particular answer. However, several researchers have noted the difficulty of distinguishing between open and closed questions, a difficulty that presumably also faces participants. Barnes and his colleagues (Barnes, Britton, & Rosen, 1971) noted that some questions may be open in form but closed in function when subsequent discourse reveals that the teacher was in fact seeking a particular answer. Edwards and Furlong (1978) argued that it is impossible to assess the true function of a question without considering
it in the context of what is said before and after it: “[M]any questions which appear to be open are closed because of the context in which they are asked . . . or because the teacher has clear criteria of relevance or adequacy or correctness of expression to which he refers in evaluating the answers” (p. 41).

Nystrand (1997) called open questions authentic questions and identified them as one of three criterial dimensions in their construct of dialogic instruction. The other two dimensions are uptake (mentioned earlier) and the level of cognitive demand posed by the question. Although quantitative analysis of teachers’ questions in multiple classrooms showed no relationship between type of question and student engagement, a qualitative examination of the data from one classroom showed that frequency of authentic questions was not the only influence. Student engagement was higher when the teacher’s questions made higher cognitive demands.

Teacher Feedback

Revoicing is one kind of teacher feedback to student utterances. Other types have been identified by British researchers Tunstall and Gipps (1996) based on their observations in London schools for children ages 5 to 7. In a useful matrix that can be found in their article in The British Journal of Educational Psychology, Tunstall and Gipps classified 19 types of teacher feedback categorized across two overlapping dimensions. Within Evaluative types of feedback, they distinguished between Positive and Negative types. Within Descriptive types of feedback, they further distinguished feedback that focuses on Achievement from that which focuses on the need for Improvement. Also included in their scheme is a type of feedback for the purpose of socialization into “the values, attitudes and classroom procedures” that are an important part of primary teachers’ practice. This scheme represents a more complex continuum of feedback than the continuum of teacher responses to children’s sharing time narratives discussed in the first section of this chapter.

Tunstall and Gipps (1996) found that, “within evaluative types of feedback, judgments are made according to explicit or implicit norms,” whereas in contrast, “within descriptive types, feedback more clearly relates to [the student’s] actual competence” (p. 393). They argued that whereas evaluative feedback involves affective and cognitive dimensions of learning to an equal degree, descriptive feedback places a greater emphasis on the cognitive dimension. Although their typology was derived from observations of young learners who were often
working on individualized activities, they believe that their scheme can be applied to older learners and group work.

Pace and Sequence

The temporal relationship between teacher questions and student utterances and between student utterances and teacher feedback is another important source of discourse variation. Two aspects of that temporal relationship have been studied. The first and more microrelationship (a matter of seconds) is the interval between a student response and the subsequent teacher feedback. Rowe (1986) named this interval *wait time* and found dramatic beneficial cognitive effects from increased wait time in science lessons.

The second and more macrorelationship (a matter of minutes or even a day) is the timing of assistance more extended than a single teacher utterance. Ulichny’s (1996) study of conversations in an adult ESL class, for example, shows how a teacher used both immediate brief corrections and conversational replays to correct minor grammatical flaws in the students’ talk and longer instructional detours to teach a new, useful phrase to the whole class.

Classroom Routines

Making some elements of classroom activities automatic and routine can free up students’ mental energy to attend to higher order thinking. If students can be socialized into a set of activity structures that become familiar and predictable, yet open for improvisations in the moment and for development over time, both teacher and students can give more of their attention to academic content. Teacher researcher Rutherford (1992) described several routines that she used as the teacher of a sixth-grade class. These included: “Benchmarks,” whole-class discussions about a broad topic in which student teams select subtopics for research; “Research Rotations,” in which students rotated through different research-related tasks such as working at the computer, working at the library, and working in RT groups with the teacher to comprehend challenging texts; and “Jigsaw,” in which a child from each research team meets with students from other teams to teach their peers the results of their work in progress. Rutherford’s (1992) students were almost all bilingual, and her analysis suggests that the activity routines, with their predictable structures, “enabled the children to anticipate what was expected, thereby giving the stu-
dents more choice of how and when they would enter the discourse community” (p. 41).

The Discourse of Teaching Versus the Discourse of Testing

Although we recommend including independent measures of students’ learning to validate claims about how discourse patterns promote students’ learning and cognitive engagement, we must also be alert to the effect of particular assessment systems on classroom discourse patterns. Poole’s (1994) study of the influence of a mandated unit test on the discourse in a 7th-grade geography class is one attempt to track such influence. Not surprisingly, Poole found that in test review sessions, topic order was determined by the order of test items rather than logical connections among curriculum topics. She also found that in the test review sessions the teacher made bald unqualified assertions about the truth of historical events, whereas in the lessons he frequently used evidentiary markers (e.g., maybe) to qualify his statements. Given the frequency of tests in virtually all students’ lives today, researchers should try to understand more about their situational influences.

TALK WITH PEERS AND COMPUTERS

The meaning of the vague preposition with in this subhead refers to two different kinds of assistance: with peers means together with peers, whereas with computers means by means of computers. Despite these differences, working with peers and computers often go together in classrooms. Because only a few computers are available in most classrooms, students’ use of computers is, by necessity, usually collaborative. Moreover, in assisting peers, student computer expertise can supplement the limited availability of the teacher, and an inherent feature of the technology is that work in progress on the screen is public in a way that paper on a student desk is not.

In general, differences between learning in teacher-led lessons and learning in peer groups are becoming less marked. Teachers do not always maintain the stance and voice of authority, and exchanges among students are more frequent even in discussions led by the teacher. In previous sections of this chapter, in both the nontraditional ST and nontraditional lessons, students are expected to respond to what has been said by peers, and some teachers deliberately revoice one student’s contribution to a discussion to position students to speak directly to each other. This fluidity, rather than a sharp contrast between talk with the teacher and talk with peers, characterizes many of the
current innovations in classroom organization and participant structures.

We first discuss the activities in which peers talk together without computers and then activities in which students interact in various ways with this technology.

Talking With Peers

Common classroom activities in which students are expected to talk with their peers without a participating teacher include spontaneous helping, peer teaching, reciprocal critique, and collaborative problem solving. After examples of research in each category, we note the increased importance of students’ social relationships in the official curriculum.

Spontaneous Helping. Students of all ages regularly ask for help and give it to each other—regularly, that is, unless such help is outlawed as cheating in the behavioral norms expressed by the teacher or assumed by the students. Schlegel (1998) described a typical instance of spontaneous helping in a fifth-grade classroom in Los Angeles in which 90% of the students are Latino. In her description, children in a small group are working together to remember a word that describes something they saw on a field trip to the desert, a word search Schegel interpreted as an example of socially distributed cognition (or “socially shared cognition”; Resnick, Levine, & Teasley, 1991). Using their shared memories of the class trip and their common knowledge of popular culture as resources, the students help one child, Bonita, describe what she is trying to name. Another child, Alvaro, takes their collectively compiled description to the teacher, and the teacher’s suggestion (“Tumbleweed?”) is ratified by Bonita. This is an example of collective problem solving that originates with the students and in which the teacher plays a minor albeit important role. In his analysis, Schlegel emphasized the role of the children’s gestures and body positioning in communicating their ideas and reinforcing their social bonding. Both gestures and body positioning are features of interaction that deserve more research attention (see McNeill, 1992).

Peer Teaching. Another way in which students can request and give help to one another is through peer teaching, in which roles are assigned. Heath and Mangiola (1991) described such projects and report ensuing academic benefits to both tutees and tutors, and “a growth in the willingness [of the tutors] to speak out in class and take leadership roles” (p. 23). The ways in which success was measured in this pro-
gram highlight the importance of longitudinal measures for studying the effects of peer talk, as for the effects of any intervention or variation in traditional interaction patterns.

In the classroom whose lessons were analyzed by Mehan, instructional chains were set up for both pedagogical and research purposes in which a teacher taught a language arts task to one student and then asked that student to teach one or more peers. Cazden et al. (1979) analyzed (solely from the videotape record) the tutors' shift to a more formal discourse style as they took on the expert role. Carrasco analyses another tape from the same classroom in which bilingual Veronica shifted to crisp pronunciation of English words in giving a spelling test, in Spanish, to a Spanish-speaking peer.

Reciprocal Critique. The adjective social as applied to the role of discourse in learning has several meanings in the now dominant social constructivist theory of learning referred to in the introduction to this chapter (see e.g., Wertsch, 1985, 1991). One meaning refers to the influence of the discourse in which we engage when we interact socially with others. Another meaning refers to the mediation of tools that are themselves socially constructed, including physical tools such as computers and symbolic tools such as language. A third meaning suggests the orientation of a speaker or writer to the response she or he may receive from an audience.

In speaking, that response can be immediate, but in the activity of writing (or painting or creating a video), response often comes too late to be helpful during the creative process. So writers benefit from feedback given by someone else who can take on that audience role. One general name for the enactment of that role is critique. Like portfolio, the term critique comes to education from the arts. In addition to comments on the comprehensibility and interest of a text, it can include comments on more formal/aesthetic and functional/use criteria.

A familiar school activity for critique is the peer writing conference. Australian researcher Kamler (1980) showed how second-grade students benefit from peer conferences because the peer reader makes the idea of an audience visible and audible, and because the questions peers ask of each other's texts can be internalized into questions that writers come to ask of themselves as they compose. Soep (1996) reported similar benefits for teenage writers. Successful critiques or conferences do not happen naturally, however. As Atwell (1998) and Lampert et al. (1996) argued, students must be socialized into the conventions for speaking and responding to their peers in helpful and supportive ways. These authors drew on a variety of strategies, such as minilessons, modeling, and rehearsing, to gradually induct their stu-
Collaborative Problem Solving. Two common small-group activities in which students are expected to work together as equal-status collaborators are problem-solving groups in math and book-talk groups in language arts/English. In Cobb and Bauersfield’s (1995) book about math teaching and learning in one second grade, Cobb’s chapter presents four case studies of pairs of students who worked together for 10 weeks. Basing these case studies on observations, interviews, and videotapes, Cobb found that, at their best, these interactions were what he called *multivocal*, by contrast with the univocal dominance of a single perspective so often the teacher’s. Conducting a more detailed analysis of these videotaped data, Cobb and Whitenack (1996) found that peer discussions among these second-grade students could be categorized by four themes: social relationships, mathematical meanings, learning opportunities, and mathematical learning.

Social Relationships Among Students. In more traditional classrooms, students’ social relationships are external to the school curriculum. What is important is the relationship between the teacher and each student as an individual. Yet in all these peer activity settings, as well as in the nontraditional lessons, each student becomes a significant part of the learning environment of the others, and teachers depend on student contributions to the learning of their peers.

In Lewis’ (1997) study of literature discussions in a combined fifth-grade classroom, she discovered through interviews with focal students that complex peer relationships were being enacted in and were influencing the peer talk. For example, low-achieving Jason was not the only student sometimes reluctant to participate with his peers. High-achieving MacKensie, often asked by the teacher to lead literature groups, expressed trouble of a different kind, sometimes feeling freer to speak her mind when the teacher was present. Lampert, whose nontraditional math lessons were discussed earlier, also found powerful effects of and on peer relationships when she asked how her students felt about public disagreements.

There is no single answer to what influences the success, in interaction and learning outcomes, of peer group activities. The contextual influences are too varied. In a review of research on “group processes in the classroom,” Webb and Palincsar (1996) presented evidence for 13 possible influences, including ability and gender discussed by Lewis and the teacher’s role stressed by Atwell, Lampert, and Cobb. Peer collaboration is also another context in which researchers can trace the
quality as well as the distribution of speaking rights and listening responsibilities.

**Talk With, At, Through, and in Relation to Computers**

The various prepositions in this heading refer to the various ways in which computer technology can be incorporated into classroom learning. Students may interact *with* computers in one-on-one exchanges, they may interact with peers in sessions *at* the computer, and they may communicate *through* computers in telecommunication. Studying the impact of computer technology on classroom talk and students' learning also involves studying talk *in relation to* computers. In other words, where does the use of computers stand in relation to the larger picture of talk and learning? Our typology is adapted from Crook (1994). This book, along with Wegerif and Scrimshaw (1997), are two excellent British resources on discourse aspects of using computers for learning.

**Talk With Computers.** One of the first uses of computers in education was computer-assisted instruction (CAI) in which a solitary student interacts with some kind of computer software. Usually following the basic IRE/F model, such computer use has been widely criticized as merely an electronic workbook, a high-tech form for a low-tech function. As Mercer and Fisher (1997) noted, the nature of the feedback provided in CAI programs continues to be limited.

**Talk At Computers.** When two or more students work together at a computer terminal, interaction changes in a qualitatively significant way. A new component, *D* for discussion, can be inserted between the computer's Initiation and one of the students' Responses (Wegerif & Mercer, 1996):

I:  Initiation (by the computer)
D:  Discussion (between the children)
R:  Response (collaboratively constructed by the children)
F:  Follow-up (feedback and next initiation by the computer)

Such interactions at the keyboard can be analyzed in the same way as other peer collaboration with one exception. One division of labor is more definite because one participant sits in front of the keyboard and becomes the one to physically enter responses; during collaborative composition with word-processing software, one person is the pri-
mary typist. It is important to track how students negotiate their work at the keyboard and who assumes what role.

**Talk Through Computers.** Telecommunication—talk through computers—makes possible communication among students at a distance. These are most often class projects, not exchanges between individual pen pals. Although most of the reported projects not surprisingly exchange texts in the English language, a few are designed to encourage the use of other languages. The texts exchanged are necessarily written, yet they affect oral discourse as well, although this is less often included in available reports. Some of the writing, such as personal introductions of one class of students to others and brief comments and questions along the way, are the kind of oral/written hybrid that has become familiar to e-mail users. Because they are class projects, part of the regular curriculum, they bring enhanced opportunities for talk in relation to computers discussed later.

Some of the telecommunication projects described in published reports are by middle and high school teachers of English communicating through BreadNet, the electronic network for the Bread Loaf School of English (Middlebury College, Vermont). This network was established to connect classrooms in otherwise isolated rural schools. Christian (1997) was a teacher in one of these exchanges in which nine teachers from Alaska, Mississippi, Vermont, and New Mexico enabled their students to share their responses to *The Diary of Ann Frank*. Christian went on to document the ways in which the students’ writing grew as a result of the extended peer audience.

Teacher/researcher Sayers (Cummins & Sayers, 2000) has worked since 1985 with De Orilla a Orilla (From Shore to Shore), a network for encouraging exchanges across languages as well as cultures to encourage bilingualism and positive cross-language attitudes. One of the first projects involved an exchange between a class in Puerto Rico and a bilingual class in Connecticut that included recent immigrants from Puerto Rico as well as English-speaking children from Puerto Rican backgrounds. In the course of producing a bilingual newspaper, Sayers shows how the bilingual children in Connecticut became valued experts and the attitudes toward Spanish of their English-dominant peers became more positive.

More unusual is the use of an electronic network in the larger social project of revitalizing an indigenous language. The Leoki bulletin board system was developed at the University of Hawaii–Hilo as one component of its support for the Hawaiian immersion schools scattered among the islands. In these immersion schools, a seriously endangered heritage language is the medium of instruction for initially
English-speaking students (Warschauer, 1999). For these students, there is still little opportunity to hear, speak, read, or write Hawaiian outside of school. Thus, being able to exchange writing in Hawaiian among the schools can enlarge the audience significantly. Moreover, because computers symbolize the world of the future, they change students’ perception of the potentialities of the heritage language they are now learning.

A different use of telecommunication, more common in post-secondary education, is when it becomes the sole medium of instruction. Black, Levin, Mehan, and Quinn (1983) provided an early comparison of the discourse structures when the same class content was taught electronically and in a regular classroom. Some of the variations they found in the IRE/F structure could be predicted. Via telecommunication, students gave longer and more thoughtful answers to questions, they received more comments from their peers, and there was a longer time lag between Initiations and Responses. Less obviously, the electronic discussions pursued multiple threads of ideas rather than only one at a time, shifting the criterion of response relevance to the class material as a whole, and teacher evaluations were almost totally absent. As telecommunication courses become more common, such research merits further attention especially, it should be noted, when the computer replaces the labor-intensive research work of recording and transcribing. Moreover, this classroom/telecommunication comparison is important not only because it offers a glimpse of what may become a more common medium of instruction, but even more because it calls new attention to features of the too-familiar classroom.

**Talk in Relation to Computers.** In the telecommunication exchange projects described earlier, participants engage in much talk about the project away from, but in relation to, computers simply because each of the projects is central to the class curriculum for a limited period of time. Where computer experiences are not so central to the class curriculum, Crook (1994) warned against their “dislocation from a main stream of educational discourse” (p. 106). This dislocation, he argued, is less physical than conceptual, involving two kinds of learning loss. One is what he called *lateral loss*—the loss of transfer across learning contexts. The second is *longitudinal loss*—the failure to build what he called *intersubjectivity* and what Edwards and Mercer (1987) called *common knowledge*. Crook (1994) stressed the importance of longitudinal research through which such continuity and cumulative learning for a class as a whole can be documented: “The consequent achievements only become visible if we research beyond the moment-to-moment level of conversation; if we concentrate on more protracted
structures of social exchange” (p. 111). The importance of both lateral and longitudinal continuity for learning is in no way limited to computer use. The reason that Crook stressed the value of such continuity in discussing how technology can enhance learning is that computer experiences are particularly vulnerable to isolation.

**Differential Treatment and Cultural Differences**

*Differential treatment* and *cultural differences* refer to different perspectives on the single problem of achieving greater equity in students’ opportunities to learn. As they are frequently used, the terms refer to perspectives that contrast with each other. The differential treatment perspective usually refers critically to *over*-differentiation—to the ways in which schools and classrooms give some students greater access to more cultural, linguistic, and material resources, thereby reinforcing and even increasing the inequalities of knowledge and skills present when students start school. The cultural differences perspective, in contrast, usually refers critically to *under*-differentiation and asserts that some students would be better served if qualitative differences among students were taken into account to a greater, rather than lesser, extent. Continuing research and practice contribute to a more complex picture of both.

**Differential Treatment**

Although differential treatment has been documented at many levels of educational structures (see Darling-Hammond, 1995, for a review; Duke, 2000, for one case study), we discuss here only those structures that involve opportunities for oral participation that are differentiated in frequency or quality. The prior section on “Variations in Discourse Features” included discussion of research on the inequitable distribution of speaking rights and listening responsibilities. Here we focus on a form of classroom and school organization called *tracking* (or *streaming* in some countries) in which children are grouped for instruction homogeneously by achievement either within the classroom or within the whole school for part or all of the school day. It has a long history in both elementary and secondary schools.

In elementary schools, tracking within each classroom is especially common in reading groups. In reading research, differences in literacy instruction between low- and high-achieving groups have been shown to create a “Matthew effect” (Stanovich, 1986) in which “the rich get richer and the poor get poorer.” In a still earlier study of the teacher’s responses to children’s oral reading, Allington (1980) found three di-
dimensions of difference across 20 primary classrooms. Teachers interrupted the lower group readers at the point of error far more frequently (66% vs. 22%) than the higher group readers, they more frequently corrected even semantically appropriate errors (e.g., house for home) made by lower readers (55% of the time vs. 11% of the time for higher group readers), and they provided different kinds of decoding help to students in the different groups (more graphophonemic help to lower group readers, but more semantic or syntactic cues that drew higher group readers' attention to the meaning of the text).

Partly in response to research studies such as these, homogeneous reading groups were widely criticized in educational practice and replaced by whole-group instruction in many schools. As we write, they are coming back in popularity and rationale as part of balanced literacy programs, and grouping children by reading ability for short periods of time is receiving research support (e.g., Juel & Minden-Cupp, 2000). Such differential treatment needs to be carefully monitored for both instructional features and learning outcomes.

Arguably even more common and controversial than within-class reading groups in the elementary school is whole-school tracking within curriculum areas in high schools. Nystrand (1997) and his colleagues studied the relationship between tracking and the quality of teachers' talk in high school English classes. They found that students in higher tracked English classes ended up with significantly higher achievement gains at the end of the year even after they controlled for differences in students' knowledge at the beginning of the year. One explanation may be found in the types of tasks that lower track students were asked to do. Students in the higher tracks were asked to write more extended discourse, whereas lower track students were given assignments that were more like recitation and clerical work. These researchers also found that teachers asked more substantive questions about literature in the higher tracked classes, whereas in the lower track classes teachers' substantive questions tended to be about personal experience or off-task topics (Nystrand, 1997).

To avoid the inequalities found in homogeneous groups, heterogeneous grouping is widely advocated. Examples of positive reports on heterogeneous grouping include: Cohen’s long-term research on achieving equity in elementary school small-group science instruction (Cohen & Lotan, 1995), Reddy and her colleagues’ (1998) teacher/researcher reports of their implementation of Cohen’s ideas in an elementary science magnet school, and Rex and McEachen’s (1999) study of the latter’s high school English class. We still need more research, by academic researchers and teacher/researchers alike, that can show in detail how to make heterogeneous groups work.
In reading or conducting such research, it should be helpful to consider Lave and Wenger's (1991) construct of legitimate peripheral participation. Although they are not writing with classrooms in mind, their words speak to how heterogeneous grouping needs to be conceptualized by teacher or researcher. From their perspective, it is too much to expect that children who initially vary in levels of knowledge and skills, or in self-identities as competent students, will, when grouped together, immediately participate equitably—that is, without regard to those initial differences. Some students, because of previous experiences, may initially engage in more peripheral participation that should be considered legitimate. Over time, however, teachers can work for students' rates of participation and contributions to group tasks to become less associated with initial achievement status. At the same time, researchers need to observe longitudinally whether heterogeneous groups become dynamic sites with opportunities for all members to gain new competencies and new identities as legitimate, rather than marginal group members and students. Legitimate participants are those whom the group perceives as “potential members” (Holmes & Meyerhoff, 1999, p. 182). In contrast, marginal participants are those whom the group prevents from fully participating. This comparison of legitimate peripheral and marginal status comes from Holmes and Meyerhoff's (1999) introduction to a special journal issue exploring the relationship of Lave and Wenger’s (1991) concept of community of practice to related terms and theories in the study of language use.

Cultural Differences

The effect of cultural differences on students' school experience has been a challenging issue for researchers at least in the 30 years since the publication of Functions of language in the classroom (Cazden, John, & Hymes, 1972). Classrooms are places of increasing intercultural contact, and for many students, especially those in younger grades, they may be the sites of students' first intensive experience with cultures different from their own. Because there is documented variation in Discourse patterns across cultures, it is reasonable to suggest that generating a hybrid classroom discourse may help create more inclusive classroom environments. Yet documentation of successful hybrids is still hard to find, and controversy continues on their relative importance among other influences on students' engagement and learning (e.g., Jacob & Jordan, 1987). One now-classic success story is Heath’s (1982, 1983) decade-long work as ethnographer and teacher educator in low- and middle-class Black and White communi-
ties in southeastern United States. When White teachers in newly de-
segregated schools complained that rural Black children from Track-
ton did not participate in lessons, she helped these teachers under-
stand what she had learned from her fieldwork.

Heath realized that the children were not used to answering adult’s
known-answer questions about the labels and attributes of events. She
encouraged the teachers to become researchers and observe the ques-
tions they asked in their own homes and at school. She then helped
them design and try out a combination of new patterns of classroom
interaction and explicit teaching of the forms of Discourse they ex-
pected in school:

- Start with content and kinds of talk familiar to the children;
- Go on to new kinds of talk, still about familiar content, and pro-
vide peer models of children answering questions as the teacher
expected, making these models available later for repeated re-
hearing on audiotapes;
- Provide opportunities for the Trackton children to practice the
new kinds of talk, first out of the public arena and also on tape,
and then in actual lessons;
- Finally, talk with the children about talk itself.

In this work, teachers’ new understandings and changed practices
came from two sources. One source was what Heath could tell them
from her own fieldwork. Conveying such information is the more com-
mon role for discourse researchers. But alone, it carries the danger of
stereotyping and of making the teachers less observant of their stu-
dents and themselves rather than more. A second source, one less of-
ten mentioned in the many references to Heath’s work, is what she
helped the teachers learn for and about themselves through their own
research. In the upper grades, Heath and the teachers extended the lat-
ter, more situated type of learning by engaging students as core-
searchers about contrasting language practices. For example, they
compared the vocabulary used to describe farming in their textbook
and in their homes.

Ballenger (1996, 1999) provided an example of such situated learn-
ing in her observational records of her work teaching Haitian children
in a day-care center. Although proficient in Haitian Creole, Ballenger
was an outsider to the students’ culture. Ballenger was familiar with
Heath’s work, but knew that engaging in such long-term community
ethnography was impossible. By consulting with Haitian staff at the so-
cial service agency in which the day-care center was located, as well as
with Haitian parents in the Child Development class she taught, she
gradually learned to see the strengths that her children brought from their home culture. She also learned how to draw on those strengths as resources for managing classroom interaction and developing a literacy curriculum.

For example, Ballenger was accustomed to controlling children’s behavior by praising children who conformed and specifying her expectations and the reasons for them to those who did not. When her familiar patterns failed, she turned to the Haitian adults for help in learning a new language of control. Whereas she expected to justify her reprimand to the children for not waiting before crossing the parking lot “Because the cars are dangerous,” the children’s justification was “Because you like us” (Ballenger, 1999). To describe what happened here in terms of Tunstall and Gipps (1996) discussed earlier, Ballenger learned that, for these Haitian children, any specifying improvement reprimand had to include an affective personal expression as well. More generally, the usefulness of putative universal discourse typologies in any particular situation depends on knowing what counts as an instance of any category in the local culture.

What was studied by Heath in Trackton and Ballenger in the Haitian day-care center (and what most of the cultural difference literature is about) are what Ogbu and Matute-Bianchi (1986) called primary cultural characteristics. These authors defined such characteristics as “those cultural features that existed before two populations came into contact.” These features contrast with secondary cultural characteristics, which are defined as “those different cultural features that came into existence after . . . contact, especially in contact involving the subordination of one group to another” (pp. 96–97).

Ogbu and Matute-Bianchi attributed secondary cultural characteristics only to groups called caste-like minorities, such as Native Americans, Mexican Americans, and African Americans in the United States. The secondary experience of subordination is what children from all such groups share in the differential treatment described previously.

Yet one can extend the construct of secondary cultural characteristics to the dominant group as well (including the writers of this chapter). Nondominant groups may take on certain behaviors (like silence or resistance) and internalize certain self-perceptions (such as images of inferiority) from prolonged, sometimes multigenerational experiences of low-status contact. In the same way, members of a dominant group may adopt certain behaviors and internalize certain self-perceptions in response to prolonged, multigenerational experiences of high-status contact, more monocultural lives, and frequent White privileges.
One research study that probes these controversial issues at the intersection of differential treatment and cultural characteristics in the classroom comes from New Zealand. Alton-Lee, Nuthall, and Patrick (1993) studied the temporal and substantive relationships between official and unofficial levels of classroom discourse by having students wear individual wireless microphones and then transcribing their talk muttered to themselves or whispered to a peer in synchrony with the official lesson simultaneously in process. They reported encouraging news about the frequency and cognitive value of many of the muttered utterances. Yet in this classroom, the utterances also include potentially discriminatory or harassing talk that goes unnoticed by the teacher—talk that may be sparked or reinforced by the curriculum content.

Through their wireless recording equipment, the New Zealand researchers overheard how Ricky, the one Maori boy in the class of White, middle-class children, became the object of discriminatory remarks by White peer Joe sitting next to him. They traced the timing of those remarks to the teacher’s talk in the social studies lesson on the settlement of New York City. As the teacher shifts from “White people” to “we” in his first two comments, he conveys “an implied positioning of the children in the class and himself as ‘white people, Europeans, we’ ” (Alton-Lee et al., 1993, p. 74). Although there is no way of knowing whether the teacher’s words influenced Joe’s remarks, we can suggest that the teacher’s enactment of the curriculum through talk serves in this instance to reinforce particular attitudes in the wider society and legitimate their expression in school. When Ricky tries verbally to defend himself and Joe kicks him in return, the teacher becomes aware of some disturbance. His decision to reprimand Maori Ricky but not Joe has to be considered in this context another indication of biased effect even if unintentional and out of awareness.

**STUDENTS’ DISCOURSE DEVELOPMENT**

In the introduction to this chapter, we adopted Gee’s distinction between *discourse*, as any sequence of talk, and *Discourse*, as a particular form of talk that enacts and expresses a speaker’s identity and interactional role. We argued that new expectations for students’ Discourse should become part of the explicit curriculum. Most of this chapter is about research on classroom discourse, with a small $d$, but along the way we have pointed out instances of teacher’s expectations and students’ appropriations of forms of Discourse, with a capital $D$, that we hope become additions to (not substitutions for) students’ more familiar repertoires.
Examples of Discourse forms in the research cited here include more topic-centered ST narratives, the specialized register of science, and extended genres of explaining and even challenging both in teacher-led math lessons and in peer collaboration. Not included in this chapter is research on arguably the most noticeable Discourse difference—nonstandardized versus standard dialects of English. Three recent books cover this work well: Perry and Delpit (1998); Wolfram, Adger, and Christian (1999); and Adger, Christian, and Taylor (1999).

All teachers, whatever their formal curriculum subject, must now consider themselves also as mentors of Discourse socialization. Admittedly, this makes an already complex job even more complex. For readers of this chapter, it adds an additional focus for classroom discourse analysis.

REFERENCES


5. CLASSROOM DISCOURSE


This page intentionally left blank
There is no question that we now live in an age when information is as precious a commodity as gold or silver. Events occurring oceans away unfold in real time before our eyes, data that once took months to assemble and interpret can be analyzed within minutes, and the holdings of vast libraries can be carried in a briefcase or sent across.
space in less time than it took to type this sentence. As a consequence of these transformations in our postindustrial society, the fifth graders in Susan Pinker’s classroom learn in ways that past generations could not even imagine.

Yet how well prepared is this generation to tackle this veritable flood of information? What should teachers like Susan know to facilitate their students’ learning? Can these young minds make reasonable sense of their history textbook and acquire an understanding of the events and concepts it describes? What strategies do they need to navigate the vast ocean of information online? Do they understand or remember much of what they see and hear on videos or TV? What of the information they encounter in the form of classroom discussions? How do they go about capturing and retaining the ideas that Susan and other students are sharing with them on life in early Jamestown? These are among the critical issues that we consider in this exploration of learning from text.

WHAT CONSTITUTES TEXT

As may be apparent, we approach the construct of text in a somewhat nontraditional manner. That is, we do not restrict ourselves to a consideration of linear print typified by books or newspapers. Clearly, this textual form remains a centerpiece of human learning, and the source of most of the research in text-based learning. However, traditional linear text is but one segment of the literate world in which we function. The less linear, more dynamic, and more transient messages we encounter daily in discussions, as well as in the visual and online media (i.e., hypertext), are an increasing part of the text base. Thus, students must become conversant with all modalities of text if they are to learn effectively in the decades to come.

There is a range of definitions of text that could guide this review. The more conservative of these equate the term with composed, printed language intended to be processed linearly—that is, from left to right with requisite sweeps and stops (Caverly, Orlando, & Mullen, 2000). The more liberal conceptions, by comparison, consider text to be any artifact or symbolic representation as exemplified by Bloome and Egan-Roberston’s (1993) description:

People textualize experience and the world in which they live, making those phenomena part of a language system (broadly defined). The result of textualizing experience can be a set of words, signs, representations, etc. But it might be other forms and products not usually associated with texts: architecture, rock formations, the stars in the sky, the wind, the
ocean, emotion—these can all be texts, but their being texts depends on what people do. (p. 311)

As it pertains to learning within academic settings, the thrust of this chapter, our definition of text lies between these extremes. For our purposes, text represents the inscription of ideas in linguistic form. Although other modes of representation (e.g., numeric or spatial/graphic) often accompany text, these are not texts in and of themselves. It is the words accompanying such graphs or pictures, or the conversations they engender, that we hold as text. Certainly all textbooks, magazines, journals, tradebooks, videos, and computer sites in Susan Pinker’s classroom qualify as text under our definition. What also qualify are the discussions ongoing in that learning environment. In addition, any teacher- or student-generated materials that become part of the class content, such as Susan’s overheads and handouts and the products her students create, become a basis for text-based learning.

Salient Attributes

Texts that engender learning can be either formal or informal, oral or written, permanent or dynamic, individualistic or social, and sanctioned or unsanctioned. That is, text may be carefully composed for public airing or may arise spontaneously with little or any forethought. Texts may also be written down or shared orally. Moreover, these texts may remain physically unchanged as with the pages of the tradebooks or magazines in Susan Pinker’s room. Many of the texts from which we learn are of this type. They are the static artifacts that exist on shelves in schools, homes, or elsewhere. They may be in formation as when students craft and recraft the textual portion of their class projects or build on one another’s ideas during discussions or online.

We should note that some literacy researchers perceive all texts, even those that sit on library shelves, as dynamic (Berkenkotter & Huckin, 1995; Fish, 1980). These researchers argue that the roles, interpretations, or values of a text continually change as the society changes (Bakhtin, 1981, 1986). Many literacy researchers also see the process of learning from text as dynamic because it requires a continuing interaction between person and message (Nystrand 1986, 1989). Our research in persuasion and conceptual change provides support for this claim (Alexander, Murphy, Buehl, & Sperl, 1998; Murphy & Alexander, 2001). Whether participants were novices or experts, and whether the topic pertained to science or social issues, it was in the confluence of person and message variables that we saw change in knowledge, interests, or beliefs. In other words, it was in the interchange between person and text that learning occurred.
Finally, texts that become catalysts for learning can either be sanctioned or unsanctioned. Sanctioned texts are essentially the authorized or official linguistic materials that are part of the instructional environment. Despite the array of materials in play in Susan’s classroom, and despite the increasing presence of nontraditional forms, textbooks remain the primary form of sanctioned text in formal educational settings (Yore, 1991). Although they lie outside the realm of approved materials, unsanctioned texts also impact students’ learning. For example, Moje (2000) demonstrated how graffiti functioned as a powerful form of text for gang-connected adolescents. Similarly, Hartman (1997) found that texts valued in the broader African American community, like sermons, had no place in the educational setting. Unsanctioned texts may also include unauthorized books, magazines, Web sites, or messages that find their way into students’ thoughts and social interactions and, therefore, their learning. However, due to the variability and scope of unsanctioned texts, we restrict this examination to more authorized or sanctioned materials.

Dimensions of Learning

Learning, as conceptualized in this chapter, represents both a process and product. That is, learning is evidenced as students mentally, physically, or verbally engage in the activities at hand, as exemplified by the way Susan’s students listen to her remarks, pose questions, or pursue information online. Yet, engagement alone does not constitute learning. Learning necessitates some discernible growth in students’ knowledge, a deepening of their interests, and enhanced strategic abilities (Alexander, 1997b; Alexander, Jetton, & Kulikowich, 1995; Wade, Trathen, & Schraw, 1990).

Despite its many manifestations, the process of text-based learning can be described on the basis of common dimensions. For instance, learning from any text is inevitably intertwined with knowledge, both in terms of the knowledge one brings to the text as well as the knowledge one derives from it. Learning is also facilitated by an awareness of common text structures and features, a repertoire of general and domain-specific strategies, and positive motivational goals and dispositions. Granted these dimensions configure differently as one moves from traditional print to hypertexts or oral texts. They also configure differently within academic domains so that the texts Susan Pinker’s students read during history have a quality distinct from the texts they use in mathematics (VanSledright, 2002).

Still these dimensions of content and text-structure knowledge, strategies, and motivation are invariably present and critical to effective
text-based learning. For that reason, they serve as the framework for our exploration of texts that fall into three broad categories: traditional text, hypertext, and oral text. Of course much more has been written about learning from text when the source is rather traditional. Thus, in this chapter, we highlight learning under conditions when the text is delivered by hypermedia or orally. Still whether we are dealing with more or less traditional versions of text, knowledge, especially prior knowledge, remains foundational to the learning process. Consequently, we consider its general role first in this discussion.

Knowledge

Of all the factors relevant to text-based learning, none exerts more influence on what students understand and remember than the knowledge they already possess (Alexander, 1996; Stanovich, 1986). Whether Susan Pinker’s students are reading in their textbooks or surfing the Web, their knowledge is at work facilitating their actions and guiding their thoughts. To be more accurate, we should say that there are knowledges at work because there is actually a family of knowledge forms that impact text-based learning. The matriarch of that family is prior or background knowledge. The term prior knowledge, in effect, represents individuals’ mental histories—their “personal stock of information, skills, experiences, beliefs, and memories” (Alexander, Schallert, & Hare, 1991, p. 317).

Our personal mental histories are as unique as our fingerprints because no two of us have had the same life experiences. Nor do we perceive, interpret, or internalize the same event in the same way. Prior knowledge, which is ever present, “serves as the foundation of all future learning by guiding organization and representations, by serving as a basis of association with new information, and by coloring and filtering all new experiences” (Alexander & Murphy, 1998, p. 28). This is true whether knowledge comes from print or hypertext, traces to sanctioned or unsanctioned sources, is accurate, real, and verifiable or flawed, imagined, or unscientific. Undeniably, humans are not the blank slates or tabula rasa that Locke (1894/1975) envisioned them to be. Nor can we ever truly divorce ourselves from our mental histories. Our knowledge base remains our perpetual guide in text-based learning.

Due in large measure to information-processing research conducted in the 1970s and 1980s and the more recent resurgence of interest in domain-specific learning, we understand much about the power of the knowledge base. For instance, we recognize that prior knowledge impacts what students comprehend and remember from text (Afflerbach & VanSledright, 2001; Anderson, Reynolds, Schallert, &
Goetz, 1977). In addition, it is tied to the judgments students make about the relevancy, accuracy, interest, or importance of what they read or hear (Alexander & Jetton, 1996; Murphy & Alexander, 2000; Wade, Buxton, & Kelly, 1999). Knowledge is also significantly linked to students’ willingness to engage in text-based learning and their level of strategic behavior (Alexander & Judy, 1988; Guthrie & McCann, 1997). This is true whether we are dealing with traditional linear text or more nonlinear hypertext. It is also the case whether the text is written or orally conveyed (Alexander, Kulikowich, & Jetton, 1994).

**Content Knowledge.** Under the overarching construct of prior knowledge, there are more specific states and forms of knowledge that shape the process of learning from text. In fact, based on their review of the literacy research, Alexander et al. (1991) compiled a litany of such knowledge constructs. For instance, one of the primary reasons students engage in text-based learning is to acquire knowledge about specific domains (e.g., history) or topics pertinent to those domains (e.g., Jamestown). The labels *content* and *subject-matter knowledge* have been applied to this particular category. Moreover, this content knowledge can exist in three states: declarative, procedural, and conditional.

Declarative knowledge, for example, is the *whats* of our understanding and can take the form of specific labels, facts, definitions, explanations, or descriptions (Ryle, 1949). This state of knowledge still constitutes the bulk of what students encounter in the educational experience, especially when that experience is more teacher directed as in a teacher lecture or teacher-controlled discussion. Further, the textbooks and curricular materials found in classrooms are repositories of declarative knowledge. Still there is a value to declarative knowledge that should not be overlooked. Certainly more competent learners possess more declarative knowledge than less competent learners, but quantity of knowledge is not the sole issue. The knowledge of competent learners is also more richly formed and more strongly interconnected around core principles—what others have called *principled knowledge* (Gelman & Greeno, 1989).

Ryle (1949) also categorized knowledge as procedural. Procedural knowledge, the *hows* of our understanding, is critical to efficient and effective functioning. If you consider the number of actions or procedures you engage in on a daily basis, you can begin to appreciate the role of procedural knowledge in any field or endeavor. Even as we framed this chapter, for example, we applied our knowledge of writing conventions as we summarized bodies of research, devised examples, and transitioned from one sentence or paragraph to the next. The utilization of these writing conventions or linguistic processes are indicative of our procedural knowledge.
Research in problem solving and strategic learning has also created awareness of a third state of knowledge that deals with understanding *when*, *where*, and *for what reason* knowledge should be brought into play. This has been christened conditional knowledge because it pertains to the conditions under which knowledge is used (Newell & Simon, 1972; Paris, Lipson, & Wixson, 1983). These conditions are fluid because they vary with the circumstances in the immediate environment, which are also dynamic. What makes conditional knowledge so complex is that it falls on students to read these situations effectively and bring their declarative and procedural knowledge to bear in some suitable fashion. The difficulty of achieving transfer for much of academic learning may well be a reflection of students’ inability to activate relevant conditional knowledge under appropriate situations (Marini & Genereux, 1995). Sadly, students are given little training in determining when and how to use the declarative and procedural knowledge they possess. Thus, they falter in their self-regulation whenever tasks or circumstances move away from the routine or habitual, making the probability of transfer even more unlikely.

**Text-Structure Knowledge.** One aspect of knowledge that is requisite to meaningful learning is text-structure knowledge. Experienced readers recognize that the texts they encounter are rife with linguistic cues and organizational features that can promote meaning construction. Still as with other forms of knowledge we discuss, text-structure knowledge has the potential to promote students’ learning from text. Whether that potential is realized, however, remains to be determined. That is why Goldman and Rakestraw (2000) distinguished between text structure as available versus usable. As they stated, structural cues “cannot be effective if readers lack the prior knowledge needed to recognize and interpret these cues” (p. 313).

In our research, we have focused on one particular aspect of text-structure knowledge—genre. The term *genre* denotes these recognizable frameworks for oral or written communication, distinguished by their purpose, structure, and content (Jetton & Alexander, 2001; Swales, 1990). Further, there are various genres that are of particular significance in that literature. Here we focus on three: narration, exposition, and mixed texts.

Narratives are essentially stories that consist of common structural elements, which we collectively refer to as a story schema or story grammar. The elements of these story grammars include characters, setting, plot, conflict, and resolution. The power of these common narrative structures to instigate learners’ expectations and predictions has long been established in the research literature (Graesser, Golding, & Long, 1991). It has also been established that many
younger or less able readers are often unaware of this underlying structure, the emphasis on story reading and storytelling in early reading instruction notwithstanding (Graesser et al., 1991).

Expository text, by comparison, is generally written to inform or persuade, rather than to entertain. Textbooks and teacher-created texts in the form of handouts or overheads are typical forms of classroom exposition (Jetton, 1994a). Unlike narration, however, which has an overarching prototypic structure, exposition may incorporate various structures, including descriptive, cause–effect, chronological, comparison–contrast, and problem–resolution forms (Anderson & Armbruster, 1984; Meyer, 1975; Meyer & Rice, 1991). Moreover, these text forms can shift from paragraph to paragraph.

In actuality, many texts in classrooms are neither purely narrative nor purely expository. Instead, they are some hybrid of the two aptly called mixed texts (Hidi, Baird, & Hildyard, 1982). Biography is a common example of mixed text in that it contains information about a famous person in the form of a narrative. Mixed texts are increasing-ly finding their way into school texts (Jetton, 1994a; Jetton, Alexander, & White, 1992). Hence, as Susan Pinker’s students read how Jamestown’s settlers struggled to survive the winter of 1609, they encounter a poem by Stephen Benet depicting this struggle. Later as they read about motion in their science books, they find a side commentary about Galileo and his struggles with church authorities, who imprisoned him for his scientific beliefs.

The research of Goldman and colleagues (e.g., Goldman & Varma, 1995) has demonstrated how critical knowledge of text structures, such as genres and organizational schemes, is to text-based learning. Not only does this knowledge trigger hierarchical connections, as students form a mental representation of the content, but it also aids in interpretation and knowledge retrieval (Goldman, 1997). In her research on argument, Chambliss (1995) also determined that knowledge of text structures helps students weigh the evidence authors forward and, thus, judge the validity of the information presented.

**Strategies**

Whenever we encounter barriers to understanding or performance and we purposefully and intentionally take steps to compensate for or overcome those barriers, we are behaving strategically. As a review of the psychological research ascertained, “the ability to reflect upon and regulate one’s thoughts and behaviors is essential to learning and development” (Alexander & Murphy, 1998, p. 31).

Strategies are essentially special cases of procedural, or what Ryle (1949) called how-to knowledge that can be either heuristic or
algorithmic in form. As we engage in the process of learning from texts, we are more apt to rely on heuristic rather than algorithmic strategies. In other words, we are likely to use general guidelines or procedures to facilitate our learning, like outlining or summarizing, than we are to use step-by-step, formulaic techniques that render a correct solution. That seems reasonable given that learning from text is an ill-structured problem for which some agreed-on outcome is unlikely (Frederiksen, 1984). Building a mental representation of early Jamestown based on textual information, in effect, is quite distinct from solving distance = rate × time problems.

One trend in the strategy literature that has relevance for text-based learning is a growing interest in task- and domain-specific strategies (Alexander & Murphy, 1998). Therefore, as we consider the nature of text-based learning, we must keep in mind that there are strategies more pertinent to one modality than another or one domain than another. When dealing with oral text, for example, Susan Pinker’s students must have a means of capturing and retaining the content they hear, just as they must acquire strategies for sourcing information online and for judging its veracity. Likewise, in history, students must be able to compare evidence from multiple sources and consider the descriptions and claims made in light of these sources. In science and mathematics, by comparison, they must develop strategies for translating linguistic information into some alternative representation (e.g., graphic or numeric) or vice versa.

Motivation

Even students who come to text with a relevant base of background knowledge and a sufficient repertoire of general and domain-specific strategies may still not learn. The reason is that these students lack the purposefulness or interest that drives them to reach into text to unearth a message—what Garner and Alexander (1991) termed the will and thrill of text-based learning. According to Bergin, Ford, and Hess (1993), motivation can be broadly understood as “the psychological process involved in the direction, vigor, and persistence of behavior” (p. 437). Motivation as a field seeks to understand the reasons, motives, or drives for human thought and behavior (Pintrich & Schunk, 1996). As it pertains to learning from text, several facets of the motivational literature prove especially enlightening, including the research on goal theory, self-determination, and interest.

Goals. All students in Ms. Pinker’s classroom, regardless of their ultimate success at learning, come to that instructional environment with particular intentions that guide the manner and extent to which
they engage in text-based learning. These intentions are reflective of students’ goals, which can assume one of four orientations (Meece, Blumenfeld, & Hoyle, 1988; Wentzel, 1991): learning, performance, work avoidant, and social.

Learning or mastery goals represent a student’s desire for competence or depth of understanding (e.g., Ames & Archer, 1988; Archer, 1994; Dweck & Elliott, 1983). Other goals are reflective of a student’s desire to be evaluated positively or judged well by teachers and peers. These are referred to as performance or ego-involved goals (e.g., Dweck, 1986; Nicholls, 1984; Nicholls, Patashnick, & Nolen, 1985). Still some students’ approach to text-based learning is more indicative of their intention to exert minimum academic effort (Meece et al., 1988) or engage for the purpose of pleasing others or doing what is socially responsible (Wentzel, 1991). The former group is said to manifest work avoidant goals, whereas the latter group is motivated by social goals.

What the motivation literature demonstrates is that students who evidence either learning or performance goals generally fare well in schools, although the depth of their learning may vary. Social goals have also been linked with student achievement in that students who act in socially acceptable ways tend to receive higher grades in classrooms than students who are more socially irresponsible. It is not clear, however, whether the favorable ratings they receive from teachers are actually reflective of greater learning. Not surprisingly, work-avoidant students, as compared with peers, do not necessarily garner as much from the learning environment. This suggests that teachers like Susan Pinker must be concerned that students, even those with knowledge and strategic ability, are purposefully and meaningfully invested in their learning and comprehend the relationship between effort and achievement.

**Self-Determination.** Self-determination has also been positively linked to student learning and a desire to seek knowledge. As a form of intrinsic motivation (i.e., a task performed because it is rewarding in and of itself), self-determination represents individuals’ beliefs that they have some say so or control over their learning (Deci, 1992; Deci & Ryan, 1987). Opportunity for choice is often considered one contributor to students’ self-determination that can be implemented in classrooms (Corno & Rohrkemper, 1985). Another promoter of self-determination in classrooms includes the downplaying of external reinforcers (Ormond, 1998). In other words, Susan Pinker should give her students chances to explore topics of their choosing (i.e., opportunities for choice), and she should not always link evaluations or rewards to all classroom activities.
Most important, self-determination is not conceived of as an either/or condition, but one that exists in varying degrees. Thus, although it may be easier to engender stronger feelings of self-determination during student-led discussions or independent project work, self-determination can be experienced even when learning centers around traditional texts or during teacher lectures or recitations. Moreover, self-determination does not require complete freedom of choice or the lack of any external controls. Therefore, students who feel they have an important part to play in class discussions can benefit from verbal exchanges that are not entirely student-generated or student-led.

**Interest.** The relationship between interest and learning has a long history in the literature as represented by the writings of James (1890) and Dewey (1913). It is also presently enjoying a resurgence among literacy researchers (Jetton & Alexander, 2001; Renninger, Hidi, & Krapp, 1992; Wade et al., 1999). In general, the term interest signifies “the processes by which the underlying needs or desires of learners are energized” (Alexander, Murphy, Woods, Duhon, & Parker, 1997, p. 128). An array of interest-related constructs has been examined in the text literature, including cognitive and emotional interest (Kintsch, 1980), involvement (Reed & Schallert, 1993; Reed, Schallert, & Goetz, 1993), and engagement (Almasi, McKeown, & Beck, 1996; Guthrie et al., 1996).

Much of the research in text-based learning in academic domains has focused on the concepts of situational and personal interest (Hidi, 1990; Schiefele, 1991; Wade, 1992). Situational interest has been characterized as a transitory, short-lived state instigated by the specific characteristics of an event or object within an immediate context (e.g., Albin, Benton, & Khramtsova, 1996). Schraw and Dennison (1994) have shown that various factors, including the comprehensibility of the text and background of the reader, predict how situationally interesting a given text will be perceived.

Whether a situationally interesting text produces deeper learning has been the focus of debates in the literacy research. This is certainly the case when those aspects of the text most enticing to readers are not core to the topic or domain at hand (Garner, Gillingham, & White, 1989; Goetz & Sadoski, 1995; Wade, Alexander, Schraw, & Kulikowich, 1995). Those concerned with learning from text should, at the very least, heed Dewey’s (1916/1944) warning that attempts to make something interesting may send the unintentional message that the topic or task is not inherently interesting or valuable.

Over time individuals also develop enduring and deep-seated interests that are not as susceptible to fluctuations in the given context.
These personal or individual interests typically emerge from repeated interactions with an object or topic. Further, they are often associated with an investment of time, a desire for competence, or a pursuit of knowledge (e.g., Albin et al., 1996; Hidi, 1990). Given that so much of text-based learning in schools is about examining ideas that are unfamiliar or novel, there are questions about the opportunities students have to seek knowledge in areas of personal interest (Alexander, 1997a).

With this framework in hand, let us consider how the dimensions of knowledge, strategies, and motivation play out in three arenas. Specifically, let us look first at some emerging issues and concerns regarding knowledge, strategies, and motivation as they relate to traditional, linear text. From there, we turn our attention to alternative modes of text delivered by means of hypermedia or through classroom discussions.

**TRADITIONAL TEXT**

Research that examines student learning with traditional text has been ongoing for decades. In fact, most of the prior reviews of learning from text have dealt almost exclusively with this form (Meyer, & Rice, 1991). What we generally know about the effects of knowledge, strategies, and motivation on text-based learning was, in fact, derived from research with traditional print (e.g., Anderson et al., 1977). Therefore, rather than revisit familiar territory here, we consider several recent trends or issues related to the roles of knowledge, strategies, and motivation in learning from traditional text.

**Knowledge**

The powerful influence of prior knowledge on what students remember and ultimately learn from printed text is certainly well established (Alexander et al., 1991). However, current emphases on the sociocultural nature of learning and on socially shared and socially constructed understanding have led educators away from views of learning centered on the individual mind (Reynolds, Sinatra, & Jetton, 1995; Sfard, 1998). That is, researchers have become increasingly interested in the processes rather than the products of learning, and they are more concerned with understandings created in social interactions, rather than with any individual memories or interpretations resulting from them (Rogoff, 1990).

Indeed some research communities (e.g., critical theory or postmodernism) cast school knowledge in the role of a handmaiden to
political and cultural authorities bent on oppressing those who stand outside the realm of power (Woods & Murphy, 2002). These differing voices and views of knowledge have contributed to the maturation of research on knowledge in the last decade (Alexander, 1998). The research now considers not only the knowledge of the individual or the one, but also the shared understanding of the many, and embraces a multidimensional and sociocultural perspective on knowledge and knowing that would have been unheard of decades before (Alexander & Murphy, 1998).

Still researchers studying the nature of knowledge and the role of knowledge in text-based learning have only recently begun to tap into the multidimensional character of knowledge. Relatively few text-based studies specifically target multiple forms of knowledge or consider their developmental nature. For example, Jetton, Rupley, and Willson (1995) investigated how content, topic, discourse, and strategy knowledge accounted for differences between good and poor readers in Grades 3 to 6. The researchers determined that these forms of knowledge operated differently. For instance, Jetton et al. found that content knowledge was generally a significant predictor of differences between good and poor readers for narrative texts, whereas strategy knowledge was a significant predictor of performance differences for expository text.

Similarly, the research program by Alexander and colleagues has investigated changes in individuals’ subject-matter knowledge, concurrent with shifts in their strategic processing and interest, as these individuals progress toward competence or expertise (Alexander, 1997b). To date, these studies have explored learning in an array of academic domains, including history, biology, physics, and educational psychology (Alexander et al., 1995; Alexander, Kulikowich, & Schulze, 1994; Alexander et al., 1997). Overall this body of work demonstrates how learning from text is much more than a matter of altering the quantity and quality of individuals’ knowledge, although this is a clear effect. Movements toward expertise also require a shift in learners’ strategic profiles—from the use of more surface-level strategies to those associated with deeper processing and the formation of a coherent mental representation. It also manifests itself as an increase in individual interest in the topic or domain and a concomitant waning in the reliance on transitory or temporal interest.

**Strategies**

Research has clearly ascertained the relationship between strategic processing and comprehension (Alexander & Judy, 1988; Dole, Duffy, Roehler, & Pearson, 1991; Pressley, Goodchild, Fleet, Zajchowski, &
Evans, 1989; Pressley, Johnson, Symons, McGoldrick, & Kurita, 1989). Yet it would appear that few teachers today are aware of these strategies, encouraging students to use such strategies, or modeling them for their students (Dole, 2000). This situation may be partly attributable to the fact that strategy instruction is difficult to implement effectively in classrooms (Allington & McGill Franzen, 2000). The problems that researchers and practitioners have had in replicating the results for popular strategy approaches like reciprocal teaching (Palincsar & Brown, 1984) serve as evidence of this difficulty (Rosenshine & Meister, 1994).

It is also apparent that other educational issues have taken center stage in schools, like the concern for high-stakes assessment, producing a decreased interest in text strategies. Further, it is the case that few studies have explored the complexity of strategy instruction in the dynamic and uncertain context of the classroom (Elmore, Peterson, & McCarthey, 1996). However, there is evidence that the explicit instruction of strategies through demonstration, guided practice, and independent practice improves students’ text-based learning (Pearson & Gallagher, 1983).

The problem is that the idea of teaching anything explicitly runs counter to some educators’ ontological views that effective teachers should only guide and never direct or specify what their students should learn (Edelsky, Altwerger, & Flores, 1991; Goodman, 1986). Those holding to this worldview do not necessarily deny the importance of strategy knowledge, but they believe that students can internalize strategies naturally as they participate in literacy activities that allow them to freely share their feelings and interpretations of text (Karolides, 1997). Even those who are not opposed to the concept of explicit teaching remain unsure about “what to be explicit about and when and how” (Allington & McGill-Franzen, 2000, p. 145). Thus, the field of reading is still at odds as to whether strategies should be taught explicitly or whether teachers should provide implicit instruction of strategies as the need arises in the classroom.

As a result of all these factors, strategy research has come to a virtual standstill. This is unfortunate because there is much more to learn about strategic processing of traditional text and how to instigate such processing in students. Thankfully, some research is underway that may spark new interest in strategic processing (e.g., Baumann & Ivey, 1997; Graham, 1990). For example, Guthrie and others (1996) have incorporated strategy instruction in their Concept-Oriented Reading Instruction (CORI) program by teaching students how to locate and search relevant texts. Similarly, Harris and Graham (1996) continued their work on the explicit teaching of strategies for improved
writing, particularly for students with special needs. Indeed, Harris and Graham spoke quite strongly to this issue of explicit strategy instruction by attributing many of the writing problems seen in the upper elementary, middle, and secondary grades to a lack of explicit instruction in earlier grades.

**Motivation**

Although research on motivation has resulted in insights pertaining to academic achievement, empirical studies related to specific motivational constructs have only recently come to the forefront in the reading literature (Guthrie & Wigfield, 1997; Oldfather & Wigfield, 1996). Faced with mounting evidence that students do not always read often or well, researchers have increasingly turned to non-cognitive determiners as plausible explanatory factors. Initially, these investigations focused on global traits such as extrinsic and intrinsic motivation or questions related to affect (e.g., why individuals liked or disliked reading or perceived certain reading texts or tasks favorably or unfavorably). More recently, researchers have investigated more specific motivational constructs (e.g., self-efficacy) in designing studies of text-based learning (Wigfield & Guthrie, 1997). Consequently, such motivational constructs as will (Corno & Rohrkemper, 1985), expectancies (Eccles, Wigfield, & Schiefele, 1998), and goals (Ames, 1992; Dweck & Leggett, 1988) have appeared with increasing regularity in the text literature.

One motivation variable that has garnered considerable attention is the construct of interest (Renninger et al., 1992; Wade et al., 1999). The literature on interest has sought to characterize and define interest, particularly text-based interest (Hidi, 1990; Renninger et al., 1992), and examine the relationship between knowledge and interest (Alexander, Kulikovich, & Jetton, 1994; Tobias, 1994). Wade and her colleagues (1999) explored text characteristics that were positively or negatively associated with interest. They found that characteristics such as importance/value, unexpectedness, readers' connections, imagery/descriptive language, and authors' connections were positively related to interest. Other characteristics such as lack of adequate explanation/background information, difficult vocabulary, and lack of coherence and credibility were negatively associated with text interest.

In addition, various researchers have investigated the relationship between interest and importance (Wade, Schraw, Buxton, & Hayes, 1993). Several studies concluded that seductive details (i.e., highly interesting but unimportant information) can detract the reader from the more important information in the text (Garner, Alexander,
Gillingham, Kulikowich, & Brown, 1991). Our work (Alexander, Jetton, Kulikowich, & Woehler, 1994; Jetton et al., 1992) and that of others (Hidi et al., 1982; Schellings & van Hout-Wolters, 1995) has also pinpointed a possible disconnect between importance and interest in texts. Such disconnect may be associated with the text genre.

Specifically, in the case of narration, there is evidence that important information (e.g., characters, setting, or plot) is frequently aligned with the elements of the story grammar (Hidi et al., 1982). This does not tend to be true for mixed or expository texts, however (Alexander & Jetton, 1996). For mixed and expository texts, segments judged as important within the text structure (i.e., structurally important) often receive low ratings of interest from students. Teachers can exacerbate this problem. That is, we have found that teachers often give less structurally important but more intriguing content relatively greater emphasis in their lessons than those textual segments that are quite important but rather dry, complex, or boring (Jetton & Alexander, 1997; Schellings, van Hout-Wolters, & Vermunt, 1996). This kind of disconnect between importance and interest has significant effects on the concepts that students come to see as important for that instructional setting (i.e., instructional importance). However, we still understand little about how teachers choose the content to be shared during instruction or the long-term effects this pedagogical pattern may have on students' learning.

**HYPERTEXT**

If there were any questions about the presence of information technologies in text-based learning, the production of this chapter would have put them to rest. Not only did we prepare this text digitally, but we also searched the Web for relevant information, communicated with research colleagues via e-mail, downloaded research papers from Web sites, ordered several pertinent volumes via the Internet, shared text segments electronically, and engaged in online editing. It is likely that the students in Susan Pinker's classroom perform many of these same activities both in and out of school. As Postman (1993) aptly noted, it is not a matter of *if* technologies will change our lives, but rather *how*.

Despite the ubiquitous nature of information technologies, we are only beginning to understand how the process of learning from hypermedia unfolds. It is encouraging to note that research in learning from hypertext, which was rather slow to take hold in the 1980s, is now growing by leaps and bounds. This expanding literature affords important insights about hypertext, including an awareness of the way its structure, content, and processing influence the learning process.
Hypertext is not actually one text, but many texts that are digitally produced and electronically linked (Bolter, 1991). Further, each unit or page of hypertext can contain a variety of linguistic, graphic, pictorial, and auditory information, animation, interactive segments, as well as multiple links (paths) to other units or informational nodes (destinations). The entire World Wide Web is, in effect, one massive, interconnected hypertext that poses almost infinite possibilities to learners in terms of reading goals and navigational paths. Nelson (1974) coined the term hypertext in the 1960s as a marker for any text to be processed in a nonsequential or nonlinear way. For this reason, some researchers (ourselves included) refer to this form of text as nonlinear (Alexander, Kulikowich, & Jetton, 1994; Gillingham, Young, & Kulikowich, 1994) or multilinear (Bolter, 1998).

As the words nonlinear or multilinear imply, there is a dynamic, fluid, and multimedia character to hypertext (e.g., Nix & Spiro, 1990; Purves, 1998). This character becomes both the strength of hypertext, as a learning tool, and its weakness. Almost anything you want to read about or explore can be accessed through the Internet, from the hottest topic to most obscure subject. Moreover, hypertext can maintain a level of currency that traditional texts cannot because they can be instantaneously updated or revised (Garner & Gillingham, 1998).

In addition, teachers and students do not have to work for multiple interpretations or perspectives when using hypermedia. It is simply an artifact of the vast storehouse of information that can be found in hyperspace—information that is often competing or conflicting in view (Hawkins, 1996). There is also a collaborative nature to learning from, or better with, hypertext that does not occur as spontaneously with traditional text (Garner & Gillingham, 1996; Goldman, 1997). This may arise from the online exchange of ideas through e-mails, bulletin boards, or list servs, from shared problem solving, or from the co-authoring of texts or other hypermedia products (Anderson-Inman & Reinking, 1998; Cognition and Technology Group at Vanderbilt [CTGV], 1996; Scardamalia & Bereiter, 1991), or simply from the fact that students, like those in Susan Pinker’s classroom, must often share technology.

Yet this dynamism, fluidity, and multirepresentation come at a price. Simply stated, those with the knowledge, strategic ability, and motivation have much to garner when learning from hypertext, whereas those with some problem in one or more of these areas can surely find themselves lost in hyperspace (Alexander, Kulikowich, & Jetton, 1994). In other words, the experience of learning from hypertext can be
enriching if you (a) know how to navigate this informational labyrinth, (b) know how to cull the relevant and accurate from the trivial and erroneous, (c) have a desire to learn, and (d) know how to weave this information into some viable mental and physical representation.

Certainly information technologies have done more to reshape conceptions of text and learning from text than any other event in the recent history (Lemke, 1998; Reinking, 1998). The conception of learning from text that involves one person processing one text has given way to images of multiple people interacting around multiple texts (Goldman, 1997). Further, from the standpoint of text structure, we cannot talk so easily about text genres (e.g., narrative) or text forms (e.g., cause–effect) as we have with traditional text. Instead, we are confronted in hypertext with genres and forms that are hidden or implied (Gillingham et al., 1994). Some would even argue that hypertext lacks any discernible structure (Burnett, 1993; Landow, 1992). To Burnett (1993), for example, hypertext is “explicitly non-sequential, neither hierarchical nor ‘rooted’ in its organizational structure, and therefore may appear chaotic and entropic” (p. 1).

At the least, the common structure of hypertext is mixed both in terms of the confluence of text genres and forms it contains and the multimedia it infuses. This raises serious concerns, however. Specifically, we know that understanding of text structure enhances student learning when the medium is traditional text (Coté & Goldman, 1999; Graesser et al., 1991). We also know that an awareness of text structures is limited among students and teachers (Anders & Evans, 1994). That raises questions about what students and teachers know about the structure of hypertext, what knowledge they should acquire, and, in turn, whether such knowledge translates into better learning.

In Landow’s (1992) judgment, teachers and students should develop an understanding of the contours of hypertext. Those contours include an understanding of how the links and nodes function, what Landow (1992) called the rhetoric of departures (i.e., where a link may go) and the rhetoric of arrivals (i.e., what readers do when they come to the new location). Especially until such time as prototypical forms in hypertext become recognizable, teachers and students need to consider the distinctive features of hypertext, including its modifiability, dynamism, navigability, and collaborative nature (Anderson-Inman & Reinking, 1998).

Despite this intriguing character of hypermedia, Kamil, Intrator, and Kim (2000) reminded us that the literature investigating the cognitive consequences of reading hypermedia remains limited. Moreover, available findings differ as to the relative merits of learning from hypermedia versus traditional texts. Part of this ambiguity rests in
the difficulty of making clear and meaningful comparisons between these two modes. For instance, as Bolter (1998) and others have noted (e.g., Gillingham, Young, & Kulikowich, 1994; Goldman & Rakestraw, 2000), nonlinear processing may pertain more to the movement from one informational node to another. Once learners arrive at a particular Internet site, they may revert to a rather traditional, linear processing of text. In general, this serves to remind us that much more needs to be learned about the relationship between the structure of hypermedia and text-based processing if we are to maximize whatever potential exists.

Content Knowledge

What of the content found in hypertext? How does this content impact hypertext learning? In view of the fact that there is unlimited information available through hypermedia and anyone can contribute to that existing information base, the issues of relevance, credibility, and authenticity must be raised. For example, as Bolter (1991) and others (Purves, 1998) have pointed out, so much of what makes its way into the World Wide Web has no clear place in classrooms. Thus, part of the problem in managing learning environments that include hypermedia is regulating without stifling the flow of information. This, of course, is much easier said than done.

When the medium is traditional text, there is at least the perception of greater control over content flow and a stronger sense that the unseen writers of these texts are authorities who are to be respected (Garner & Gillingham, 1998). For instance, Luke, de Castell, and Luke (1989) claimed that textbooks have unrivaled status as the rightful sources of school knowledge. What topics or concepts are legitimized through the curricular process and relevance of that topic or concept to student learning often goes unquestioned with traditional text. This is not the case with hypertext, however, where sense of authority is eroded (Purves, 1998). That is, if students can choose their own paths to follow through hypertext and can contribute, modify, or delete text at will, then what does that say for the unquestioned status of hypertext? Even when the text is a posting on the Internet that cannot be directly manipulated, an individual’s response can be similarly posted and read by an audience as extensive as the original. The overall result is an undermining of textual authority (Landow, 1992). As compelling as these claims appear to be, much more specific evidence needs to be acquired before we clearly understand the nature of textual authority not just in hypermedia, but with regard to traditional text as well (Murphy, Long, Holleran, & Esterly, 2000).
One conclusion we reached in our examination of linear and nonlinear text processing is that the ability to make determinations as to what is relevant versus irrelevant or important versus tangential with regard to some topic or domain is predicated on certain factors (Alexander, Kulikowich, & Jetton, 1994). Specifically, the more familiar students are with the topic or domain under study and the more personally interested they are in that topic or domain, the more capable they are of separating pertinent from nonpertinent and important from nonimportant information.

Those with stronger knowledge or higher interest are also more likely to engage in information sourcing. Sourcing is a procedure that experts in history use to ascertain the credibility or potential biases of document authors. Let us assume that the students in Susan Pinker’s classroom search the Web for information on the early colonies. They come to several nodes with quite different stories about early Jamestown. One of those sites includes the reproduction of the diary from Jamestown’s governor during the winter of 1609. Another is a synopsis of the Disney movie “Pocahontas.” If Susan’s students engage in sourcing, they would ask themselves critical questions about these two sites and the information included there. For example, which is more likely to provide a realistic versus fictionalized or romanticized account of this period, is there conflicting information presented, and what perspectives are the authors of these texts apt to assume? Regrettably, although experts evidence sourcing almost 100% of the time, students rarely do so (Goldman, 1997; Wineburg, 1996). There is some reason for optimism, however, in that students can improve in this sourcing ability when provided with explicit instruction (VanSledright, 2002).

**Strategies**

Beyond the procedure of sourcing, there are other strategic abilities associated with learning from hypertext (Lawless & Kulikowich, 1996). Among these strategies are interpreting and translating information across multiple media (Flood & Lapp, 1995) and creating an organized and integrated representation of the resulting information (Goldman, 1997). Precisely because hypertext entails an integration of linguistic, pictorial, graphic, numeric, and auditory content, it requires a multimedia literacy. That is, students learning from hypertext must be able to function effectively and fluently within multiple representational systems and move back and forth across those systems (CTGV, 1996).

Some researchers even argue that the nonlinguistic elements of hypertext may take precedence over the linguistic components. For
example, Reinking (1998) believes that hypermedia signals a shift away from alphabetic prose to more pictorial representations. Similarly, Bolter (1998) stated that, “Literacy in electronic environments may have more to do with the production and consumption of images than the reading and writing of either hypertextual or linear prose” (p. 7). Regardless of whether these judgments hold up under empirical scrutiny, it is likely that the processing of hypertext will require a new form of literacy that incorporates this multimedia configuration. CTGV (1996) referred to this as representational literacy.

With this representational literacy, however, comes increased risk of difficulties in learning from hypertext, particularly for certain school populations. Those populations of at-risk learners include those typically labeled in this manner (e.g., students with limited English proficiency or those identified with cognitive, physical, or socioemotional problems). They also include students with limited exposure to information technologies inside and outside the classroom (Tierney & Damarin, 1998). One approach to assisting students who face potential problems learning within a hypermedia environment is through the use of supported texts (Anderson-Inman & Horney, 1998). Supported texts are electronically altered hypertexts that include a variety of resources or prompts presumed to assist readers with certain processing needs. For instance, these supported texts may incorporate glossaries or speech synthesizers that aid with reading fluency, prompts or questions to stimulate strategic processing, and supplementary graphics, illustrations, or text to extend the topics discussed.

Anderson-Inman and Horney (1998) determined that students who availed themselves of these supports in a thoughtful manner recorded the highest levels of comprehension when involved in hypertext learning. The question, however, is which students use such online resources wisely and what can be learned from their performance profiles. In terms of traditional expository text, Alexander and Kulikowich (1994) found that those with limited knowledge and linguistic abilities were those least likely to demonstrate strategic use of text-processing supports. If this pattern holds for hypertext, we would expect supported texts to prove more viable for those with more knowledge of the content, more strategic ability, or more interest in the subject.

One benefit of hypermedia is that it can make students’ strategic moves somewhat transparent (Bolter, 1998). For example, computers can maintain a record of the paths students travel through hyperspace, the sites they visit or revisit, the amount of time they spend at certain sites or tasks, and the resources they employ. The computer, in effect, can construct a history of students’ strategic actions. This allows teachers and researchers to explore students’ processing
profiles in relation to the understanding they ultimately demonstrate. Anderson-Inman and Horney (1998) did just that by monitoring at-risk middle-school readers’ moves through supported texts. What these researchers found were six processing profiles. Two of these profiles (i.e., skimming and checking) involved only a superficial scanning or surveying of hypertext; two profiles represented some, but not extensive, involvement with the text and its supports (i.e., reading and responding); and two indicated in-depth exploration of the content and appropriate resources (i.e., studying and reviewing).

There is one concern for strategic processing in hypermedia environments that Bruce and Hogan (1998) insightfully raised—the ultimate disappearance of technology. This is not to imply that technology will go away in our lives, just the opposite. Hypermedia will become so ingrained in our day-to-day lives that we will become less aware of its existence and more insensitive to its effects on learning. Based on this projection, we have in effect a window of opportunity to explore individuals’ interactions with this tool for better or worse. Beyond that point, as Bruce and Hogan stated, there will be less of a tendency to perceive the tool as the source of learning problems and an increased propensity to blame the user. Yet this also raises the intriguing question as to what elements learners will eventually come to rely on as hypermedia becomes more transparent than apparent. In effect, with the disappearance of hypermedia, will methods for organizing information become more content-based and structural cueing become more critical, rhetorical devices?1

Motivation

Since they began to find their way into homes and schools in the 1970s, computers have been strongly linked with motivation. Students manifest a fascination and skill with technology that many older individuals cannot fathom. Further, students can literally spend hours online cruising the Web, exchanging e-mails, conversing in chat rooms, or playing games. When making a case for hypermedia and learning, researchers and educators often use words like self-determination, choice, interest, and stimulation to capture the motivational qualities on hypermedia. Yet how do these qualities translate into student learning from hypertext?

This is a question that Alexander and Wade (2000) recently asked in their examination of select after-school and out-of-school programs that sought to build on students’ interest in technology as a mechanism for

---

1The authors thank Susan Goldman for posing this and other intriguing issues.
enhanced learning. Writing the commentary to a special issue of *Computers in Human Behavior*, Alexander and Wade offered their impressions of several of these technology programs, including KLICK (Zhao, Mishra, & Girod, 2000), Computer Club (Gillingham & Youniss, 2000), Fifth Dimension (Gallego & Cole, 2000), Project SEARCH (Bruce & Bruce, 2000), and the Odyssey program (Grosshandler & Grosshandler, 2000). Although their analysis pertained specifically to those after-school or out-of-school programs, several of Alexander and Wade’s concerns warrant consideration here. First, the more that these computer programs placed academic learning in the foreground, the harder they had to struggle to maintain students’ interest and engagement. Students often wanted to participate in these after-school and out-of-school programs to chat online or play video games and not to learn science or some school content.

Second, there seemed to be a misperception among some project directors that self-determination required virtually complete freedom for students. Therefore, any effort to direct or specify parameters was viewed as antithetical to self-determination, thus dampening students’ motivation. Such fears, according to Alexander and Wade, seemed unfounded because students can still experience self-determination when there are parameters or guidelines established as long as they feel that they are, to some extent, in control of their learning. In fact, it might be argued that younger or less experienced students might benefit when there are reasonable parameters placed on their use of hypermedia and their navigation of the Web. Thus, Susan Pinker might teach her students how to judge the suitability and relevance of appropriate Web sites and show them techniques for making intelligent moves through hyperspace. Ample opportunities for choice and self-determination still remain.

As a means of exploring the relationship between learning and motivation in hypermedia, Lawless and Kulikowich (1996) used cluster-analytic techniques to examine the knowledge, interest, and strategic profiles of undergraduates learning educational psychology via a hypermedia system. The researchers uncovered three performance profiles reflective of learners’ motivations: knowledge seekers, feature explorers, and apathetic hypertext users. Knowledge seekers were those students who seemed to have a goal of learning the content well and who, indeed, succeeded at that goal. Their navigational patterns showed that these students delved deeply into the online system, using information resources and support features often and purposefully. Feature explorers, by comparison, invested quite of bit of their time on the special features in the hypersystem and relatively less on the pertinent informational pages. The authors attributed some of this
behavior either to a lack of understanding of the computer environment or a failure to discern which components of the hypermedia system would prove most facilitative to learning. The third cluster, as their label implies, were unmotivated to learn. Meece and Holt (1993) would probably classify these students as work avoidant in that they simply did not care to learn and showed little, if any, evidence of engagement.

Overall, there is still much to learn about the motivational nature of learning from hypertext, especially as information technologies become embedded in the culture and classroom (Bruce & Hogan, 1998). For now we believe that educators would be wise to build on students’ interest in hypermedia as they orchestrate the learning environment. That is, we agree with Goldman (1997) and others who recommend that hypermedia be put to intelligent use in the classroom by allowing students to apply such technologies to the solution of challenging and meaningful problem-solving tasks that cannot be as effectively accomplished with traditional texts or tools. Thus, in many ways, it would seem that motivation in learning from hypertext, as with traditional text, rests largely on making the processing enticing and personally relevant to students. Thus, we speculate that the mode of text might not significantly alter the basic issues regarding human motivation and learning, although such a conclusion demands empirical verification.

ORAL TEXT

Learning from text in classrooms is not relegated to traditional print or hypermedia. It also embraces verbal exchanges between teachers and students in the form of lectures and discussions. Lectures, which are essentially monologues, share many of the same characteristics as traditional linear texts with one primary difference—their lack of permanence. In contrast, discussions are socially constructed and involve multiple speakers whose ideas become interwoven into the resulting oral text.

Researchers have characterized discussions as the communicative events that occur within a sociopolitical environment (Alvermann, O'Brien, & Dillon, 1990). Others have more conservatively defined discussions as verbal interactions that must be longer than short words or phrases. They also hold that students must talk at least 40% of the time and interact with one another as they express multiple points of view (Dillon, 1981). Given this conservative definition of discussions, most of the verbal interactions that teachers and students report having in their classrooms would not be considered discussions. In our research (Jetton, 1994a, 1994b; Jetton & Alexander, 2001), we
borrowed from Martin’s (1985) description of conversations by defining discussions as verbal interchanges among students and teachers who converse in a cooperative and constructive manner to listen and learn from one another. According to this view, these interchanges provide opportunities for students to gain knowledge that may not be shared in a printed text.

As researchers explore how students learn from the oral text in the classroom, they acknowledge that learning occurs when participants interact through language (Vygotsky, 1962). Language becomes the mediation tool by which individuals transform or modify “the stimulus situation as part of the process of responding to it” (Cole & Scribner, 1978, p. 14). Thus, it is through this social process that individuals internalize meaning. Discussions in which teachers, as more knowledgeable others, guide students to higher thinking processes can augment understanding by promoting advanced planning strategies, encouraging verbal interaction, and providing verbal instruction necessary to learning (Ellis & Rogoff, 1986; Radziszewska & Rogoff, 1988). Similarly, peer-led discussions in which students freely exchange ideas can be catalysts for enhanced learning by allowing students to grasp alternative perspectives (Almasi, O’Flahavan, & Arya, 2001).

**Text-Structure Knowledge**

Just as there are discernible patterns in traditional text and contours in hypertext, there are genres that are indicative of oral text—genres that serve to specify the role of teachers and students and influence the quantity and quality of subject-matter information shared. Further, the discussions that occur in school have a character that is unique from communicative practices in the home and community (Gee, 1990).

For some children, the discursive culture of the home and community is highly facilitative of the communicative practices in the school classrooms. For example, children who recite the alphabet or read along with their parents are familiar and comfortable with the recitation patterns of classroom discourse. However, children who learn by observing others in the home rather than by actively engaging in knowledge construction may not understand the form or purpose of the discussions that take place in classrooms. For these children, the tone, rhythm, and register of communication outside of school clashes with the discourse of the academic community, and they may struggle to gain access to the power of literacy afforded others (Gee, 1990; Hicks, 1997).

Therefore, students need to be apprenticed in the nature and forms of discourse that arise in the classroom if they are to optimize their
learning (Shaughnessy, 1977). One type of knowledge that students require is an understanding of the structures of oral texts that take shape in classrooms. Various researchers have explored the forms or structures of oral text indicative of literacy and content-area instruction (e.g., Alvermann et al., 1990; Cazden, 1988; Lemke, 1990; Mehan, 1979; Rogoff & Toma, 1997; Wells, 1993). In our research, we investigated discussion patterns that unfolded in high-school science classes (Jetton & Alexander, 2001). Through this examination, we identified four discussion genres that were similar to the structures described by other researchers: teacher-controlled, student-shared, student-led, and teacher-scaffolded.

In the teacher-controlled genre, the teacher dominates classroom talk by asking students questions and verifying their responses. The result is a back and forth exchange between teacher and students. Although close in form to lecture, the teacher-controlled structure does incorporate some student involvement, but in an abbreviated form and typically at a rather superficial level. This genre is comparable to the recitation (Alvermann et al., 1990), Initiation-Response-Evaluation (Cazden, 1988; Mehan, 1979), and Initiation-Response-Follow-up structures (Sinclair & Coulthard, 1975; Wells, 1993) described in the literature.

For the student-shared genre, there are frequent student-to-student interchanges, with minimal interjection by teachers. This oral text form seemed intended to extend or elaborate on a skill or content already taught by the teacher. Further, we observed this pattern most often during small-group activity where students worked to complete some assigned task. This pattern is comparable to the peer discussions documented in other studies (Almasi, 1995; Almasi et al., 1996; Almasi et al., 2001). In the student-led genre, by comparison, one student assumed leadership, functioning as a surrogate teacher for the group. Typically that student leader helped to frame the group’s focus and determine the topics of the resulting oral text. Several popular pedagogical approaches, most notably reciprocal teaching (Palincsar & Brown, 1984) and the Jigsaw approach, are based on this mode of verbal interaction.

The last discussion genre we identified was the teacher-scaffolded genre. Like the teacher-controlled genre, the teacher-scaffolded genre usually occurred during whole-class instruction. However, in contrast to the teacher-controlled genre, the teacher’s role in the teacher-scaffolded genre was that of facilitator. That is, the purpose of the discussion was to facilitate student discourse by first asking students to provide knowledge they constructed about a particular concept and
to explain their reasoning for those constructions, followed by an open
discussion of ideas among students. The teacher only intervened by
clarifying or extending student knowledge when the students reached
an impediment to further knowledge construction.

Content Knowledge

Part of becoming competent in any domain is learning the language of
that particular community. Every field of study—from history to
mathematics—has its lexicon and language conventions that mark
someone as a participating member (Ball, 1993a; Cobb, Stephen,
McClain, & Gravemeijer, 2001; Lampert, 1998; Murphy & Alexander,
2000). Consequently, discussions in content area classrooms not only
help students learn history, mathematics, or science, but also help
students learn to talk history, mathematics, or science (Lemke, 1990).

Without question, it is beyond the scope of this chapter to examine
the rich literatures on content domains such as history (e.g.,
VanSledright, 2000; Wineburg, 1996), mathematics (Ball, 1993b;
Cobb, Yackel, & Wood, 1992), and science (Lampert, 1998; Lemke,
1998) in any depth. Rather, we look at general issues of oral text that
cross these content domains.

Further, as we examine the role that knowledge plays in students'
learning from the oral texts of discussions, we must acknowledge that
the knowledge shared through verbal interactions is critical to the
knowledge that individuals eventually internalize. Yet few of those
engaged in the study of discussions have considered the nature of
the information shared during these verbal exchanges. One exception
is the work of Jetton and Kulikowich (1998) who studied the
relationship between discussion genres and the quantity and quality of
knowledge shared.

Using multidimensional scaling that considered the level and form
of information contributed, Jetton and Kulikowich (1998) detected
rather distinct differences with respect to the content shared by
discussion genre. For example, they found that the teacher-controlled
genre was rather barren with regard to knowledge. For this genre,
declarative knowledge took center stage, and students relied heavily
on the printed text as the source of that declarative information.
Moreover, as might be expected, the factual information contributed
was most often in the form of single words or simple phrases, and
there was little cohesion across the textual segments.

For the student-shared genre, the construction of a student-created
text became the focal point. To construct this oral text, which had a more
cohesive structure, students relied on not only declarative knowledge, but also procedural and conditional knowledge as well. Further, because many students helped construct this text, the sources of the content knowledge varied, including textbook materials, teacher lecture, and personal experiences. Procedural knowledge was most often in the form of strategies the students invoked to extend or clarify the emerging content, whereas the conditional knowledge often dealt with the validity or importance of the information students contributed.

With the student-led genre, Jetton and Kulikowich (1998) found a rich exchange among the student leader and discussion participants. However, there was a clearer delineation in the contributions of group members to the knowledge base. Specifically, the leader was primarily the source of declarative knowledge, and the leader and group members employed procedural knowledge in the form of strategies to think more deeply about the concepts under discussion. Finally, during the teacher-scaffolded genre, both teacher and students were contributors of the content knowledge that became the basis of the oral text. Both teachers and students referenced the emerging oral text to extend their understanding of the topic or domain. That is, as one student offered information, another student challenged or questioned the information in that student’s oral statement. The teacher helped facilitate student learning when students reached an impasse or needed guidance in problem solving. Thus, the students and teachers were sharing declarative knowledge through the oral text, employing procedural knowledge in the form of strategies that clarified or elaborated information, and utilizing conditional knowledge to consider why and when the knowledge should be used to reason through a problem-solving situation.

**Strategies**

Strategies for learning from oral texts may appear similar to those used to learn information from traditional printed text and hypertext. For example, students summarize information or predict the information that might be offered next. However, the difficulty in employing strategies for learning from lectures and discussions is that the information provided in the oral text lacks permanence. An oral text is not an artifact, such as a traditional text in that the words remain for the student to review or reread. Nor does oral text result in a navigational map of the path students follow in their learning as can occur in the processing of hypertext. In effect, no permanent artifact results from oral text that can be revisited by students or teachers.
This fleeting nature of oral text makes it imperative that students employ strategies to capture that knowledge.

**Grounding or Anchoring Oral Text**

One of the ways teachers and students deal with the transitory nature of the oral text is to ground it within or anchor it to a more permanent artifact, either one that is provided in the form of a book or outline or one that is constructed in the form of notes or journals (Palincsar, 1999). The printed artifact helps students retain the information by providing a framework for the verbal discourse. For lectures or teacher-led discussions, an advanced organizer either as a handout or on an overhead projector can give students some sense of the path that the oral discourse will follow. The teacher’s oral text thus extends or elaborates that framework, situating it in a richer subject-matter context of conceptual relationships and examples. While guiding students’ history learning, for example, Susan Pinker might refer them to a figure or graphic in their book or ask them to draw supporting evidence from the text. Once again, the figure or book becomes a touchstone to which the teacher and students can refer as they construct their oral texts.

The grounding or anchoring of oral text takes place whether the artifact is traditional text (Almasi et al., 1996) or hypertext (Bruce & Bruce, 2000). However, the purpose of this anchoring seems to vary depending on whether discussion occurs as part of literacy instruction or as part of a content lesson. Discussions during literacy instruction often center on narrative text and are seemingly more concerned with the process of constructing meaning than with any specific knowledge outcomes (Almasi et al., 1996; Eeds & Wells, 1989; O’Flahavan, 1989). What students take away from the text (i.e., efferent response) is pushed aside as the goal becomes engaging students in the oral and printed text through their aesthetic responses (Rosenblatt, 1985). The printed text becomes the focal or reference point for the discussion participants. Thus, students’ abilities to employ strategies such as providing support, evidence, or justification for their stated positions are sometimes dependent on the printed text as a tool by which these strategies can be enacted or addressed.

By comparison, one of the main goals of content area instruction is to deepen students’ understandings of the principles that undergird a particular domain (Alexander, 1997b; Gelman & Greeno, 1989). Discussions become the means for extending, reorganizing, and summarizing information garnered from a variety of sources so that students take away a more principled understanding.

Social Negotiation Strategies

Generally, strategies for capturing the knowledge of oral texts are different for lectures and discussions, in part, because of the purposes for each. Because a teacher’s lecture is primarily intended to convey information to a large group of students in a short amount of time, there is an unstated assumption that students restrain themselves from interjecting their own oral texts during the course of the lecture. For that reason, students pose minimal questions as the lecture unfolds because they would deter from the primary goal of the lecture. Many studies of teacher-controlled oral text support this notion (Alvermann et al., 1990; Jetton & Alexander, 2001). Instead, students who are actively processing information during a lecture employ alternative strategies for capturing fleeting message, including note-taking, creating a visual representation, or writing summary points. Clarifications must be sought at some more opportune time, as when there is pause or a request for questions.

During discussions, there are more strategic moves available to students because they have more opportunities to seek clarification or elaboration as the oral text is taking shape. In fact, one of the real benefits of a discussion, particularly over traditional text, is that students can strategically enter the discussion to resolve any misunderstandings or contribute pertinent information as the oral text unfolds. For example, as Jennie participates in a discussion in Susan Pinker’s history class, she becomes confused about whether the British colonists were friendly or hostile toward the Native Americans who lived around Jamestown. When this topic of the relations between the colonists and local tribes is raised, she strategically interjects a question so as to resolve her confusion. At this point, Susan and the other students focus on Jennie’s question, building the oral text around this theme. Thus, strategic moves, like those shown by Jennie, can be highly effective during discussions because there can be an almost immediate response to her query.

When and how students like Jennie enter into the discussion is also no simple matter. Much of the recent research has focused on these social negotiation strategies because researchers are interested in understanding the social processes involved in discussions that complement the cognitive processes involved in learning the content information. This literature sheds light on the strategies that teachers and students employ to control or sustain the discussion and to participate in socially acceptable ways. For example, during teacher-controlled discussions, students employ these social negotiation strategies by raising their hands when they want to contribute to the discussion. In contrast, during a peer-led discussion, students sustain
the discussion by employing social negotiation strategies such as turn-taking, eliciting a response from a student who had not previously participated, or initiating points for the discussion (Almasi et al., 1996). These social negotiation strategies become just as critical to students' learning as the cognitive strategies because they allow students to contribute effectively to the oral text that is in formation.

Almasi (1995) contended that the meaning of the oral text depends on the sociocultural contexts of the activity, as well as the cognitive processes of the participants. Her research focuses on social-cognitive factors during classroom discussions. In her analysis of the social negotiation strategies occurring during peer-led and teacher-led discussions, she conducted a sociolinguistic examination by indicating the manner in which discourse was initiated and sustained through questioning in the forms of chaining, arching, embedding.

Borrowing from Mishler's (1978) work on how questions serve to connect and extend discourse, Almasi (1995) described chaining as the strategy by which a participant's confirmation contains a question. She reported that a teacher uses this social negotiation strategy to link discourse together through a chain of questions, answers, and confirmations to maintain control over the discussion. In contrast, during the peer-led discussion, the participants employ either an arching or embedding strategy to sustain discussion. Through arching, a participant's response contains a question; through embedding, the participants provide two or more responses to a question.

In a later study, Almasi et al. (2001) examined strategies for attaining coherence during more or less proficient peer-led discussions. They examined ways in which students managed topics either through changing or sustaining the topic. They found students using strategies such as shifting to new or old topics, linking topics, and embedding topics by extending or elaborating on an existing topic. They also found that students managed group processes by using strategies for determining the parameters of what they could talk about and for discussing the interactive processes within the group.

Alvermann and Hayes (1989) also examined the social negotiation strategies involved in sustaining discussions. They listed the strategies that teachers and students used to lead a discussion along a planned course. These included asking critical questions that cause participants to respond and judge or evaluate one another's statements, and the intonations of the speakers that signal whether the discussion should continue or end. These researchers also noted that participants sustained discussion by clarifying through elaboration of examples or reinforcing a response through verbal and nonverbal reinforcement. Participants also engaged in strategies such as paraphrasing, defining
certain terms, and calling on others to participate. Alvermann et al. also noted that participants engaged in pacing strategies to determine the tone and speed of discussions.

Motivation

Most researchers and teachers recognize that learning from text is not a purely cognitive activity, but rather an orchestration of factors that work in synchrony as one engages oral or written text. One of those factors is motivation. Certainly the goals that teachers and students hold for participating in lectures and discussions influence their learning. The students in Ms. Pinker’s class, who evidence learning goals or performance goals, put forth extra effort to listen, take notes, and remember the content of her lectures or discussions. However, those students with work-avoidance goals may choose to copy the skeletal outline provided as a transparency and ignore most of the surrounding oral text. As Graves, Juel, and Graves (1998) stated, “Purpose is what motivates us, helps focus our attention, or gives us a goal, something tangible to work toward” (p. 228). To facilitate learning, students need to understand why they are listening to a lecture or participating in a discussion. These purposes must be pertinent to the oral text and appropriate for the students.

Alvermann and Hayes (1989) identified three distinct purposes for discussions in middle-school teachers’ classrooms. Teachers’ purposes included reviewing course content, helping students comprehend the reading material, learning definitions, identifying parts, and filling in missing gaps of information. Similarly, Jetton and Alexander (2001) looked more specifically at teachers’ purposes in relation to the particular discussion genres they employed in science classrooms. For example, Jetton and Alexander determined that teachers had several reasons for using a teacher-controlled discussion. One such reason was the teachers’ belief that this genre allowed them to dispense scientific information in an efficient and effective manner because they could cover a lot of information in a short amount of time.

The back-and-forth pattern of this genre also permitted teachers to engage a large number of students in the construction of the oral text. Other purposes seemed to arise from the teachers’ perceptions of themselves as content experts. As one teacher put it, having a teacher-controlled discussion allowed him to impart his scientific knowledge to the students. Other teachers noted that their purpose was to ascertain whether students understood the information they had acquired through a small-group activity. In other words, their purpose
was to review and check the validity of the information that students had constructed in their groups.

One of the stated purposes of having a student-shared discussion, by comparison, was to allow students to gather information on their own or solve a particular problem posed by the teachers independently. Another purpose was to engage students in open exchange of information and allow students to review previously learned concepts. When teachers chose student-controlled discussions, their purposes were more often to empower students by allowing them to be the authority over the content of the oral text, thus making the content more interesting or personally involving. This discussion genre was also seen to give students greater ownership of the information represented in the oral text. That is, teachers felt that students would be more engaged with the information if their students constructed the related text, rather than receiving it from them. These teachers’ views are consistent with that of various researchers who would argue that discussions in which students lead or share the oral text, with little or no teacher involvement, provide students with a greater sense of autonomy and control (Turner, 1995).

Among the purposes for teacher-scaffolded discussions was the activation of and reflection on background knowledge. The teachers in the Jetton and Alexander (2001) study stated that they wanted to facilitate student discourse, but not control it. Another teacher noted that her purpose for having this type of discussion was to help students think about information that they already knew before instruction began. Other purposes for employing this discussion were to explicate what students knew and did not know about the scientific information. Further, teachers stated that teacher-scaffolded discussions tended to inspire deeper thinking about concepts among students.

Several researchers attest that the social interaction that occurs as students use oral text to discuss content can lead to increased motivation. As the teachers and students introduce new ideas through the oral text, these ideas can spark curiosity and interest (Berlyne, 1960; Turner, 1995). The oral text of students and their peers can provide models by which students can gauge their own progress or thoughtfully evaluate the response of others (Brown, Collins, & Duguid, 1989). Recently, researchers examined how motivation works in tandem with one’s knowledge, strategic processing, and social interactions (e.g., Guthrie & Wigfield, 1997). Guthrie and Wigfield described engaged learners as those motivated to learn new knowledge, able to use prior knowledge to construct new meanings, strategic, and socially interactive as they share oral text with others to learn.
CONCLUDING THOUGHTS

In this review of learning from text, we explored the dimensions of knowledge, strategies, and motivation across three forms of text: traditional texts, oral text, and hypertext. The literature amassed around these three textual forms is indeed vast. Therefore, we recognize that we have only touched the surface of the issues and trends that exist. Thankfully, there are extensive reviews and informative studies that interested readers can pursue in each of these topical areas, and we encourage such a pursuit. Moreover, our understanding of the less traditional modes of text, in the manner of oral text and hypertext, will grow during the next 10 years not only with regard to their evolving nature, but also the roles they come to play in school learning. Whatever transpires in that next decade, we feel strongly that the dimensions of knowledge, strategic processing, and motivation will not fade in importance, but only configure differently in the processes of teaching and learning and the research those processes engender.

REFERENCES


Technology-mediated communication is a fact of life. The human communication apparatus is constrained in several ways. There are limits to the distance at which speech is audible, and visible behaviors such as gesture, gaze, or facial expressions are perceivable. Furthermore, these natural communication behaviors are ephemeral and do not persist over time. Given these limitations, we must rely on some form of mediation if we are to communicate at distance and across time. People have therefore invented technologies that attempt to circumvent these limits to allow remote synchronous communication (e.g., telephone, videoconference) or asynchronous communication (e.g., letters, telegraph, e-mail, fax, voicemail).

Understanding the principles that govern mediated communication has important practical implications. Because of the pervasiveness of mediated communication, we need to determine whether, how, and why it differs from face-to-face communication. If there are differences between mediated and face-to-face communication, we need to provide guidelines to inform decisions about the circumstances in which it is appropriate to use mediated communication and the effects of using it in those situations. Systematic insights into mediated communication should also enable us to improve the designs of current and future technologies supporting mediated communication. An understanding of mediated communication should also inform more general theorizing about the psychology of communication. Most current communi-
cation theories regard face-to-face communication as an integrated set of speech, gaze, and gestural behaviors. As we see, studies of mediated communication allow us to isolate the individual contribution of different nonverbal behaviors such as gaze and gesture to communication. They also help clarify the overall influence of visual information in communication and the effects of communication interactive processes such as feedback on communication.

The structure of the review is as follows. The main part of the review is organized around three main theoretical approaches to mediated communication. We first talk about the general characteristics of mediated communication theories and methods in this area. We then characterize each theory in detail and evaluate the evidence for it. We conclude with a discussion of outstanding practical and theoretical issues.

THEORIES OF MEDIATED COMMUNICATION

The current section outlines: (a) mediated communication theories, and (b) technologies used to support communication in the various situations. There are many different theories of mediated communication. Rather than describing each in detail, we restrict ourselves here to an exposition of the general form that these theories take and the general set of claims that they make. The fundamental goal of mediated communication theories has been to explain the relationship between the affordances (Gibson, 1966; Norman, 1988) of different mediated technologies and the communication that results from using those technologies. Most theories are comparative, addressing how and why mediated communication is different from face-to-face communication. The general form of these theories is to (a) characterize how technologies differ in terms of their communication affordances, (b) describe how the affordances of a given technology differ from those of face-to-face communication, and (c) explain how these different affordances produce differences between mediated and face-to-face communication in process, content, or outcome of communication. For example, social disinhibition theory (Kiesler, Siegel, & McGuire, 1984; Sproull & Kiesler, 1986) argues that e-mail has different affordances from face-to-face communication. E-mail is asynchronous so it cannot provide immediate interactive feedback to speakers about whether their communication has been understood or accepted. E-mail also does not provide nonverbal information such as head nods, gestures, facial expressions, and postural information that have been argued to be important to the transmission of social and affective information (Argyle, 1990; Short, Williams, & Christie, 1976). According to the theory, the absence of feedback and the lack of social and emo-
tional information has two effects. It: (a) inhibits social processes such as consensus formation, and (b) leads to stronger emotional expression in e-mail.

Table 7.1 shows the affordances of different mediated communication technologies as identified by these theories. One major technology affordance is the different *modes* that technologies support. The main difference here is between linguistic and visual modes. For example, e-mail, phone, and voicemail transmit purely linguistic information, whereas videoconferencing is a combination of a linguistic speech channel and a visual channel showing participants’ images. Another example of a technology that combines linguistic and visual information is Shared Workspaces, which combine a synchronous speech channel with shared visual material such as a documents or graphics (McCarthy, Miles, & Monk, 1991; McCarthy, Miles, Monk, Harrison, Dix, & Wright, 1993; Stefk, Foster, Bobrow, Kahn, Lanning, & Suchman, 1997; Tang, 1991; Tatar, Foster, & Bobrow, 1991; Whittaker, Brennan, & Clark, 1991; Whittaker, Geelhoed, & Robinson, 1993).

Note that there are no instances of pervasive technologies that are purely visual: we discuss reasons for this later. The second major affordance is *interactivity*. For example, the phone, instant messaging, and chat are interactive because they support synchronous bidirectional communication, allowing participants to provide immediate feedback to speakers. In contrast, e-mail and voicemail are noninteractive because they do not allow such concurrent feedback. Another example of a noninteractive technology is videomail, which allows people to leave messages that include both speech and pictures (Hopper, 1994). Noninteractive technologies are necessarily permanent because message recipients are not present when the message is generated. In contrast, interactive ones can be ephemeral (e.g., the phone) or permanent (e.g., instant messaging and chat). Although some theories (e.g., Clark & Brennan, 1991) propose more complex sets of affordances, modality and interactivity are sufficient to explain

<table>
<thead>
<tr>
<th>Affordance</th>
<th>Interactive</th>
<th>Noninteractive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode</td>
<td>Phone, audioconference, chat, instant messaging</td>
<td>E-mail, answeringphone, voicemail, FAX, letter, Usenet</td>
</tr>
<tr>
<td>Linguistic and visual</td>
<td>Videoconference, video-phone, shared workspace</td>
<td>Videomail</td>
</tr>
</tbody>
</table>
most of the research we report here. The conclusions section discusses the limitation of the current affordance taxonomy and describes how it might be enriched. Finally, in this review, we exclude technologies such as radio, TV, newspapers, and film; these technologies only allow one-way broadcast, and our concern here is with two-way interaction.

Having characterized technologies and their affordances, the theories explain how those different affordances affect communication behaviors. By determining how different affordances affect communication behaviors, mediated communication theories predict the effects of these behaviors on critical communicative phenomena involving communication content, process, and outcome. The relationship among affordances, behaviors, and their communicative effects are derived from studies of the role that different communication behaviors play in face-to-face communication. By identifying how visual behaviors and interactivity affect these behaviors, the theories make predictions about how communication processes, content, and outcome are altered in technologies that do not support visual modes or interactivity. For example, face-to-face research has shown how behaviors such as head nods, gaze, and gesture mediate turn taking (Argyle & Cook, 1976; Argyle, Lalljee, & Cook, 1968; Argyle, Lefebvre, & Cook, 1974; Beattie, 1978, 1981; Duncan, 1972; Kendon, 1967). However, each of these behaviors is visible and depends on the presence of visual information. Therefore, we might expect that technologies that do not support visual information would show impaired turn taking (Jackson, Anderson, McEwan, & Mullin, 2000; O’Conaill, Whittaker, & Wilbur, 1993; Sellen, 1992, 1995; Whittaker & O’Conaill, 1993, 1997). These relationships among affordances, communication behaviors, and their effects on core communicative phenomena are shown in Table 7.2. For example, the table shows that the visual mode affects multiple communication behaviors, including head nods, gaze, and gesture and that one core communication phenomenon affected is turn taking. Because all theories make predictions about the differences between different forms of technology, the table does not show the affordances of linguistic modes. Such information would be redundant as the linguistic mode is common to all technologies. We now describe three different types of mediated communication theory with respect to these two tables. We discuss in more detail how different affordances affect behaviors and hence communication process, content, and outcome. The remainder of the review is organized around these three theories: bandwidth hypothesis, cognitive cueing, and social cueing.

One early theory, the bandwidth hypothesis, addressed how modes affected communication. The bandwidth hypothesis proposed a direct
relationship between the modes that a technology supports and the communication that results from using that technology. The hypothesis posits that, regardless of task, the closer the set of modes supported by a technology approximate to those of face-to-face communication, the greater the efficiency of the communication using that technology. Therefore, we should expect a technology that supports both visual and linguistic modes should always outperform one supporting only the linguistic mode. Later we review evidence showing that the bandwidth hypothesis is false.

Subsequent theorizing has been more directly influenced by current psycholinguistic theory. A large number of recent studies have examined the effect of using mediated communication on various types of cognitive cueing (Clark & Brennan, 1991; Doherty-Snaddon, Anderson, O’Malley, Langton, Garrod, & Bruce, 1997; Isaacs & Tang, 1993; Kraut, Galleger, Fish, & Chalfonte, 1992; Kraut, Miller, & Siegel, 1996; McCarthy et al., 1991, 1993; O’Conaill et al., 1993; Olson, Olson, & Meader, 1995; Sellen, 1992, 1995; Whittaker et al., 1991, 1993; Whittaker, 1995; Whittaker, Terveen, Hill, & Cherny, 1998; Whittaker & O’Conaill, 1993, 1997). Later we review a number of hypotheses about the effects of visual information on communica-

### TABLE 7.2

<table>
<thead>
<tr>
<th>Affordance Type</th>
<th>Communication Behaviors Affected by Affordance</th>
<th>Core Communicative Phenomena Affected</th>
</tr>
</thead>
<tbody>
<tr>
<td>VISUAL MODE</td>
<td>Facial expressions</td>
<td>Attention, understanding, agreement</td>
</tr>
<tr>
<td></td>
<td>Head nods</td>
<td>Conveying affect, attitude</td>
</tr>
<tr>
<td></td>
<td>Gaze</td>
<td>Attention, understanding, agreement</td>
</tr>
<tr>
<td></td>
<td>Gesture</td>
<td>Turn taking</td>
</tr>
<tr>
<td></td>
<td>Visual access to objects in a shared physical environment</td>
<td>Turn taking, reference</td>
</tr>
<tr>
<td></td>
<td>Physical presence</td>
<td>Reference, attention</td>
</tr>
<tr>
<td>INTERACTIVITY</td>
<td>Feedback via backchannels, completions, interruptions</td>
<td>Attention, understanding, agreement</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Turn taking, reference, repairs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Socioemotional feedback</td>
</tr>
</tbody>
</table>
tion in cognitive tasks. Following predictions derivable from Table 7.2, concerning the relations among the visual mode, communication behaviors, and core communication phenomena, we examine whether technologies that do not provide visual information disrupt turn taking (Jackson et al., 2000; O’Conaill et al., 1993; Sellen, 1995; Whittaker, 1995; Whittaker & O’Conaill, 1997) or reference (Bly, 1988, Whittaker et al., 1993), as well as initiating impromptu conversations (Fish, Kraut, Root, & Rice, 1992; Tang, Isaacs, & Rua 1994; Tang & Rua, 1994). We also look at how noninteractive technologies disrupt turn taking (O’Conaill et al., 1993), reference (Oviatt & Cohen, 1991), and understanding for cognitive tasks (Kraut, Lewis, & Swezey, 1982; Oviatt & Cohen, 1991).

A second research area has been into social cueing processes in mediated communication. Here the focus is on the role of visual information in supporting the transmission of affective and interpersonal information, as well as the function of interactivity in providing social feedback. Again following predictions derivable from Table 7.2, we review the evidence for several theories that argue that the absence of visual information changes the expression of affect and its role in communication (Morley & Stephenson, 1969, 1970; Rutter, 1987; Short et al., 1976; Sproull & Kiesler, 1986). We also examine how the absence of such visual interpersonal information combined with reduced socioemotional feedback affects higher level social processes such as negotiation (Morley & Stephenson, 1969, 1970) and reaching consensus (Kiesler et al., 1984).

**METHODS IN MEDIATED COMMUNICATION RESEARCH**

Mediated communication research has employed a variety of methods: laboratory studies, field trials, interviews, surveys, and ethnographic techniques. Most of the data reported in this chapter come from laboratory studies where users are given predefined technologies, tasks, and instructions. Often such experiments collect both objective data (e.g., time to complete the task, quality of solution) and subjective data (how useful was technology Y? how easy was task Z?). Researchers also carry out content analyses of conversations using different mediated communication technologies. Clearly such experiments offer control and the ability to systematically manipulate relevant variables. However, other research reported here includes field trials of novel systems where experimental control is sacrificed for data about usage patterns and subjective reactions to technology used for extended periods of time. One critical element of such field trials is that the systems are used for people’s everyday work. Although such
ecological validity is highly valuable, it obviously reduces experimental
total control over the tasks for which the system is used. Nevertheless, such
extended experience with mediated communication systems is vital
given recent empirical research showing that many reported differ-
eences between mediated communication and face-to-face communica-
tion are vastly reduced after extended periods of technology usage
(Kelly, Futoron, & McGrath, 1990; Kelly & McGrath, 1985; Walther,
1992, 1994, in press). Other research has employed survey tech-
niques to examine reported patterns of usage for established technolo-
gies such as e-mail and voicemail (e.g., to understand reasons for using
these technologies in certain situations; Markus, 1987, 1994; Rice,
1993). A final (small) class of data reported here is ethnographic. Eth-
ognaphies employ descriptive, interview, and observational tech-
niques to generate fine-grained analyses of communication behaviors,
technology usage, and reactions to technologies (Nardi, Kuchinsky,
Whittaker, & Schwarz, 1996; Nardi, Whittaker, & Bradner, 2000;
Whittaker & Sidner, 1996). The weakness of such techniques con-
cerns their generality because ethnographic data often investigate only
small numbers of users. However, ethnographic data differ from other
methods in offering detailed qualitative findings about users’ tech-
niques and strategies for using various technologies. We now review
evidence for each theoretical approach.

The Bandwidth Hypothesis

*Hypothesis:* The closer the modes supported by a technology corre-
spond to those of face-to-face communication, the more efficient the
communication with that technology. Specifically, adding visual infor-
mation to speech should improve the efficiency of communication.

Early research in mediated communication focused on technology
modes. It took as its starting point the five human senses (sight, hear-
ing, smell, touch, taste) in face-to-face communication. Ryan and Craig
claimed that different technologies can be characterized in terms of
their correspondence with face-to-face communication, in which “in-
formation can be transmitted and received by any of the five senses”
(pp. 2–3; cited in Williams, 1977). In mediated settings, sensory input
is clearly impoverished. Technologies do not support taste or touch,
and visual information is often missing. According to the theory, this
should result in reduced efficiency of communication independently of
task. It also follows that technologies supporting multiple senses
should be more efficient than those providing access to few senses. We
refer to this as the bandwidth hypothesis. As we see, the work on the
bandwidth hypothesis has exclusively concerned the senses of sight and hearing by investigating the effects of adding visual information to auditory communication. There has been no work into the effects of other sensory modes.

The bandwidth hypothesis was disproved by work examining the effects of various combinations of communication modes on the process and effectiveness of mediated communication (Chapanis, Ochsman, Parrish, & Weeks, 1972; Chapanis, Parrish, Ochsman, & Weeks, 1977). This work showed that adding visual information to speech does not necessarily lead to performance improvements. It involved a variety of dyadic, synchronous communication tasks. Success was measured in terms of quality of solution and time to solution. The set of technologies evaluated was a permutation of face-to-face communication, remote speech, remote handwriting, remote teletyping, and video/speech. (The term video/speech is used throughout this chapter to refer to the combination of video information with speech. We prefer this to the term video, which can be ambiguous as it is often used to refer exclusively to the visual mode of an audio/visual communication). Communication efficiency was tested for 10 different technology combinations (e.g., speech with remote handwriting, handwriting with remote teletyping, etc.) and for a variety of cognitive tasks that required information exchange and problem solving. On one task, participants had to jointly construct a mechanical object, when one partner was given the unconstructed physical components, and the other, the instructions. In another task, one person had to identify a map location satisfying a number of criteria (e.g., the nearest dentist to a given street address). Again one participant had the map and the other the relevant selection criteria. Participants were at different physical locations (except, of course, when the communication condition was face-to-face communication) and communicated only with the technologies provided.

There were two main conclusions of these multiple studies. First, adding the visual mode to speech to increase bandwidth did not necessarily lead to more efficient communication. Face-to-face communication or video/speech was no more effective than speech-only communication for these tasks. These are striking results given that many of the tasks used in the experiments (e.g., finding a location on a map, jointly constructing a complex physical object) would intuitively seem to require access to visual information, and hence to favor mode combinations that transmitted visual data. The result is also of general significance because it shows that mediated communication is not always different from face-to-face communication. The second major conclusion was that speech is the critical medium for these types of task.
Mode combinations that included speech were always more efficient than those that did not (Chapanis et al., 1972). If participants used spoken interaction, the addition of high-quality video, text, or writing modes made little difference to task outcome or process. Even face-to-face communication was no different from speech in task outcome or solution quality. Furthermore, modes that included speech were two to three times more efficient than nonspeech modes.

These results showing little impact of visual information on cognitive problem-solving tasks have been replicated in many other laboratory experiments (Reid, 1977; Short et al., 1976; Williams, 1977). For example, Reid (1977) reviewed evidence from 28 different studies showing no effects for the addition of visual information to speech for cognitive tasks. This is not an issue of video quality: Even face-to-face communication is no better than speech-only communication for this type of task (Reid, 1977; Williams, 1977). These studies do, however, show some effects for the addition of visual information in tasks in which social cueing is critical. We review this in a later section.

Data from naturalistic studies also show that having access to visual information does not always improve communication when speech is available. A field trial of high-quality videophone conducted over several months showed few objective usage differences compared with the telephone (Fish et al., 1992; Fish, Kraut, Root, & Rice, 1993). The videophone supplemented speech with a 12-inch TV-quality image of the remote participants on the user's computer screen. Overall there were few differences between speech and video/speech for either logged usage or user perceptions. Phone and videophone calls had similar durations and were used for the same set of communication tasks. The researchers also administered a questionnaire asking people to state the tasks for which they felt that different communication technologies (e.g., videophone, telephone, face-to-face communication) were appropriate. Multidimensional scaling techniques applied to people's answers indicated that videophone is viewed by users as more similar to the telephone than face-to-face communication. Again this suggests that visual information adds little for this type of task.

In conclusion, there is little general evidence for the bandwidth hypothesis. A huge amount of data seem to support two complementary and counterintuitive conclusions. The first is that adding a mode does not increase communication performance: Specifically, visual information generally seems to contribute little to the outcome of a variety of cognitive tasks when speech communication is available. The second is the primacy of speech: adding or removing other modes has little effect compared with the addition or removal of speech (Chapanis et al., 1972, 1977).
Cognitive Cueing

The prior results offer strong evidence against the global claims of the bandwidth hypothesis. A more recent set of studies have argued the need to apply face-to-face communication models more directly to mediated communication (Clark & Brennan, 1991; Daly-Jones, Monk, & Watts, 1998; Doherty-Sneddon et al., 1997; Isaacs & Tang, 1993; McCarthy et al., 1991, 1993; O’Conaill et al., 1993; Olson et al., 1995; Tang & Isaacs, 1993; Whittaker, 1995; Whittaker & O’Conaill, 1997; Whittaker et al., 1991, 1993). Consistent with the approach illustrated in Tables 7.1 and 7.2, this work claims that instead of focusing on the five senses, we need to identify the contributions of various communication behaviors (e.g., speech, gaze, gesture, and backchannel feedback) in supporting core communication phenomena, such as turn taking, reference, and understanding. If we can determine how the affordances of various mediated communication technologies affect communication behaviors, we can make more precise predictions about how different types of technologies affect communication. Rather than focusing on general effects of modes, this work argues that we should specify precise mechanisms by which affordances affect behaviors, which in turn influence core communication phenomena.

These studies have tested three specific hypotheses about different functions of visible information on various cognitive aspects of mediated communication. These concern turn taking, conversation initiation, and the role of a shared environment. They are all directly derivable from the information presented in Table 7.2. The turn-taking hypothesis argues that visible behaviors such as gaze and gesture support turn taking in face-to-face communication. Removing these (e.g., in phone conversation) results in impairments in the ability to switch speakers. The initiation hypothesis evaluates the importance of visible information about the presence and availability of other conversational participants for impromptu communication. The hypothesis claims that initiation of impromptu conversations is impaired in technologies (such as the phone) that do not support this type of visual information. The shared environment hypothesis (McCarthy et al., 1991, 1993; Whittaker et al., 1991, 1993) maintains that a shared physical environment with mutually viewable objects and events contributes to conversational common ground (Clark & Marshall, 1981). The shared environment supports various types of conversational inferences as well as allowing certain types of reference such as deixis. The absence of the shared environment (e.g., when conversing by
(phone) makes such conversational inferences harder, reference less efficient, and conversations more explicit.

Other cognitive cueing work concerns interactivity. The *interactivity* hypothesis argues that technologies such as voicemail or e-mail that prevent immediate feedback via backchannels, completions, and interruptions decrease shared understanding and disrupt various reference processes.

In what follows for each hypothesis, we first characterize relevant behaviors in face-to-face communication and describe how they affect turn taking, availability, reference, and interactivity, respectively. We then predict how removing visual information or using noninteractive technologies affect mediated communication compared with face-to-face communication.

**Turn Taking**

_Hypothesis:_ Technologies that preclude access to visual information about gaze, gesture, and head nods disrupt turn taking given the demonstrated role that these behaviors play in managing transitions between speaking and listening in face-to-face communication.

There is a great deal of work elucidating the role of gaze and gesture in mediating turn taking in face-to-face communication. Speakers and listeners in face-to-face communication show different patterns of gaze, with listeners spending more time looking at speakers than vice versa (Argyle & Cook, 1976). In addition, gaze can mediate transitions. Speakers tend to look more at listeners as speakers draw to the end of their turn to await confirmatory information that the listener is ready to continue (Kendon, 1967). Gesture can also serve to coordinate turn-taking transitions. The termination of gesture acts as a signal that the speaker is ready to hand over the conversational floor, and it is therefore a turn-yielding cue (Beattie, 1978, 1981; Duncan, 1972).

The results are mixed for the turn-taking hypothesis. Sellen (1992, 1995) conducted a series of laboratory studies in which groups discussed contentious issues and tried to reach consensus. They communicated using speech, video/speech, or face-to-face interaction. As expected, there were differences in turn taking between face-to-face communication and speech. More important for the hypothesis, however, contrary to expectations, video/speech was no different from speech-only communication. For measures such as pausing, overlapping speech, and interruption management, there were no differences between the video/speech systems and speech-only communication.
When compared with face-to-face communication, both video/speech and speech systems showed reduced ability of listeners to spontaneously take the conversational floor as measured by number of interruptions. Both mediated communication systems also led speakers to use more formal techniques for relinquishing conversational initiative, such as naming a possible next speaker or using tag questions (e.g., “isn’t it?”, “aren’t they?”, “couldn’t you?”—involving an auxiliary verb and question syntax at the end of an utterance). Consistent with this, O’Conaill et al. (1993) and Whittaker and O’Conaill (1993) also found that speakers in naturalistic video-mediated meetings used more formal turn-taking techniques than were observed in face-to-face interaction. They also found that listeners showed reduced ability to spontaneously take the conversational floor. Similar data showing differences in turn taking between face-to-face communication and video/speech are reported by Cohen (1982) and Isaacs and Tang (1994).

More recent work has examined the turn-taking hypothesis in multiperson interactions (Jackson et al., 2000). The prediction was that turn-taking problems would be exacerbated in multiparty video/speech interactions compared with dyadic video/speech interactions. This is because the presence of multiple participants should both lead to increased competition for the conversational floor as well as exacerbating the difficulties in tracking gaze and gestural behaviors of all participants. It was therefore hypothesized that reduced video quality induces more turn-taking problems in multiparty than dyadic groups. However, the study showed the opposite, with reduced video quality having greater effects on dyadic than four-person interactions.

However, subjective data offer more support for the role of visible information in facilitating turn taking. Sellen (1992, 1995) collected subjective data about turn taking gathered from questionnaires addressing subjects’ impressions of the impact of video on conversational processes. Video/speech was perceived to be better than speech in a number of ways. It was thought to: (a) support interruptions, (b) lead to more natural, interactive conversations, (c) increase the ability to listen selectively to particular speakers, (d) allow one to determine whether one is being attended to, and (e) generally keep track of the conversation. People also believed that they were better able to track the attention of others. Similar qualitative data are reported by Isaacs and Tang (1994), who found that video users felt they were better able to manage pauses in video than speech-only communication. Despite this, Tang and Isaacs (1993) found that high-quality video was again not perceived as equivalent to face-to-face interaction: Subjective data showed that video was not seen as being as effective in supporting
interactivity, selective attention, and the ability to take initiative in the conversation.

Availability in Opportunistic Conversations

Hypothesis: Technologies that preclude access to visual information about people’s presence and availability reduce the ability to initiate impromptu workplace conversations given the demonstrated role that visual information about recipient availability plays in opening such conversations. Providing visual availability information should lead to fewer failed attempts to initiate impromptu conversations.

The second cognitive cueing hypothesis concerns the role of visible behaviors in initiating impromptu communications. One can infer the presence of another person (and hence their potential availability for communication) if they are visible. Information about the proximity, current activities, and movements of other people affect various aspects of impromptu conversations (e.g., their initiation and termination) as well as how interruptions are handled (Heath & Luff, 1991; Isaacs, Whittaker, Frohlich, & O’Conaill, 1997; Kendon & Ferber, 1973; Kraut, Fish, Root, & Chalfonte, 1993; Tang & Rua, 1994; Tang et al., 1994; Whittaker, Frohlich, & Daly-Jones, 1994). Workplace interactions are generally unplanned (Kraut et al., 1993; Sproull, 1983; Whittaker et al., 1994), and visible information provides mechanisms for initiating such impromptu conversations. First, sightings of others can lead one to fall into spontaneous conversation (e.g., in public areas such as hallways; Kendon & Ferber, 1973). In addition, seeing a colleague may remind one of an issue that needs to be discussed (Kraut et al., 1993; Whittaker et al., 1994). Visible information is also helpful in determining whether colleagues might be receptive to an unplanned conversation, offering vital clues as to how available or interruptible they are. Clearly all these types of visible availability information are absent from phone-only conversations. One striking statistic is that 60% of business phone calls fail to reach their intended recipient (Rice & Shook, 1990). Calls are either unanswered or transferred to voicemail or a secretary. This failure of the phone has been attributed to the absence of availability information, in contrast to face-to-face settings, which allow participants to time unplanned phone calls for when the recipient is present and receptive to conversation.

Three separate classes of video application have been built to provide availability information to facilitate impromptu interactions between remote coworkers: (a) glance, which enables a user to briefly
look into the office of a coworker to assess their availability; (b) open links, in which persistent video/speech channels are maintained between two separate physical locations simulating a shared physical office, allowing constant availability information; and (c) awareness applications, in which video images of coworkers' offices are periodically sampled so that snapshots of their office can show their recent movements and availability. The difference between awareness and glance is that awareness information is not continuous: It may be a single frame updated periodically.

Fish et al. (1992, 1993) tested the use of different types of glance and their differential success in promoting impromptu interactions. Brief glances at a user-selected recipient were the most frequently chosen type: 81% of user-initiated interactions were of this type, but only 54% of these initiated conversation. This rate is no better than with phone-only interaction. All other types of glances were much less frequent and had even lower likelihoods of resulting in conversations. One type of glance was intended to simulate chance meetings such as bumping into another person in a hallway. This showed high failure rates, with 97% being terminated immediately without conversation. The relationship between glances and impromptu communication was also explored by Tang et al. (1994) for a system operating across multiple sites in a local area. Participants could first look into the office of a remote coworker, with the option of converting this into an extended conversation. Altogether only 25% of glances were converted into conversations. Again this is worse than connection rates using only the phone.

Permanently open video links support the ability to waylay a potential recipient who is out of their office. By monitoring the link, one can see when they return, ensuring that a vital communication takes place (Bly, Harrison, & Irwin, 1993; Fish et al., 1992; Heath & Luff, 1991; Mantei, Baecker, Sellen, Buxton, Milligan, & Wellman, 1991). There are also claims that working with a video link may be less intrusive than sharing a real office, but still offer many of the same benefits in terms of access to other participants (Heath & Luff, 1991). However, evaluation data suggest that open links may be the exception rather than the rule. Fish et al. (1992) reported that only 5% of connections lasted more than 30 minutes, and Tang and Rua (1994) reported only 2% lasted more than 30 minutes. Furthermore, in both cases, these data may overestimate the frequency of open links because they include extended continuous interactions as well as intermittent conversations over open links. Thus, both sets of usage data suggest that brief interactions, rather than open links, are the main uses of the system, again suggesting that availability data may not be enormously valuable.
Open links can also be constructed between public areas (Abel, 1990; Bly et al., 1993; Fish, Kraut, & Chalfonte, 1990; Root, 1988). Cameras were installed in common areas, transmitting images of people who were in equivalent public areas at remote sites. This was intended to promote impromptu conversations of the type that can occur when people meet in public areas of the same site (e.g., coffee room or water cooler conversations). These studies report the frequent use of open links for social greetings between remote sites, with 70% of usage being of this type (Abel, 1990; Bly et al., 1993). However, the use of the open link was mainly limited to these brief social exchanges, and the link was seen by users as being ineffective in supporting work (Fish et al., 1993). Furthermore, sightings over a videolink were again less likely to convert to extended conversation than face-to-face sightings (Fish et al., 1990).

Overall these results offer little evidence for the availability hypothesis. Nevertheless there are methodological problems with the studies reported. These negative results may have occurred because of the weakness of particular implementations. Video quality was often low, and it is well documented that conversations over low-quality video are disrupted compared with face-to-face communication (Cohen, 1982; Daly-Jones et al., 1998; O’Conaill et al., 1993). Therefore, people may be unwilling to initiate conversations when these are held over low-quality video.

Shared Environment and Shared Attention

Hypothesis: Technologies that preclude access to visual information about gaze and gesture in a shared physical environment disrupt reference given the demonstrated role that these behaviors play in supporting reference in face-to-face communication. Reference should be less efficient, taking more turns with speech-only communication and deixis should be much reduced. Communication using a shared environment should also be more efficient and implicit for tasks requiring physical manipulation because the shared environment provides a direct method to jointly perceive changes to the environment, which therefore do not have to be referred to directly.

In face-to-face communication, the fact that participants have access to a shared physical environment means that other types of visible information are available, such as shared information about physical objects and events. Sharing the same physical environment enables people to coordinate conversational content by making inferences about the set of objects and events that others in the same environment are likely to know about and want to talk about (Argyle & Graham, 1977;
Clark & Marshall, 1981; Cooper, 1974; Kahneman, 1973; Whittaker et al., 1991, 1993, 1994). Information about the visible environment often interacts in important ways with verbal and visible behaviors (e.g., when participants gesture at, orient toward, and manipulate aspects of their environment). A crucial aspect of conversational content coordination is the ability to achieve joint reference, and gaze can play a role in reference to external objects or events. People are good at determining where others are looking (Bruce, 1995; Watt, 1995). Gaze serves to coordinate the joint attention of conversational participants. When speakers and listeners achieve joint attention to an object, event, or location, speakers can refer to it by pointing or deixis. Even when they do not resort to pointing gestures, joint attention to an object or event also allows participants greater flexibility in how they verbally refer to it (Clark & Marshall, 1981). Joint attention also contributes to inferences about common ground. If both participants have observed a change to an object or event, they can assume that such changes are part of the conversational common ground. Therefore, they do not have to be mentioned explicitly.

A number of studies have examined the role of a shared visible environment in remote design tasks requiring depiction and spatial reasoning. Bly (1988) observed dialogues in three different technology configurations: face-to-face communication, speech only, and speech/video. In all cases, participants were provided with paper on which they could write or draw; in the speech/video condition, the video camera was directed at the writing/drawing surface so that it could be seen by the remote participant. In the face-to-face and speech/video conditions, participants could therefore observe the other person’s writing, drawing, and gesturing behaviors, whereas in the speech condition, they could not. Bly observed the relative frequencies of gesturing versus writing and drawing in the three conditions. Levels of gesturing were high when there was a shared visible environment: 61% in face to face and 31% in speech/video, but gesturing still occurred 11% of the time in the speech-only condition. Deixis was frequently used to reinforce talk: People would underline or circle objects they were currently discussing even when they did not refer to these with linguistically deictic forms such as this and that.

A study by Whittaker et al. (1993) compared communication effectiveness when people interacted using speech with and without an electronic shared workspace for three tasks: design layout, collaborative text editing, and brainstorming. The shared workspace enabled people to share visual material such as documents or designs as well as to type, draw, write, and annotate these. In the design layout task, participants had to configure a room by arranging pieces of furniture of differ-
ent shapes and sizes given a room plan of specified dimensions. Pieces were given different scores, and participants had to maximize their joint score by inserting high-value pieces. Participants obtained higher scores faster when they had the workspace. Conversational processes also differed between the two conditions. Participants were able to depict and identify visual objects more easily with the workspace because it allowed the straightforward expression of spatial relationships and locations, and this was often achieved by gesture. Efficiency of reference was demonstrated by the finding that participants with the workspace took fewer turns to identify a given piece and fewer turns overall. This was partly because participants were able to refer to pieces deictically. Similar effects were observed in the collaborative editing task. Here the workspace facilitated the straightforward gestural depiction of spatial relationships for text edits that would have been laborious to describe, such as where a sentence should be relocated within a paragraph or where a paragraph should be moved. In contrast, for the simple brainstorming task, there were few benefits for the workspace because of the absence of a spatial component to the task. Olson et al. (1995) reported similar results for a complex visual design task. They compared speech, speech/video, and face-to-face interaction for visual design tasks. Face-to-face groups generated higher quality designs than speech-only groups, but only when they had access to a shared workspace, allowing participants to jointly view, annotate, and gesture at shared designs. There were also differences in conversation processes. Speech-only groups spent more time stating and clarifying issues, again suggesting the utility of the workspace in supporting common ground.

Kraut et al. (1996) also investigated the effects of a shared visual environment provided by video on the performance of the physical task of repairing a bicycle. Novices had to make the repairs to the bicycle while connected to a repair expert by one of three technology combinations. These were (a) speech, (b) speech and high-quality video, and (c) one-way audio and high-quality video. There were no overall differences in quality of solution in the technology conditions, but there were differences in the process of advice giving. In the two video conditions, help giving was more proactive: Experts did not have to be prompted by novices who needed help. In addition, novices were much more implicit in their descriptions of the state of the repair. These observations can be explained in terms of common ground. In the video condition, the experts could perceive when the novice was experiencing setup problems without having to be told this. Also novices were able to assume that experts could see the current state of the system without having to describe it in detail. In both cases, having shared common
ground allows participants to be more implicit about what needs to be said.

These findings are also confirmed by ethnographic research looking at the use of video images to share information about patients’ brains and spines during neurosurgery (Nardi, Schwarz, Kuchinsky, Leichner, Whittaker, & Sclabassi, 1993; Nardi et al., 1996; Whittaker, 1995; Whittaker & O’Conaill, 1997). We compared communication in surgical practice with and without the provision of this visual information. In one setup, the team assisting the surgeon had no visual information about the actions that the surgeon was currently carrying out on the patient. In the other, we provided all team members with broadcast images of the surgeons’ actions on the patient’s brain or spine produced by a magnifying microscope. We observed two different types of communicative use of the video image. First, the dynamic image of the surgeon’s actions on the patient allowed detailed coordination of interleaved physical action between the assisting nurse and operating surgeon. By monitoring the surgeon’s current actions via a shared video image viewed through the microscope, the assisting nurse could anticipate the surgeon’s requirements and provide the correct surgical instrument without it being directly requested. A second communicative function of the video image was that it served to disambiguate other types of surgical data that were supplied to remote consultants, such as neurophysiological monitoring data. The interpretation of these neurophysiological data depends critically on precise information about the physical actions that the surgeon is currently executing, such as the exact placement of a surgical clamp. Without the video image depicting these actions, the remote consultant had to rely on telephone reports from those who were present in the operating theater, and the inadequacy of the descriptions meant that the consultant often had to resort to physically visiting the operating theater to observe the actions directly. In both cases, the video provided situational information that was previously unavailable to certain participants. They no longer had to ask for this information directly, allowing communication to be more implicit and hence more efficient.

Interactivity

Hypothesis: Technologies that preclude access to listener feedback in the form of backchannels, completions, and interruptions can disrupt reference, repairs, and understanding given the demonstrated role that these feedback behaviors play in face-to-face interaction. Technologies that fail to support interactivity show less incremental reference, more redundancy, longer turns, reduced interruptions, and reduced speaker switching.
The importance of interactivity in face-to-face communication is well documented. Listeners offer speakers regular feedback as to whether conversation is on track by giving verbal evidence in the form of backchannels (*mm, uhu*) and visual evidence in the form of head nods and attention (Clark, 1992, 1996; Clark & Wilkes-Gibbs, 1986; Duncan, 1972; Kendon, 1967; Krauss, Bricker, McMahon, & Garlock, 1977; Kraut et al., 1982; Schegloff, 1982; Walker & Whittaker, 1990; Whittaker & Stenton, 1988; Yngve, 1970). Feedback is often given concurrently with the speaker’s utterance to indicate understanding and, in some cases, acceptance of what the speaker has said (Clark & Shaefer, 1989). In addition, the fact that listeners can interrupt to clarify or question the moment that they experience breakdown of understanding has been shown to be an important mechanism for achieving shared understanding (Clark & Wilkes-Gibbs, 1986; Clark & Shaefer, 1989; Oviatt & Cohen, 1991; Schegloff, 1982). A different and stronger sense in which conversation is interactive is that all participants have the chance to incrementally contribute to, negotiate, and modify the content of what is being said. Thus, listeners sometimes interrupt to clarify what a speaker meant, dispute an assertion, ask a question, or complete an utterance for the speaker. Speakers also offer frequent opportunities for listeners to make this type of contribution.

Despite the crucial role of interactivity in face-to-face communication, there has not been a huge amount of research into interactivity in mediated communication. The lack of research is surprising because pervasive technologies such as e-mail and voicemail do not support interactivity, and it would seem to be important to document how this affects communication using those technologies. One of the few attempts to directly test the effect of interactivity is an experiment by Oviatt and Cohen (1991). This directly compared communication processes and outcome using interactive and noninteractive versions of the same linguistic communication mode. Experts provided instructions to novices about how to assemble a toy water pump, either interactively by telephone or noninteractively by recording an audiotape (analogous to leaving a voicemail message). In the audiotape condition, novices listened to the recorded instructions using a tape recorder, which they could stop and start as they wished as they carried out the expert’s instructions. In the interactive condition, instructional exchanges occurred in real time allowing incremental feedback. The results show that interactive communication was more efficient. Interactive pairs were much faster to complete their tasks. Analysis of communication processes also showed that in the interactive condition there was moment-by-moment coordination of communication. Definite reference was often incremental (Clark & Wilkes-Gibbs, 1986),
with the novice confirming at each stage of the reference process that
they were following the description. This incremental feedback al-
lowed misunderstandings to be quickly redressed. In contrast, in the
noninteractive condition, there were large inefficiencies of commu-
ication. Experts tended to generate highly redundant instructions: Ref-
ence descriptions were often elaborated, with repeated descriptions
being applied to the same object. Experts also tended to generate more
structural and summary remarks in the noninteractive condition.
Oviatt and Cohen argued that this overelaboration and structural
marking was a compensatory response to the absence of incremental
feedback. Without feedback, experts in the noninteractive condition
could not establish whether they had been correctly understood. In
consequence, they overelaborated instructions to block potential mis-
derstandings. Similar results of overelaboration are reported by
Krauss and Weinheimer (1966) for a comparison of interactive versus
noninteractive speech.

Another test of interactivity is to investigate the effects of preventing
conversational feedback in face-to-face, video/speech, or speech
communication. Several laboratory studies have shown that the ab-
sence of persistent feedback also leads to overelaborated and highly
redundant messages (Krauss et al., 1977), as well as decreased mu-
tual understanding (Kraut et al., 1982). Another source of evidence
for the importance of interactivity comes from studies of delayed or
disrupted speech communication. Disruptions to interactivity can
occur either because the speech channel is unidirectional (preventing
concurrent feedback, completions, and interruptions) or because of
transmission lags (resulting in feedback or interruptions being de-
layed and arriving at inappropriate times). Interactive phenomena
such as backchannels, repairs, and interruptions characteristically
require precise timing, and minor disruptions to the transmission
characteristics can seriously affect these (Jaffe & Feldstein, 1970;
Norwhine & Murphy, 1938).

Krauss and Bricker (1967) investigated the effects of pure transmis-
sion delays and half-duplex (unidirectional) speech in phone conversa-
tions. In the half-duplex condition, only one person could speak at a
time, and listeners could not interrupt until the previous speaker had
been silent for delays between 250 milliseconds and 1.8 seconds. Both
manipulations were found to have disruptive effects, leading to
speaker overelaboration. These were more serious in the half-duplex
case: Lack of bidirectional access to the channel was disruptive at in-
tervals as short as 250 milliseconds. Similar data showing the effects
of delayed feedback are reported in studies of videoconferencing. Co-
hen (1982) investigated the effects of a 705-millisecond lag on various
conversational measures. (This duration of lag was characteristic of videoconferencing systems in 1982.) The lag led to longer conversational turns and reduced overlapping speech when compared with face-to-face communication. The absence of overlaps was interpreted as compromising key aspects of interactivity that are necessary for concurrent feedback, turn completions, and incremental reference. Whittaker and O'Conaill also investigated these underlying processes in a similar, naturalistic study of videoconferencing, comparing lagged half-duplex videoconferencing with zero-lag, full-duplex videoconferencing and face-to-face communication. Half-duplex speech and lags of between 410 and 720 milliseconds led to reduced backchannelling, fewer interruptions, fewer completions, and fewer overlaps (O'Conaill et al., 1993; Whittaker & O'Conaill, 1993). Listeners found it harder to time incremental feedback or interject in a manner that was not disruptive to the speakers. Therefore, they would wait until the current utterance was complete before making their contribution. Lack of feedback led to longer speaker turns because speakers tended to overelaborate.

Finally, comparisons of spoken and typed interactions also demonstrate the effects of reduced interactivity. Interactive behaviors require precise timing, and the lack of fluency of typed communication should disrupt their execution. Oviatt and Cohen (1991) compared differences between spoken and typed conversation for the toy water pump construction task described earlier involving experts and novices. Typist experts generated instructions that were more verbose and complex. Typist novices produced many fewer backchannels and clarifications. This lack of interactivity led to more elaborations and redundancy in typed than speech instructions, leading pairs of typists to take 3.6 times as long to complete their tasks.

**SUMMARY OF COGNITIVE CUEING HYPOTHESES**

Evidence for the cognitive cueing hypothesis is mixed. Although there is strong evidence for the effects of interactivity and shared environment, the results for turn taking and availability are much less convincing. There are two potential reasons for the negative findings: (a) poor implementations of video systems, and (b) misanalysis of the role of visual behavior in communication. First, the failure to find turn-taking and availability effects may result from flawed implementations because current video implementations may not capture important aspects of face-to-face visual behavior. For example, mutual gaze is impossible, and proxemics and gesture are distorted (Heath & Luff, 1991; O'Conaill et al., 1993; Sellen, 1992, 1995; Whittaker, 1995).
Another more compelling argument for the negative results concerns the role of visual information in communication. Most of the experimental research on cognitive cueing (e.g., the turn taking and availability hypotheses) concerns the role of visual information about other conversational participants. Yet the results suggest that, for cognitive tasks, it is more important to show a shared environment rather than to depict other participants. Why might this be the case? One possibility is that participant information is unimportant for this type of task because participants spend relatively small amounts of time gazing at others in the presence of relevant visible objects. Argyle (1990) reported that it is highly unusual for listeners to spend more than 50% of their time looking at others, with other studies reporting much lower frequencies (Anderson, Bard, Sotillo, Doherty-Sneddon, & Newlands, 1997). Gaze at others falls to 3% to 7% of conversational time when there are interesting objects present (Argyle & Graham, 1977). Mutual gaze (when both participants are simultaneously looking at each other) is even lower, falling to below 5% (Anderson et al., 1997).

Gaver, Sellen, Heath, and Luff (1993) tested what video images people select when they are engaged in a complex physical design task. Users could choose between an image of the other participant and various views of the object under study. Participants chose views of their coparticipant only 11% of the time. Mutual gaze, where both participants were simultaneously viewing each other, occurred only 2% of the time. Instead, people were much more likely to choose an image of the object, spending 49% of their time with the object views. This supports the view that for cognitive tasks, information about the gaze and gestures of others is less important than information about the shared physical context. Given that participants spend so little time looking at others, they have restricted opportunities for eliciting visual information about them. This may explain the lack of evidence for the turn-taking and availability hypotheses. It also suggests that when building video applications for cognitive tasks, it might be more profitable to provide information about shared objects rather than people (Anderson, Smallwood, MacDonald, Mullin, Fleming, & O’Malley, 2000; Nardi et al., 1993, 1996; Whittaker, 1995; Whittaker & O’Conaill, 1997).

SOCIAL CUEING

The previous section focused on cognitive cueing in mediated communication. However, communication is not restricted to the exchange of propositional information. One important aspect of communication concerns the interpersonal and social attitudes of other participants,
including their feelings, emotions, and attitudes. As with conversational intentions, participants generally do not make this information verbally explicit, thus it often has to be inferred. Exchanging this type of information can have significant effects on the content and process of communication. The social cueing hypothesis concerns visible behaviors such as facial expressions and gaze, which convey affective information, and interactive behaviors providing socioemotional feedback. The hypothesis states that technologies that fail to support interpersonal information and feedback alter both the emotional content of communication as well as higher level social processes that mediate negotiation.

There have been two major theories proposed for social cueing. Social presence theory (Short et al., 1976) argues that technologies differ in the extent to which they present a sense of other communication participants’ goals, attitudes, and motives. Because face-to-face communication and video/speech provide rich visual interpersonal information afforded by gaze and facial expressions, they are high on the scale of social presence, with text at the opposite end of this scale. According to social presence theory, using a technology that fails to communicate social presence changes the content and outcome of communication for tasks that require access to interpersonal information. Another focus has been the absence of social cues in mediated communication (Culnan & Markus, 1987; Kiesler et al., 1984; Rutter, 1987). This work argues that the lack of visual interpersonal information in mediated communication, combined with attenuated socioemotional feedback, transforms emotional content and also higher level social processes such as negotiation. Although most of the work on social cueing has investigated remote synchronous communication (such as chat), some of the research described here concerns asynchronous settings. For conceptual clarity, we describe all this work together.

There are three main social cueing hypotheses. These concern (a) communication content, (b) negotiation and deadlock, and (c) social processes such as participation patterns and acceptance. The hypotheses are derivable from Table 7.2. The content hypothesis addresses the role of interpersonal information about participants’ affective and attitudinal state provided in face-to-face communication by visible behaviors such as gaze and facial expression, along with socioemotional feedback provided by backchannels. Technologies that limit access to these behaviors lead to conversations with different emotional content from face-to-face communication. The negotiation and deadlock hypothesis predicts that technologies that reduce information about participants’ attitudes and affect cause participants to overlook others’ perspectives. This changes the outcome of tasks such as negotiation.
that require extensive perspective taking. In extreme cases, the inability to take another’s perspective leads to deadlock. Finally, technologies that limit access to interpersonal information and social feedback are hypothesized to change high-level social processes such as conversational participation patterns and acceptance of others’ conversational contributions.

Content Differences in Mediated Communication

Hypothesis: The content hypothesis concerns the role of interpersonal information provided in face-to-face communication by visible behaviors such as gaze and facial expression, along with interactive feedback. Technologies that limit access to these behaviors lead to communication with different emotional content than face-to-face communication.

Visible behaviors such as gaze, gesture, and facial expressions contribute to the communication of interpersonal information. Gaze is an important indicator of interpersonal attitude or affect. Gaze patterns normally show specific distributions with few gazes lasting more than a second. As a consequence, any deviation from this distribution is associated with an unusual form of interaction. Speakers tend to gaze at a listener’s face more when they are being more persuasive, deceptive, ingratiating, or assertive (Kleinke, 1986) possibly because under these circumstances it is crucial for speakers to monitor the effects of their speech closely. People also tend to look more at conversants whom they like (Exline & Winters, 1965). In addition, people tend to evaluate others by their patterns of gaze: People who look at their interlocutor only a small part of the time are judged as defensive or evasive, whereas those who look a lot of the time are friendly, mature, and sincere (Kleck & Nuessle, 1968).

Facial expressions are also highly informative about the affective state of conversational participants and their current interpersonal attitudes. The face is highly visible, and conversational participants can interpret a great deal from the faces of others. Ekman and colleagues (Ekman, 1982; Ekman & Friesen, 1975) have shown that people across a number of cultures are able to recognize seven distinct facial expressions from posed photographs (happiness, sadness, surprise, anger, disgust, fear, and interest). Affective expressions can be important for both speakers and listeners. They allow listeners to infer speakers’ emotional stance to their utterance, whereas speakers can determine their audience’s reaction to what is being said. The primacy of nonverbal affective information (whether it is conveyed by facial expressions or gaze) is demonstrated by studies showing that when this
visual information is in conflict with verbal information, people tend to trust visual information (Short et al., 1976).

Consistent with the hypothesis, there is evidence that the content of communication is different when visual information is present for negotiation tasks. Stephenson, Ayling, and Rutter (1976) compared two-person debates about union management relationships in speech-only and face-to-face communication settings. They found that speech discussions were more impersonal and task oriented. Depersonalization was manifested by reduced praise for one’s opponent, more blame for the opponent, fewer self-references, and more overall disagreement. Speech discussions also produced more purely informational interactions containing no reference to other conversational participants. Subjective data also support the content hypothesis. Participants believe that technologies providing visual interpersonal information such as video/speech and face to face are better than the telephone for tasks requiring access to interpersonal information, such as getting to know other people or person perception tasks. In addition, groups conversing using video/speech tend to like each other more and regard other participants as more intelligent than those using only speech (Reid, 1977; Short et al., 1976; Williams, 1977). Similarly, Rutter (1987) showed that speech discussions are perceived to be more socially distant than face-to-face discussions.

A number of similar studies have shown that textual communication is more impersonal and task oriented than face-to-face communication (Culnan & Markus, 1987; Hiltz, Johnson, & Turoff, 1986; Hiltz & Turoff, 1978; Rice, 1984). Textual communication obviously does not provide access to visual interpersonal information. These studies compared the content of chat and e-mail with face-to-face communication. Consistent with the hypothesis, they showed that people using textual communication are more concerned with task constraints and matters of fact, rather than the feelings and motives of their conversational partners. They also found lower proportions of socioemotional messages showing agreement. However, these studies argued that the absence of interactive feedback rather than the absence of visual interpersonal information leads mediated communication to show a prevailing task orientation (Hiltz, 1975; Hiltz, Johnson, & Turoff, 1986; Hiltz, Turoff, & Johnson, 1989).

A different and more controversial set of arguments about content in mediated communication concerns flaming: (i.e., inappropriate and strong displays of emotion). Several early studies (Siegel, Dubrovsky, Kiesler, & McGuire, 1986; Kiesler, Zubrow, Moses, & Geller, 1985; Sproull & Kiesler, 1986) argued that such behaviors are more prevalent in mediated communication because of attenuation in the social proc-
ees such as feedback or access to interpersonal information that normally moderate these displays. Inappropriate displays of affect may also be self-reinforcing: An initial display of inappropriate affect may engender equally strong responses, leading to extended interactions with highly emotional exchanges (flame wars). Several early studies reported that such behaviors are more frequent in mediated than face-to-face communication (Kiesler et al., 1985; Sproull & Kiesler, 1986). Sproull and Kiesler (1991) reported 102 flaming remarks in 24 electronic discussions compared with only 12 such remarks for the same number of face-to-face discussions. Chesebro (1985) studied the most recent 10 messages from 14 public computer bulletin boards and found that 32% were interpersonal in nature. Meyers (1985) analyzed two bulletin boards and found that jokes, insults, sexual topics, games, stories, and personal information constituted 29% and 39% of the content of the systems, respectively. However, other more recent studies did not find inflated levels of socioemotional content in e-mail and chat compared with face-to-face interaction (Hiltz, Johnson, & Agle, 1978; Lea & Spears, 1991; Rice & Love, 1987). Furthermore, in some of these studies, there was less flaming in mediated communication than in face-to-face communication. For example, Hiltz and Turoff (1978) found an average of 14% socioemotional content in eight mediated communication groups compared with 33% in three face-to-face communication groups where the groups were trying to reach a joint decision.

Negotiation and Deadlock

*Hypothesis:* The negotiation hypothesis predicts that technologies that reduce access to interpersonal information about participants’ attitudes and affect make perspective taking difficult, and this combines with the reduction in social feedback to make it harder for participants to negotiate effectively and achieve consensus.

There is good evidence for the claim that removing visual information leads to more task-oriented behaviors in mediated communication. Using video/speech as opposed to speech changes the outcome of communication tasks that require access to interpersonal information, such as negotiation, bargaining, and conflict resolution. Short (1971, 1972) studied negotiation behaviors in three technology conditions: face to face, video/speech, and speech only. Participants engaged in simulated role-playing negotiations where they represented either unions or management in an industrial dispute. Each participant was assigned a role and given a detailed description of the dispute. Unknown to the participants, the descriptions were worded so that one side always had objectively the stronger case. These studies demonstrated,
counterintuitively, that there were more settlements in favor of the side with the stronger case in the speech condition. There was also a marked tendency for speech conversations to end in deadlock, whereas face-to-face participants were more likely to compromise. Similar data were reported by Morley and Stephenson (1969, 1970), who compared the outcome of negotiation tasks in face-to-face, video/speech, and speech-only conditions. They found differences between the visual and nonvisual conditions. Speech-only communication demonstrates greater task focus, whereas video/speech and face-to-face participants are more likely to compromise. The results can be interpreted in the following way: The reduction of interpersonal information in the speech condition causes participants to behave more impersonally. They attempt to pursue task goals to win the negotiation—a situation that sometimes produces deadlock. When deadlock does not occur, the side with the factually stronger case tends to succeed. In contrast, in face-to-face communication, the greater availability of interpersonal information leads participants to become less task focused. Here the merits of the case are less prominent, and compromise is more likely to occur (Reid, 1977; Short et al., 1976; Williams, 1977).

Other research investigates the processes for arriving at consensus in mediated communication. Decision making in face-to-face settings generally shows a gradual convergence, with successive speakers being more likely to agree than disagree with the proposals of prior conversants. In mediated communication, in contrast, the absence of visible interpersonal information and social feedback make rapid convergence harder to achieve. Hiltz et al. (1986) showed that for judgment tasks, mediated communication groups using textual conferencing were less likely than face-to-face communication groups to reach consensus. They also had more difficulty in deciding among alternatives. Other related research shows that coalitions are easier to form when visual information is available whether this is video/speech or face-to-face communication (Reid, 1977). This is again consistent with the view that visual information provides access to the feelings and motives of others. However, another result from the Hiltz et al. (1986) study suggests that these findings may be task specific. Hiltz et al. found that consensus could be achieved in mediated communication when intellectional tasks were used.

**Participation and Acceptance**

*Hypothesis: Technologies that limit access to interpersonal information and social feedback compromise social processes. This leads to equalized conversational participation patterns and reduced acceptance of others’ conversational contributions.*
Multiperson face-to-face interactions have strong participant inequality. Sproull and Kiesler (1991) showed that in small-group, face-to-face discussions, half the participants contribute only 10% to 20% of talk. Furthermore, the amount that a person talks is strongly correlated with his or her perceived status within the group. For example, in face-to-face groups, managers talk more than subordinates and men more than women. This holds true even when the higher status (and generally more vocal) members are not more expert on the topic under consideration. In contrast, mediated communication groups communicating real time using text show much greater equality of participation (Dubrovsky, Kiesler, & Sethna, 1991; Hiltz et al., 1986, 1989; Kiesler et al., 1985; McGuire et al., 1987; Siegel et al., 1986; Weisband, 1992). Similarly, subjective data from interview studies of e-mail show equalization of contributions. E-mail is reported to be more democratic because minorities and lower status individuals feel they are more able to contribute (Sproull & Kiesler, 1991). E-mail also shows a greater diversity of perspectives (Sproull & Kiesler, 1991). These studies argued that mediated communication attenuates visible social status cues because personal identity is often restricted to an e-mail address or login name. Greater participant equality partly results from this reduction in visible status cues.

Other research has shown differences in patterns of acceptance of others’ contributions in mediated communication. Several studies show that consensus is harder to achieve in mediated communication (Hiltz et al., 1978, 1986, 1989; Olaniran, 1994; Rice, 1984; Siegel et al., 1986; Straus & McGrath, 1994). The imposition of status is argued to be one mechanism for ruling out competing alternatives when there is dissent. Two studies support this view. Dubrovsky et al. (1991) documented how status effects are reduced in mediated compared with face-to-face communication. In mediated communication, high-status members are less likely to be the first participant to propose a decision; if made, their proposals are not as likely to be accepted as in face-to-face communication. Weisband (1992) reported similar conclusions in a study of consensus formation comparing chat and face-to-face conversations.

**SUMMARY OF SOCIAL CUEING**

The results for the specific social cueing hypothesis are again mixed. Although there is good support for the effects of mediation on negotiation, participation, and acceptance, support for the content hypothesis is less clear. Although mediated communication often has different content from face-to-face communication, we need to refine this hy-
hypothesis to avoid apparent contradictions. In some cases, we are led to expect mediated communication to have increased emotional content and interpersonal expression (e.g., flaming), but in other cases it is predicted to show reduced interpersonal focus, leading to task-centric behaviors. How can we explain these apparent contradictions? One possibility is that there are different types of emotional effects, and that these arise from different underlying mechanisms. One final observation about this work is that the data seem to be more consistent with a task-specific account, rather than the view that social processes are generally different in mediated communication. There are few reports of changed emotional content or changes in negotiation, participation, and acceptance when the tasks are intellectual in nature (Hiltz et al., 1986, 1989; Reid, 1977; Short et al., 1976). This is consistent with the research on cognitive cueing reviewed earlier.

Recent work on social throughput has suggested an alternative mechanism to explain different content effects in mediated communication. It is argued that these arise from a failure of social throughput because typed communication is used. The argument is that content effects arise from the lack of fluency and slow speed of typing, rather than the absence of visual interpersonal information. Daly (1993) and Walther (1992, 1994, 2002) claimed that in textual communication, the effort and slow pace of typing cause participants to be more terse, leading them to be perceived as less polite. This, in turn, engenders flaming or dissent. Support for this view comes from studies that give participants who are communicating by typing unrestricted time to complete tasks. Unrestricted time should allow participants to overcome lack of fluency of typing allowing focus on emotional expression. As we see later, several studies show that, given more time, mediated communication participants adjust their typing to be less terse and more socioemotional.

Social throughput also has implications for other social cueing hypotheses. According to some theories (Rutter, 1987), changes to negotiation processes and participation levels follow directly from the impersonal content of mediated communication. Yet if changes to emotional content are the result of textual media failing to provide sufficient social throughput, it follows that negotiation, participation, and acceptance behaviors should also change when participants are given more time to complete tasks. This turns out to be the case. Studies show that, given unrestricted time, outcome differences between chat and face-to-face communication are much reduced and participant inequality reasserts itself in mediated communication (Hiltz et al., 1989; McLeod, 1992). The effects of social throughput are also found in naturalistic studies: Both Reid et al. (1996) and Steinfield (1986) showed
that socioemotional remarks decrease in mediated communication when groups are under deadline pressure. Again the slow rate of typing leads participants under time pressure to abandon socioemotional discourse to focus on the task at hand. Together these data argue that for applications such as chat or e-mail, rather than following directly from the absence of interpersonal visual information, social cueing effects may result from problems of social throughput (Walther, 2002). Nevertheless, social throughput may not be able to explain some early work (Morley & Stephenson, 1969; Reid, 1977; Stephenson et al., 1976) comparing speech and face-to-face communication. It is harder to explain how these differences could be explained by social throughput because both media conditions used speech, which is highly fluent. These early studies nevertheless showed differences in content and task outcome between face-to-face and speech-only communication. This suggests there may be two separate types of social cueing effects arising from lack of social throughput on the one hand and lack of visual interpersonal information on the other. We need more research to address this.

Taken together, the social cueing results indicate that we need better defined models of how social cueing occurs and experiments that partial out the effects of different mechanisms before we can make systematic statements about how these processes operate. More specific accounts and data might lead us to be able to tease apart some of the differences between social throughput and visual information effects.

**CONCLUSIONS**

We now present some general conclusions and suggest future theoretical and empirical work that needs to be done on mediated communication. One of our initial orienting questions concerned whether face-to-face and mediated communication are different. Our data show that this depends on the task. For cognitive tasks using interactive technologies, there are generally few differences between face-to-face communication and mediated communication (Chapanis et al., 1972, 1977). However, mediation seems to have larger effects for tasks requiring access to interpersonal information (Kiesler et al., 1984; Morley & Stephenson, 1969; Short et al., 1976) or those requiring complex joint physical manipulation (Bly, 1988; Kraut et al., 1996; Whittaker et al., 1993).

This result is important for theories of mediated communication. How can we explain these effects? The finding that mediated and face-to-face communication are equivalent under certain conditions is the-
oretically important because it shows that mediation per se does not change communication. Rather, the results indicate that we need to explain these observations in terms of the underlying affordances of various technologies and how these affordances affect a technology's ability to support various communication behaviors. The data indicate that these behaviors in turn have effects on core communication processes. For example, it is clear that interactive technologies allow incremental feedback that facilitates shared understanding and establishing common ground, as well as supporting incremental reference. Technologies that do not support interactivity (e.g., e-mail, voicemail) show disrupted processes (e.g., reference) and outcome (lack of shared understanding). Similarly, technologies that transmit visual information about facial expressions and gaze may be critical for tasks in which affect and attitude have a central role. Technologies that do not support visual interpersonal information (e.g., speech, e-mail) lead to interactions that are more impersonal and likely to end in deadlock. And technologies that provide a shared visual environment are important when tasks require complex reference to and joint manipulations of physical objects.

These results also offer practical insights about why certain intuitively plausible novel technologies have largely failed. People have strong intuitions about the value of visual information about other people's gaze, gesture, and facial expressions in communication (Isaacs & Tang, 1994; O'Conaill et al., 1993; Sellen, 1995). Nevertheless, counter to these intuitions, the cognitive cueing research indicates that such visual information is not important for either turn taking or availability. Yet many applications of video presuppose this type of nonverbal information (Whittaker, 1995). This may explain the failure of technologies that attempt to provide visual images of other conversational participants (e.g., videoconferencing, videophone) when these technologies are used for cognitive tasks in work settings (Fish et al., 1992, 1993). Our research also suggests that video might be more profitably used to communicate information about shared objects in a shared environment rather than talking heads-type applications (Anderson et al., 2000; Nardi et al., 1996; Whittaker, 1995; Whittaker & O'Conaill, 1997). The bandwidth hypothesis research also suggests reasons for the success of the phone. Results show the primacy of speech for a number of different tasks, along with demonstrating that combining other technologies with speech seldom improves performance, making it hard to improve on the phone (Chapanis et al., 1972, 1977).

We can also make other practical recommendations about the use of various mediated communication technologies. Our results show that the effect of using mediated communication depends on the task. For
social tasks, there are clearly differences between mediated and face-to-face interaction, but for many cognitive tasks (especially those that do not require access to a shared physical environment), outcomes may not be different. Therefore, technology users may decide to avoid (or at least be aware of the consequences) of using mediated communication if their current activity has a strong social component. However, there are two problems in applying these types of guidelines directly. The first is the absence of a systematic task taxonomy (Straus & McGrath, 1994), enabling potential technology users to determine whether their current communication task is cognitive or social prior to making their decision about whether to use technology for that task. A related point is that research on naturally occurring interactions shows that they are often heterogeneous: It is unusual for a single real-world communicative task to be predominantly social or cognitive, with most interactions containing elements of both (Short et al., 1976).

Mediated communication research also makes a theoretical contribution to theories of face-to-face communication. In one way, we can see technologies and their affordances as natural experiments, enabling us to dissect the contributions of different communication behaviors on communication processes, outcome, and content. As noted, with the exception of work on nonverbal communication (Argyle, 1990; Beattie, 1978, 1981; Duncan, 1972; Kendon, 1967), most theories of face-to-face interaction have tended to not isolate the effects of different behaviors on communication process and content. In contrast, the experiments reported here are extremely important in helping clarify the role of visible behaviors (facial expressions, gaze, gesture). The data reviewed here show that visible behaviors such as gaze and gesture are important for communicating interpersonal information. However, contrary to people's intuitions and the claims of various studies (Argyle, 1990; Duncan, 1972; Kendon, 1967), they do not seem to be necessary for availability or turn taking. The experiments reviewed here also document the role of interactivity as being a vital communication process in mediating understanding. The studies of shared workspaces also provide evidence for the notion of grounding: showing the effects of a shared visual environment on reference and explicitness of communication.

However, one weakness of many experiments on mediated communication has been the focus on technology instead of affordances and communication behaviors. Most experiments have treated technologies as independent variables. Although this provides important practical information for users of that technology, it does not enable us to theoretically disentangle the contributions of the technology's affordances on underlying behaviors. For example, many early experi-
ments compared e-mail with face-to-face communication, but when
differences were discovered, it was impossible to determine whether
these resulted from the lack of visual information in e-mail or the fact
that it did not support interactivity. Future research needs to focus on
underlying affordances and communication behaviors, rather than ex-
isting technologies per se.

What makes the picture more complex is that the causal relation-
ships between affordances, communication behaviors, and core com-
municative phenomena are highly intricate. A single affordance can
support multiple communication behaviors, and these in turn can
have effects on different core communicative phenomena. Thus, the vi-
sual mode supports multiple behaviors (gaze, head nods, gesture,
shared visual environment) relating to different aspects of communi-
cation (turn taking, reference, attention, availability, interpersonal in-
formation). Similarly, one behavior can affect multiple communica-

tion phenomena: Backchannels support attention, reference, turn
taking, and agreement. Furthermore, communication behaviors are
often redundant with respect to the core communication processes
they affect. Thus, gaze, gesture, head nods, and verbal backchannels
all contribute to turn taking. One consequence of all this is that it weak-
ens predictions that can be made to test the effects of technology
affordances on communication processes. For example, removing vi-
sual information may have less serious effects on turn taking than
might be expected because other communication behaviors that do
not rely on the visual mode—namely, verbal backchannels, tag ques-
tions, or vocative addressing—can substitute for the absence of vi-
sual information (O’Conaill et al., 1993; Whittaker & O’Conaill,
1997). This suggests that future research needs to be much clearer
when testing hypotheses about the relevant affordances, underlying
behaviors, and anticipated effects on communication. To determine
precise effects, it may be necessary to configure technologies to iso-
late different types of behaviors (e.g., to experiment with technologies
that support the transmission of facial expressions, but not gaze), to
determine the separate contributions of these behaviors.

Another theoretical problem is the lack of a clear underlying model
of the relationship among core communicative phenomena such as
processes, content, shared understanding, and task outcome. Much
of the cognitive cueing work presupposes a model in which tech-
nology affordances influence underlying communication behaviors,
which then affect higher level processes such as understanding and
conversational content (Clark, 1992, 1996; Clark & Brennan, 1991;
McCarthy et al., 1991, 1993; O’Conaill et al., 1993; Oviatt & Cohen,
1991; Whittaker et al., 1991, 1993). In contrast, the social cueing
work of Rutter (1987) and Short et al. (1976) proposed models in which technologies directly influence the content of communication, which then affects communication processes and understanding in turn. According to these social cueing theories, technologies lacking the visual mode are more impersonal and formal because of the absence of interpersonal content. This absence of affective information in turn affects higher level processes such as negotiation. Although Rutter (1987) presents regression analyses supporting his particular causal model, our models generally need to be clearer about the relationship among technologies, affordances, processes, content, and understanding.

There are a number of other limitations to our theoretical account. The first concerns the role of tasks. We have repeatedly observed that the effects of various technologies and their underlying affordances on communication depend on the task. The major differences occur between social and cognitive tasks. As noted earlier, however, we need to move beyond this simple distinction and provide a richer task taxonomy. Although there have been repeated attempts to do this, in general they have not proved to be successful (Kraut et al., 1992; Straus & McGrath, 1994). Media richness theory (Daft & Lengel, 1984), for example, proposes that tasks can be categorized in terms of their equivocality (i.e., the existence of multiple or conflicting interpretations). However, the dimension of equivocality does not map directly onto existing taxonomies of group communications (McGrath, 1984; Straus & McGrath, 1994), nor does it correspond well with observations of naturally occurring workplace interactions (Short et al., 1976). Although other task taxonomies have been proposed (McGrath, 1984; Morrison & Vogel, 1991), these can be subjected to both theoretical and empirical criticism (Short et al., 1976; Straus & McGrath, 1994). Further work on defining a task taxonomy is much needed.

Another fundamental limitation of our theories is the impoverished notion of affordances. An enriched taxonomy of affordances may address some of the problems observed earlier concerning the complex relations among affordances, behaviors, and communication phenomena. It should mean that we are better able to map from technology affordances to predictions about communication. The current review has focused on two types of affordance: mode and interactivity. Although the simple distinction between mode and interactivity affordances explains much of the current data, the notion of affordances needs to be refined in a number of ways. First, the notion of visual information is too broad. We need to distinguish among aspects of visual information relating to affective information (e.g., facial expressions), turn taking (e.g., gaze), and availability (e.g., physical presence). Similarly, it is
apparent that rather than being a straightforward dichotomy, interactivity is a matter of degree. Numerous studies have shown that speech is more interactive than typing, although both are real-time technologies. Speech allows for precisely timed backchannels, interruptions, and completions and these processes that are harder to achieve in typing (Oviatt & Cohen, 1991; Whittaker et al., 1991). Another underresearched affordance concerns the extent to which technologies support conversational context. One obvious property of many technologies (e.g., voicemail, e-mail, chat, instant messaging) is that they are permanent. Permanence means that prior conversational context is viewable and does not have to be remembered (McCarthy et al., 1991, 1993; Whittaker & Sidner, 1996; Whittaker et al., 1991; Whittaker, Swanson, Kucan, & Sidner, 1997). Having context directly available has been shown to change conversational processes and outcome, indicating the importance of this affordance (McCarthy et al., 1993; Whittaker et al., 1991, 1993). Another final affordance is expressivity. Comparisons of noninteractive uses of voice and text, for annotation (Kraut et al., 1992), and asynchronous communication (Whittaker, Hirschberg, & Nakatani, 1998; Whittaker, Davis, Hirschberg, & Muller, 2000), show that voice is used for different tasks than text. Voice is also rated by users as better for describing global and affective information than text. Together these data suggest that we should enrich our set of affordances to include richer definitions of visual information, interactivity, context, and expressivity.

We also need to move away from purely comparative theories. Most mediated communication research explicitly uses models derived from face-to-face communication. Yet there are strong reasons to think that asynchronous communication raises a large number of problems not commonly encountered in face-to-face settings. Research into e-mail and long-term collaborations has documented the fact that asynchronous communications enable participants to engage in multiple simultaneous intermittent interactions (Kraut et al., 1993; Whittaker et al., 1994, 2000; Whittaker & Sidner, 1996; Whittaker, Jones, & Terveen, 2002). Such interactions can take place sporadically over time intervals lasting from minutes to years. This gives rise to a whole new set of theoretical and empirical questions about how participants (a) monitor multiple concurrent conversations, (b) retain the context of those conversations over long time intervals, and (c) reintiate an inactive conversation to regain the attention of their conversational partners. Long-term intermittent conversations also raise important theoretical questions about common ground and shared understanding. Most face-to-face communication theories (e.g., Clark, 1992, 1996; Clark & Shaefer, 1989; Krauss et al., 1977) argue that shared under-
standing is achieved through interactive processes by which feedback can be given and misunderstandings quickly redressed. But in long-term intermittent interaction, such interactivity is absent, so how can discrepancies in understanding be detected and addressed and what stops misunderstandings from proliferating? These are important questions because they challenge fundamental assumptions about communication.

Another set of fundamental issues concerns the application of face-to-face theories to mass communication. Again most mediated communication models are implicitly derived from a restricted set of circumstances: dyadic face-to-face communication. However, in many new forms of asynchronous communication involving technologies such as Usenet and Lotus Notes, interactions can take place between hundreds or even thousands of participants. Recent research has shown that the principles guiding such mass interactions may be fundamentally different from face-to-face interaction, with longer, more involved discussions taking place between participants who have less rather than more common ground (Whittaker, 1996; Whittaker et al., 1998). This research argues that notions of weak ties rather than mutual understanding may better explain these data.

Together these theoretical and empirical gaps call for more sophisticated theorizing about technology affordances and tasks. We also need more systematic research into mediated communication that does not presuppose face-to-face models. One strategy to achieve this might be to explore situations in which mediated communication is better than face-to-face communication (Hollan & Stornetta, 1992). For example, recent research has shown that in certain settings instant messaging is preferred to face-to-face communication because it is less interruptive (Nardi et al., 2000). In conclusion, to move the field of mediated communication forward, we need progress in a number of areas. We need to refine our notions of task and the relations between tasks, produce more careful analyses of important underlying explanatory affordances, and define new models that better characterize asynchronous communication and mass communication. Unless we do this, we may never have theories that truly address mediated communication, rather than trying to account for mediated communication results in terms of face-to-face communication.

REFERENCES


7. COMPUTER-MEDIATED COMMUNICATION


7. COMPUTER-MEDIATED COMMUNICATION 283


7. COMPUTER-MEDIATED COMMUNICATION 285


The phrase *narrative discourse* juxtaposes two of the most varied and ill-defined terms in contemporary research on language. Definitions of both terms vary depending on research paradigm, disciplinary perspective, methodological tradition applied, purpose of the research, and the corpus of data available. Nonetheless, the study of narrative discourse has found a prominent place in the fields of history, literary studies, psychology, cognitive science, sociolinguistics, sociolinguistic ethnography, ethnomethodology, systemic linguistics, sociology, anthropology, folklore, education, medicine, and other professional domains (Lieblich et al., 1998).

One way to view the variation of definitions is that it reflects a disorganized field (or set of fields) with limited communication and convergence. However, I take the view in this chapter that the varied definitions of narrative discourse are instead an indication of vibrant theoretical and methodological debates and tensions in the contemporary study of language, part of the broader, historical debate over ideologies of language (Blommaert, 1999; Currie, 1998; Schieffelin et al., 1998). Further, I find a distinction that Hasan (1999) made between endotropic and exotropic theories helpful in framing the varied theoretical discussions of narrative discourse.

Endotropic theories are centered onto their own object of study, isolating it from all else. The phenomena they attempt to describe are viewed
as if they were self-generating, self-fertilizing, self-renewing; they are, thus, autogamous with respect to their central problematic. . . . By contrast, an exotropic theory is not confined within the bounds of its object of study. Rather, it is cosmorganic, typically embedding its central problematic in a context where the processes of its maintenance and change originate in its interaction with other universes of experience. From this point of view, the object of study in an exotropic theory is a component of what Lemke (1984, 1993) calls a “dynamic open system,” changing and being changed by its reciprocal engagement with the other components of the larger system. (p. 13)

I locate discussions of narrative discourse at the intersection of a broad range of exotropic theories concerned with narrative on the one hand and discourse on the other. Although such a location may not provide the tidiness that some may prefer, it may better reflect the state of the field and has the advantage of emphasizing the generation of ideas and insights.

Having located discussions of narrative discourse at the intersection of a broad range of exotropic theories, the purpose of this chapter is to highlight a series of problematics that may be useful heuristics for the analysis of narrative discourse. I use the term problematics to refer to tensions inherent to the phenomenon (in this chapter, the terms problematic and tension are used interchangeably). A problematic differs from a problem. A problem is approached as if it can be solved, whereas a problematic is a condition of the phenomenon and is approached to gain insight into the phenomenon. To paraphrase Robinson (1987), narrative discourse is a complex human activity caught up in all of the complexities, struggles, ambiguities, incompleteness, and problematics of human activity.

Although the discussion in this chapter includes a broad range of perspectives, it is not intended to be a comprehensive review. Instead, it focuses on a set of tensions and issues that provide useful entry points for exploring narrative discourse. I begin the chapter by discussing the boundaries of a narrative. I do so in lieu of providing a definition of narrative or narrative discourse. By setting the boundaries of a narrative, I am addressing the need to define the scope of narrative discourse. Then informed by van Dijk’s (1997) discussion of the study of discourse, this chapter is organized around two problematics central to the contemporary study of narrative discourse: (a) narrative discourse as text versus narrative discourse as event and practice, and (b) narrative discourse as structure versus substance. These two tensions overlap each other. That is, even if the view of narrative discourse as text is held constant, there are still a series of problematics that are important to examine between narrative as structure and narrative as
substance. It is similarly so if holding the view of narrative as event and practice constant. Thus, although I have organized the chapter around the two major problematics discussed previously, within the discussion of each, I discuss the complexities and tensions that arise when the two sets of tensions are juxtaposed.

THE BOUNDARIES OF A NARRATIVE

This chapter takes the controversial position that narrative is a fundamental language process underlying all genres, including those not usually considered narrative (McCabe’s [1991] concern notwithstanding that “narrative is a word in danger of being overused”; p. xv). Such a position eschews the binary distinction between narrative and non-narrative, which has been used by scholars to distinguish between narrative on the one hand and, on the other hand, exposition, scientific reports, different ways of thinking and knowing, and ways of representing the world.

At one level, a distinction between narrative and non-narrative can be useful in validating a variety of systems of thought and a variety of ways to represent meaning. At another level, however, any text can be viewed as a way to represent actions in the world, and the production and performance of a text can be viewed as an attempt to act on the world. Thus, rather than ask “What is a narrative?”, one asks, “What is being called (and not called) a narrative, where, when, how, by whom, and for what purposes?”

A distinction between narrative and non-narrative texts can be problematic even when considering scientific texts. For example, Singer and Gagnon (1999) found that “readers of expository [and scientific] text engage in inference processes reminiscent of those that accompany narrative comprehension” (p. 192). The findings suggest that distinctions between narrative and non-narrative texts on the basis of assumptions about differences in the cognitive processes readers might employ is problematic (see also Fletcher et al., 1999, with regard to mathematical proofs). All texts, scientific and mathematical, are the resultant product of human action in time, reflecting a past and a future, and constitute configurations and reconfigurations (Ricoeur, 1984) of the world and what happens in it. Consider a taxonomy of proteins—a text that would not typically be considered a narrative. Such a text reflects an evolution from earlier taxonomies based on a history of human action, an understanding of how proteins are and how they got that way, while anticipating human action to take place based on it (not just in the refinement of the taxonomy, but in what people might do to the proteins—e.g., splicing). A novice might read the
taxonomy without knowing the story it tells—not having the background to derive the story from the text. An expert, however, would be able to read the story (or stories) in the text of the taxonomy (cf. Geisler, 1994). That microbiologists write in a way that might hide the underlying story (from nonmicrobiologists) and that it is not traditionally viewed as a narrative is not grounds for dismissing the text as a narrative, but rather for investigating the relationship among the surface level of a text, the narratives it presents, the underlying story, and its use in a storytelling event.

In summary, the boundaries of a narrative and narrative discourse do not lie in the phenomena; they are as boundless as language itself. Rather, the boundaries lie in the perspectives brought to bear on human activity.

NARRATIVE AS TEXT VERSUS NARRATIVE AS EVENT AND PRACTICE

Narrative as Text

Labov (1972) defined a narrative as: “. . . a sequence of two clauses which are temporally ordered . . . a minimal narrative is defined as one containing a single temporal juncture” (p. 361). For Labov, a narrative is a textual recapitulation (temporally structured) of experience. As text, one can ask questions about the nature of the narrative text, including its structure and substance, how the text mediates social identities and relationships, what cognitive and linguistic processes an individual might employ to derive particular meanings from the text, how the text is related to other texts, and the fidelity and vraisemblance of representation, among other questions. I begin by examining questions of fidelity and vraisemblance (a sense by the reader/listener that the narrative text describes what really happened and what the place is like). Then I build on that discussion in consider other questions and problematics in considering narrative as text.

Fidelity and Vraisemblance. Questions about the fidelity of representation require more than simplistic assessments between the narrative text and what it is signifying. The fidelity question (How accurately does the text represent the phenomena it describes?) focuses attention on the information presented (including chronologies, relationships among events, explicitly stated and inferred) as well as the information omitted. Fidelity can also refer to the relationship between a narrative text and underlying themes, ideologies, or human qualities (“How ac-
curately do the underlying themes of a narrative represent the human condition?”). Although in some cases it may even be useful for heuristic purposes to ask such questions, if what one means by fidelity is the transparent encoding or the imitation of reality, then questions about fidelity are nonsequitors. As Ricoeur (1984) pointed out, a narrative is always a configuration or reconfiguration of human action—an imposition on human action. As such, a narrative is at best a presentation or re-presentation rather than an imitation or transparent encoding. Questions about fidelity might better be considered questions about *vraisemblance*—the sense a reader or listener has that the description and narrative presented is accurate.

The generation of *vraisemblance* is based, in part, on the detail presented, the credentials of the author and narrator, the adherence of the narrative to the cultural rules for the type of narrative presented, among other factors (see Atkinson, 1991). In addition, package presentation is important to *vraisemblance*. Package presentation refers to the broad range of factors that constitute the presentation of the text, including the materials of the text (e.g., cover, paper, font, hardbound vs. paperback, leather bound), its categorization (e.g., fiction, romance, nonfiction), and setting of its presentation (e.g., university course, supermarket). Building on Todorov (1968), Atkinson (1991) wrote about the layers of *vraisemblance*:

First, there is the relation between the given text and “public opinion”; second, there is the text’s degree of correspondence to the expectations or conventions of a given genre; third is the extent to which the text *masks its own textual conventions, appearing to conform to a “reality.”* (p. 39; italics added)

The masking of rhetorical tropes and textual conventions allows people to equate the explanation, causality, and morality promulgated in a narrative (which is part of what a narrative does; cf. Ochs, 1997) with explanation, causality, and morality in the narrated events. That is, explanation, causality, and morality are in a sense given to events by how the events are made into a narrative, and this giving may go undetected depending on the rhetorical techniques used, participants in which the narrative is used, and situation of its use.

Judgments of *vraisemblance* are influenced by the politics of identity. The author of a narrative brings to the production of the text her or his own frame of reference and point of view, which are influenced by her or his cultural, class, and gendered experiences (among other experiences). Thus, questions about the *vraisemblance* of a narrative text include how a narrative text reflects the producer’s framing of the
world, and this is so whether explicitly asked or not. To not overtly ex-
amine the producer’s relationship to the narrative text and its uses is
to accept an unmarked (default) judgment that the relationship is one
of authoritative transparency. It is similarly so with readers and their
interpretations of narrative texts. That is, to not overtly examine a
reader’s relationship to the narrative text and to its uses is to accept an
unmarked (default) judgment that the interpretation is one of authori-
tative transparency.

The prior discussion of fidelity and _vraisemblance_ makes clear that
even when a narrative is defined as a text, social relationships and so-
cial identities are at stake, especially those of the people being repre-
sented, the author/composer, and the reader/listener. Yet it is not just
the social relationships and social identities of people that are at stake;
it is also those of social institutions. That is, judgments of _vraisem-
blance_ are made not just by individuals, but also by social institutions.
Social institutions establish mechanisms to make and justify judg-
ments of _vraisemblance_. The review procedures of research journals
are one example; the procedures of religious synods, panels, courts,
and other bodies are also examples. I take up the issue of social rela-
tionships and social identities in more detail later.

**Social Relationships, Social Identities, and Narrative Text.** In the analysis of
social relationships and social identities in narrative text, it is useful to
make a distinction between author and author-in-the-text, reader and
reader-in-the-text (see Goffman, 1981; Robinson, 1987), and, more
generally, people and people-in-the-text. For example, consider the fol-
lowing passage from McGuffey’s First Reader (1879):

> We have come to the last lesson in this book.
> We have finished the First Reader.
> You can now read all the lessons in it, and can write them on your slates.
> Have you taken good care of your book?
> Children should always keep their books neat and clean.
> Are you not glad to be ready for a new book?
> Your parents are very kind to send you to school. If you are good, and if
> you try to learn, your teacher will love you, and you will please your par-
> ents.
> Be kind to all, and do not waste your time in school.
> When you go home, you may ask your parents to get you a Second
> Reader.
> (from Lass & Tasman, 1981, p. 280)

The reader-in-the-text is the student who presumably finished the
book—the _you_ in the passage and the _child_ in line 5. The reader-in-the-
text is in school, can read all the lessons and write them on the slate, is glad to have finished the book, has a family, and so on. Of course, the person who actually read the text printed here in this chapter is more likely to be an adult and more specifically a scholar. Even if we considered the reading of the prior text in an elementary classroom in 1879, it is unlikely that everyone reading that text could read all the lessons, write them on a slate, was glad to finish the book and get a new one, nor did everyone have a family, and so on. The narrative text creates the implied reader—what can be called the reader-in-the-text.

Similarly, the prior text creates the author or, more accurately, the author-in-the-text. The author-in-the-text seems to be the anthropomorphized book—a voice of authority and morality that looks down on student and teacher. Yet the author(s) is (are) a writer(s) working for the publisher. What the narrative text does is create the author-in-the-text. Interestingly, Goodman et al. (1988) pointed out that in modern basal readers and school textbooks, the authors listed in the book may not have actually done the writing—yet another twist on the creation of the author in the text.

However, it is not just authors and readers who are being created. A world is being created with its own cultural norms. In the passage from McGuffey’s Reader, norms are implied for feelings, moral action, and relationships among teachers, students, and parents. At issue is the assumed relationship between people and the world in the text and a represented group of people and their world. For example, consider the representation of teachers in the McGuffey Reader text and in other school texts. To what degree do those representations actually reflect the lives of teachers? Weber and Mitchell (1995; Mitchell & Weber, 1999), among others, have expressed concern for how teachers are represented in narrative texts such as the one earlier and in popular movies. They argued that such representations have an influence on what actually occurs in education. They further argued that the pastoral, placid, desexed nostalgic images of teachers may supplant people’s own experiences. Referring to the Dick and Jane reading series, they wrote:

So potent are some of these images [from the Dick and Jane reading series] in evoking memory, that even people who were not taught with this particular series may be convinced that they were, unwittingly allowing a dominant image to “overwrite” their memories of their own experiences! (Mitchell & Weber, 1999, p. 3)

The world-in-the-text and the people-in-the-text (what cognitive scientists often call the situation model) may have a powerful influence on how people define and understand their world.
Through a variety of linguistic means, including the use of pronouns and nouns, genre, verb tense, deixis, ellipsis, and omission (e.g., Hanks, 2000; Ivanic, 1994; Kamler, 1994; Keogh, 1997; Phelan, 1994; Pratt, 1994), texts either foreground the distinction between the people-in-the-text and a specific group of people, the author, and the reader or texts pretend such distinctions and social relationships do not exist. An example of a narrative text that foregrounds the distinctions is *Huckleberry Finn* (Twain, 1985/1885). Although *Huckleberry Finn* is written in the first person, the text makes clear in the first few sentences that the author in the text is fictional and separate from the author. By contrast, in *Cláper* (Freilich, 1998/1987), the distinction between the author-in-the-text and the author is blurred. The narrator (author-in-the-text) is a woman by the name of Alicia Freilich, and the author of the book is Alicia Freilich.

Although it may seem obvious to consider the author-in-the-text to be a fiction, it may be less obvious that the author is also a fiction (cf. Foucault, 1984). Mark Twain is a fiction of Samuel Clemens, but what makes Mark Twain a fiction is not the use of a pseudonym, but the fashioning of an author. The pseudonym is merely an overt sign of the fashioning. Any author is similarly a fashioning, including Alicia Freilich or myself. The problematic nature of the author is highlighted in books like Ana María Shua’s (1998) *The Book of Memories*, in which the authorship of the story (the author in the text) is a book of memories (similar to a family photograph album) whose meaning and veracity is debated and argued by two narrators, and although not explicitly the history of the Shua family is nonetheless sufficiently similar to have a sense of life history. Deliberately so, Shua appears to be problematizing authorship (as well as memory, history, truth, and *vraisemblance*).

The construction of the reader in the text depends on how the narrative discourse is structured: its register, lexicon, assumptions of background knowledge, and shared morality. In the mainstream United States, the default reader in the text is White, male, and middle class (cf. Morrison, 1992). For many readers, it is difficult to discern that the reader in the text is White, male, and middle class either because they are White, male, and middle class or because they have learned to read from that perspective. Yet readers who are not White, male, and middle class may recognize that the reader being addressed is not them. In classrooms, such differences may create a problem, requiring students to shift their identities so they can respond to a narrative text in a manner prescribed by their teacher. The issue is not just simply the preference and prerogative of the author in constructing the reader in the text, but also of the power of literary traditions and ideologies in fix-
Another aspect of hegemonic power in defining authors and readers concerns who is eligible to create particular kinds of narratives. For example, consider the Mass-Observation Project—a project that involved ordinary people across Britain writing about daily life and social history. Their writings are collected and stored at the Mass-Observation archive. Sheridan et al. (2000) described the criticisms made of the Mass-Observation Project by established academic scholars. A large part of that criticism was that the ordinary people who wrote for the Mass-Observation could not legitimately assume the position of author because they lacked the credentials and formal training of historians, anthropologists, or sociologists. They could be viewed, however, as subjects taking part in a kind of survey research. Thus, the writers for the Mass-Observation Project could be researched but, by definition, could not be legitimate researchers or authors. Beyond the specific example of the Mass-Observation Project, the point here is that traditions of reading and writing provide opportunities and limitations on who can be an author (or reader) of what narrative texts. Stated simply, it is not just that some people have no access to voice in representing themselves, others, or their world, but rather that the structuring of social institutions (such as academia) is such that to conceive of them as authors is a nonsequitor.

Cognitive Processes and Power Relations in Narrative Text. The prior section moves to the foreground problematics about power relations and identity in the analysis of narrative text. In this section, I examine power relations and cognitive processes in the analysis of narrative texts. Although studies of cognitive processes involved in the understanding of narrative texts rarely discuss power relations, there is nothing inherent in the study of cognitive processes that eschews consideration of power relations.

Cognitive studies of readers’ comprehension of narratives almost uniformly characterize the reader as constructing a situation model (a mental micro world) that requires bringing to the narrative text world knowledge and the making of inferences (to fill in omissions inherent in any narrative; see Graesser & Wiemer-Hastings, 1999, for a review of research). Whether, how, and what a reader apprehends about the social relationships and identities embedded in a narrative text, its plot, characters, setting, underlying themes, set of events, the micro-world (situation model), and so on, constructed in/with the text depends, in part, on the specific cognitive and linguistic processes and strategies the reader applies to the narrative text. Studies of cognitive
processes involved in understanding a narrative text suggest that there is variation in what cognitive processes and strategies people apply depending, in part, on the goal or task, how they interpret the task, differences in stored mental models of events, developmental level, relevant background knowledge and experience, sociocultural background, and situational factors (Bower & Rinck, 1999; Goldman et al., 1999; Graesser & Wiemer-Hastings, 1999; Hicks, 1991). Studies have also suggested that different types of narrative texts may demand different sets of cognitive strategies. For example, Hudson and Shapiro (1991) listed three types of knowledge: (a) general event knowledge, (b) memory of a single episode, and (c) general social knowledge. These three types of knowledge correspond to three types of narratives: (a) scripts, (b) personal narratives, and (c) stories, respectively. In brief, although narrative texts may differ in structure, content, style, and mode, the narrative text does not solely determine what cognitive processes need to or should be applied. Rather, it is the complex interaction of the particular narrative text, reader, task, and situation that needs to frame investigation of cognitive processes.

Although cognitive approaches to narrative discourse are helpful to understanding what texts mean and how readers and texts are defined, theorists employing a critical perspective (e.g., Fairclough, 1989, 1992; Flower et al., 2000) have suggested that what texts mean also involves power relations (which are often hidden). Fairclough (1989, 1992) suggested that explicating the power relations embedded in a text requires explicating the local and institutional contexts of its production and consumption. For example, a newspaper article reporting a heroic act by a soldier needs to be examined in terms of the social institution in which it was produced (in this case, multiple social institutions including the government, military, and mass media) and their agendas with regard to promulgating institutional rationalities and discourses. Part of what critical discourse analysis does is make clear that meaning and interpretation of narrative discourse is always a political act and always involves power relations. Critical discourse analysis does not deny the involvement of cognitive processes, but rather complicates the understanding of cognition by not allowing cognitive processes to be hollow or apolitical.

Although not specifically building on Fairclough’s work, Flower et al. (2000) provided one example of bringing together in the study of narrative texts investigation of cognitive processes and power relations. Flower et al. insisted that all texts be viewed as interpretable from different and rival stances. For Flower et al., rival hypotheses about the meaning of a text derive from culturally and politically different locations, and thus the explication of the meaning of a narrative
text is always at a site of conflicting (rivaling) frames derived from the differing cultural and political locations of people. For Flower et al., readers should not be allowed to approach the task of constructing meaning from narrative discourse from the default perspective of the dominant society as if there were only one interpretation. Rather, readers need to learn to approach texts with rival hypotheses about constructing meaning based on recognition of a culturally, economically, politically, and otherwise diverse world. One way to interpret the approach by Flower et al. is that the choice to read from a single frame or from multiple, rivaling frames, is a political choice with profound implications for the cognitive processes involved in reading a narrative text. Both the choice and the resultant cognitive processes employed can be viewed as the result of power relations between those privileged by an exclusive dominant frame (the default, single frame) and those who would contest such privileging by problematizing the exclusivity of the dominant frame.

**Chronotopes and Narrative Text.** One of the hidden aspects of narrative discourse as a text is time. Rather than viewing time as a given of nature, time can be viewed as a human construction (Adam, 1994; Ricoeur, 1984). Part of what narratives do is impose on human action a definition of time. Indeed one could claim that narratives define and produce time (see Ricouer, 1984, for a detailed discussion of narrative and the production of time). This occurs in at least two ways. First, narrative discourse insists that human events (by which I mean all events and phenomena that are given meaning by people) be represented in a temporal manner, whether sequential, simultaneous, or otherwise. This is so not only of formal narrative texts, but also of the everyday narratives in which we understand what happened yesterday, this week, or over a broader scope. Second, in considering what will be, we construct narratives that guide our actions and others’ actions. In a sense, we act out narratives at both a micro (face-to-face) and macro (societal) scale. Thus, the events and phenomena of our lives take on meaning and temporality because we narrativize them.

As such, one of the tasks in the analysis of narrative discourse is to make explicit the definition of time being imposed. Bakhtin (1935/1981) used the concept of *chronotope* to analyze how narratives impose a definition of time and space.\(^1\)

We will give the name *chronotope* (literally, “time space”) to the intrinsic connectedness of temporal and spatial relationships that are artistically

\(^1\)The discussion of chronotopes here is taken from Bloome and Katz (1997).
Bakhtin analyzed chronotopes in various literary periods, showing how time and space are differently conceptualized and how the chronotope frames character development (or the lack of it), plot, and the meaning and significance of a novel. For example, Bakhtin analyzed the chronotope of Greek romances. In brief, the hero and heroine meet, fall in love, but are separated by events and have to go through a series of adventures and overcome obstacles to rejoin each other and marry. In this “adventure-time,” as Bakhtin (1935) called it (p. 87), although the hero and heroine may spend time in the adventure and travel to different places, they do not change, their feelings do not change, and the world does not change. Bakhtin’s analysis of the chronotope of the Greek Romance highlights the power of a chronotope to frame human agency and the relationship of the individual to the world in which she or he lives. In adventure-time, things happen to people—people do not initiate action. There is no evolution or growth. The identity of people, places, and things at the beginning is affirmed throughout and at the end.

The concept of chronotope can be used to analyze a text whether it is fictional, historical, scientific, or a transcript of a classroom event (e.g., Bloome & Katz, 1997). Whether one uses the concept of chronotope or another construct is perhaps less important than making visible the usually invisible constructions of time being imposed on human action.

Narrative as Event and Practice

To define narrative as events and practices is to focus attention on how people use narrative discourse(s) to interact with each other and on the meanings and possibilities they create through those interactions. What constitutes a narrative from this perspective is not given in the set of words on a piece of paper (or analogously in spoken, pictorial, or electronic form), but rather narratives are constituted through and are part of the interaction of people with each other.

A social or cultural practice can be defined, briefly, as a group’s shared way of doing some social or cultural activity within a type of social and cultural setting. It is an abstraction of how that activity is accomplished. However, it is an abstraction that is located not just in shared minds, but in the interactional nature of the activity (people hold each other accountable for enacting a social practice appropriately), its history, and material conditions. Cultural events (hereafter
events) are the specific, particular, and situated enactments of social and cultural practices. Events can be viewed as sites where people create and re-create meanings, identities, social relationships, possibilities for future events, and practices. It is in events where people adapt and transform extant cultural practices. Thus, the study of cultural practices and cultural events are inseparable.

Bauman’s conception of narrative exemplifies the view of narratives as events and practices (Bauman, 1986; Bauman & Briggs, 1990). He argued that narrative texts do not exist alone or in isolation, but only exist in some event in which they are performed. Citing Jakobson (1971) and Benjamin (1969), Bauman (1986) stated that: “narratives are keyed both to the events in which they are told and to the events that they recount, toward narrative events and narrated events” (p. 2). Part of the importance of Bauman’s conception of narrative is to link narratives and storytelling in ways that shift the analysis of narratives from solely a literary task to a sociolinguistic and ethnographic task, similar to the implicit social and cultural work that participants in a storytelling event must do.

Building on Bauman’s framework, Solsken and Bloome (1992) made a distinction among story, narrative, and storytelling events, although they viewed the three as inseparable:

A story is a chronological sequence of events abstracted from experience. That is, experiences are not inherently packaged as stories with beginnings, middles and ends nor do experiences necessarily provide coherent relationships between events. Rather, story transforms experience into events and imposes boundaries, a chronology, and a set of coherent relationships on experiences. This is axiomatically so regardless of whether story-construction occurs in a reflective mode (constructing a story of past experience) or in real-time (constructing story during experiences as they occur), whether a group or individual is involved. . . . When people construct a story they are constructing an abstraction that, by itself, has no realization. Rather it is realized in narrative. A narrative is the text of the story. The text may or may not present the story chronologically. . . . Narratives do not exist by themselves. They exist only in storytelling events. Storytelling events can involve several people or one individual, they can be formally labeled storytelling events (e.g., show and tell) or embedded in events with other labels. In a storytelling event people act and react to each other while producing the narrative. That is, there are always at least two simultaneously activities in a storytelling event—the telling of a story (or stories) and the making of the storytelling event itself. (p. 4)

Similar distinctions have been made by Atkinson (1991), Genette (1980), and Le Guin (1981), among others, although the specific terminology varies.
One implication of definitions of narrative given by Bauman, Solsken, and Bloome, and similar others is that narrative discourse must be analyzed within the context of the social events in which they occur. Part of this implication is that storytelling can be viewed as a way in which people socially define and position themselves with regard to others, social institutions, and so on. A narrative is not simply a result of what the speaker or writer has produced, but the result of a sort of co-authorship between the speaker/writer and listeners (Ochs, 1997). As Ochs (1997) noted:

Scholars of narrative have argued that narratives are authored not only by those who introduce them but also by the many readers and interlocutors who influence the direction of the narrative (Bakhtin, 1935, 1953; Bauman, 1986; Goodwin, 1981). . . . The interactional production of narrative maintains and transforms persons and relationships (Miller et al., 1992). How we think about ourselves and others is influenced by both the message content of jointly told narratives and the experience of working together to construct a coherent narrative. (p. 183; italics added)

In part, the prior arguments question the use of narratives. That is, what do they accomplish? From the perspective of narrative as text, narratives are viewed as primarily for communicating a story, transmitting information from one person to another (even if that information is a fabrication). Yet from the perspective of narrative discourse as event and practice, other social purposes may be primary. Narratives and their telling can be used to establish social identity and social relationships such as friendships, families, group cohesion, social hierarchies, emotional bonds (e.g., love, care, joy), among other social functions (see Bloome et al., 2000; Champion et al., 1999; Dyson & Genishi, 1994). When a primary social function of a storytelling event is other than the communication of a narrative, the analysis of narrative discourse needs to follow from the social functions of the story, the narrative, and its telling. In brief, the analysis of narrative discourse, when narrative is viewed as event and practice, cannot be separated from the analysis of the social relationships and dynamics of the people involved in the storytelling event.

Consider a story reading event between a parent and her or his young child. Inasmuch as many schools, libraries, public service TV advertisements, and so on in the United States and elsewhere have promoted reading to one's children as good parenting, engaging in an event that can be counted as reading to one's children (a cultural practice) may define the parent as a good parent and the family as a good family. How the parent and child interact over, beyond, and through
the text during the event may have implications for how the child is socialized to use language and narrative discourse (Heath, 1982; McCabe, 1997). Such socialization may be an explicit purpose of engaging in bedtime story reading. The routines established for such reading events provide the child with one set of cultural practices for the use of narrative in a particular set of situations, as well as opportunities to play with those practices to accommodate a variety of personal and interpersonal goals. The stories selected to be read may provide the child with insight regarding what is considered an appropriate story by his or her cultural group for such events as well as helping to define what a story is (i.e., what incidents constitute a story and how they may be connected). Depending on what happens during the event, the child and parent may give totally different understandings of plot to the narrative (assuming they are orienting to plot). They might discuss the meaning of a book, constructing together a meaning for the narrative idiosyncratic to themselves and that event. Of course, it may also be the case that the bedtime story reading event has little to do with the narrative or story in the book, per se, and is more so about parent–child relations and socialization into being a particular kind of reader and person (Bloome, 1985).

**Storytelling, Intertextuality, and Social Institutions.** Beyond the question of shifting the analysis from text to the event is examining how particular storytelling events reflect cultural practices connected with various social institutions such as church and school. Bedtime story reading is a culturally specific practice and, when conducted in particular ways, may provide children with cultural capital (cf. Bourdieu, 1977, 1991), ways of doing narrative, that privileges their participation in school settings (Heath, 1982). Children who do not participate in such practices may find themselves at a disadvantage in school settings (Heath, 1982)—not because they lack narrative skills or because they do not value stories or language play, but rather because they have not practiced those specific narrative discourse practices whose display is valued by schools (McCabe, 1997). Thus, among the tasks involved in the analysis of narrative discourse when viewed as event and practice are (a) investigation of how storytelling is done across situations and institutional settings, and (b) the role these settings and situations play in providing opportunities and privileges.

One of the mechanisms involved in connecting storytelling practices within and across social institutions is intertextuality. Briefly defined, *intertextuality* is the juxtaposition of different texts. Traditionally, intertextuality has been defined as a property of a text. That is, intertextual links are referenced in the text of the narrative. For exam-
ple, the opening of *Huckleberry Finn* makes explicit reference to *The Adventures of Tom Sawyer*, the *Periodic Table* by Levi (1984) juxtaposes his autobiographical narrative account with a scientific text. At a broader level, a narrative text may imply another narrative text through membership in a similar genre (e.g., creation stories; see Hamilton, 1988) or geographic location (e.g., Higgs & Manning, 1975). For some literary theorists, all texts imply other previous texts (Bakhtin, 1935), so that the question to ask is not whether intertextuality exists in a narrative text, but what voices, discourses, and texts have been brought together to create a particular narrative text.

From the perspective of narrative as event and practice, intertextuality is not located in the text, but in the interactions people have with each other. Participants in a storytelling event need to make clear to each other the relationship of the narrative text and story to other narrative texts and stories, and they need to make clear the relationship of the storytelling event to other events (Bloome & Egan-Robertson, 1993).

Methodologically, researchers need to examine how conversational structures make possible a series of stories and the linguistic mechanisms that people in interaction with each other use to socially construct connections among stories and events. Bloome and Egan-Robertson (1993) listed three conditions for the social construction of intertextuality. First, a proposed intertextual link must be acknowledged. Interlocutors must jointly and publicly acknowledge that an intertextual link has been proposed. Second, a proposed intertextual link must be recognized. The texts being juxtaposed must be jointly and publicly recognized by the interlocutors—they must have some sense of knowing the texts. Third, the intertextual link must have a social consequence for the joint construction of meaning, social relationships, social identities, or some combination thereof. Bloome and Egan-Robertson also listed three aspects of the social construction of intertextuality: (a) intertextual substance, (b) intertextual process, and (c) intertextual rights. *Intertextual substance* refers to the content of narratives being juxtaposed—for example, the juxtaposition of Tom Sawyer and Huck Finn, Daniel in Doctorow’s (1996) *Book of Daniel*, and Michael Meeropol, a teacher lecture about the conduct of ethnographic research and a textbook the students have been assigned to read. Intertextual substance is only established when the three conditions noted earlier are present. *Intertextual process* refers to the ways in which the conditions for the social construction of intertextuality are established (through explicit comments, head nods, eye contact, ellipsis, etc.). *Intertextual rights* refers to the differential distribution of privileges to participate in the social construction of intertextuality.
For example, a university professor may be able to propose an intertextual link between Shakespeare’s *The Tempest* and expositions on slavery and Western colonialism (see Takaki, 1993) and have the intertextual link acknowledged and recognized with given social consequence. A low-achieving student in a low-track high school English class might not get the chance to propose such an intertextual link; if he or she did, it might not be acknowledged, recognized, or have social consequence. Indeed, it might be dismissed as a *nonsequitur*.

One way to describe the theory-method issue raised by the social construction of intertextuality is that people are never telling (or listening to) one narrative, they are always telling (and listening) to many; they are never in one event, they are always in many. People need to make clear to each other what the set of narratives and events are, and they need to make clear to each other how they will establish (negotiate) what the sets of narratives and events are.

**Language Ideology and Narrative as Text Versus Narrative as Event**

The distinction between viewing narrative discourse as text versus narrative discourse as event and practice is not just a matter of research perspective. It is a dynamic that structures ideological issues with regard to language socialization. For example, in formal academic storytelling events like sharing time, book reading, or report making, an emphasis is often placed on sequential structure (McCabe, 1997) and other referential aspects of storytelling events (e.g., understanding a story, telling a coherent story, writing a story or report with an integrated beginning, middle, and end). Such an emphasis on referential aspects in school can be viewed as a language socialization process. That is, it is a socialization process whereby children are taught to view narratives as primarily (or perhaps as only) text (cf. Bloome et al., 2000). By contrast, Bloome et al. (2000) described the storytelling of preschoolers and kindergarteners in which a referential orientation (associated with narrative discourse as text) is supplanted by a performative orientation (associated with narrative discourse as event and practice). The storytelling of the children focused on entertainment and social relationships among the interlocutors (teller and audience) at the expense of (or complementary to) formal narrative structure and coherence. Bloome et al. suggested that there are always both performative and referential aspects of narrative production, and narrative discourse is always both event and text. They argued that what is at issue is that which is foregrounded, suggesting that in school narrative as text is foregrounded, increasingly so as students move up the
grades. Michaels (1991) used stronger language. She argued that “much of what goes on in urban, public school settings promotes the dismantling or breaking down of narrative performance and artistry—in favor of alternative forms of meaning making deemed more scientific, rigorous, reliable, intelligent, or important” (p. 303; italics original). By dismantling, Michaels meant “the destruction or suppression of a particular narrative style in toto” (p. 303). Although alternatives to the dismantling of narrative performance and artistry are possible, what is at issue is the relationship of narrative discourses, especially those grounded in communities other than middle-class or academic communities, and the language and literacy practices promulgated by schooling.

It is important to note that performative aspects of narrative discourse do not atrophy on account of a school emphasis on narrative as text (although students may learn to background performative aspects in particular types of situations). Indeed, performative aspects may be of primary emphasis in some events outside of academic settings, and in some academic settings performative aspects may still be important to emphasize (as in giving a good lecture), although such an emphasis might be denied. What is at issue for researchers is the situated definitions of discourse, what implications those definitions have for the structuring of knowledge, comprehension of text, identity, social relationships, the organization of social life, power relations among individuals, institutions, social groups, and so on. As is true of the other problematics discussed in this chapter, the issue is not whether one definition of narrative discourse is right and one wrong, but rather what insights, possibilities, and limitations the tensions between the definitions provide.

**NARRATIVE AS STRUCTURE VERSUS NARRATIVE AS SUBSTANCE**

Structure and content work together in a narrative to express meaning. Thus, even for heuristic purposes, it may seem inappropriate to separate narrative structure from narrative content. I do so in this section for two reasons. The first is to foreground structural issues in the analysis of narrative discourse. The second is to examine tensions between conceptions of narrative discourse that foreground structure and those that foreground content. By narrative structure, I refer to the naming and organization of the components of narrative discourse as well as to those cognitive, linguistic, and social processes employed in producing or comprehending a text when abstracted from content. For example, when narrative is viewed as a text, one might describe the
narrative structure as consisting of rising action, climax, falling action, and coda. One might describe the cognitive processes employed in composing a text without reference to particular content (e.g., Flower & Hayes, 1981). Similarly, if narrative discourse is viewed as event and practice, one might describe narrative structure as components of a storytelling event (e.g., signaling the beginning of a story, retelling events, maintaining the floor, engaging audience, signaling the end of the storytelling). One might also generate rules that describe the turn-taking processes of a narrative practice without considering the content of the story or the social interactions.

With regard to narrative content, I make a distinction between narrative content and narrative substance. Narrative content refers to the information in the narrative. To use a metaphor from the language of cognitive scientists, it is the text base of the narrative. The concept of narrative substance assumes that there are meanings created in the use of a narrative that lie outside the text base. Building on activity theory, Beach (2000) distinguished between the content of a narrative text and student understanding of the relationship among acts, beliefs, and goals as contextualized by a social system (what I call narrative substance). He wrote:

[Activity theory] moves away from defining text understanding and learning simply in terms of readers’ cognitive processing to frame text learning as embedded in complex participation in object-driven activities (Neuman & Roskos, 1997). . . . Much of literature instruction and textbook questions operates primarily at the level of inferences about characters’ acts, beliefs, and goals in responses to questions such as “What is this character doing?” “What does the character believe about other characters?” and “What is the character’s purpose?” (Beach & Wendler, 1987; Graesser, Golding, & Long, 1991). Students also need to be able to infer the ways in which these acts, beliefs, and goals are driven by objects or motives operating within larger social and cultural activities (Smagorinsky & Coppock, 1995; Smagorinsky & O’Donnell-Allen, 1998). Students need to be able to infer how specific acts, as tress, are shaped by the woods of an activity system. (pp. 238–239)

Although the distinction between the content and substance of a narrative discourse does not necessarily need to be framed by activity theory, the content of a narrative text can be defined as the description of the plot, characters, setting, events, symbols, themes, and so on of the narrative text (the trees in Beach’s formulation). By contrast, the substance of narrative discourse can be defined as the what, how, and where of narrative discourse (the woods in Beach’s formulation) in the structuration of social systems (cf. Giddens, 1984) at the levels
of face-to-face events, cultural practices, social institutions, and other units of social systems. Given the definition of narrative substance, the heuristic distinction made earlier in this section between narrative structure and narrative content should be rewritten as narrative structure and narrative substance.

**Narrative Discourse as Structure.** Discussion of the structure of narrative discourse has primarily concerned the organization of linguistic representation—how a series of events are configured to yield the *vraisemblance* of sociocultural life, the experience of catharsis, and the poetic aesthetic. Narrative structure can be seen to function, paradoxically, both to reflect and shape human experience.

One approach to analysis of narrative discourse has been to specify what the components of a fully formed narrative are and in what order they occur (or should occur). Story grammar (e.g., Johnson & Mandler, 1980), plot unit analysis (e.g., Vine, 1994), and high-point analysis (Labov, 1972) are examples. Almost all of the major approaches to structural analysis of narratives include a climax component or something analogous. A climax, in addition to potentially creating a cathartic effect, structures events and people within a particular world view. The application of story grammar, high-point analysis, or similar approaches to the analysis of narrative discourse can be problematic. Consider their application to a story such as Singer’s (1984) “Naftali the Storyteller and His Horse, Sus.” In this story, a young boy grows up to be a book peddler. He travels from village to village in a wagon pulled by his horse, Sus. They get old, retire, and die. Although one could identify some story grammar components, critical components such as a climax are obviously absent. It would be hard to argue that their absence makes the story of less quality, importance, or entertainment. Stories such as “Naftali the Storyteller and his horse, Sus” challenge the notion that a set of structural components defines a narrative or distinguishes a good narrative from a poor one.

Of course, it could be argued that stories such as “Naftali the Storyteller and his horse, Sus” play against a typically structured narrative with a climax. Therefore, one might claim that stories such as “Naftali the Storyteller and his horse, Sus” actually do involve the components and organization structure of typically structured narratives. Assuming for the moment that such an argument has validity, it raises an important methodological construct in the study of narrative discourse. What is at issue in the application of story grammar analysis, high-point analysis, and similar systems is not necessarily identification of the components and their order within a targeted narrative, but rather how story grammars, high-point analysis, or any other struc-
tural heuristic reflect an interpretative framework that authors (tellers) and readers (audiences) might use to create meaning, aesthetic experience, and judge the adequacy of a narrative. For example, Champion (1998; Champion et al., 1995, 1999) showed how the application of high-point analysis to young African-American children’s narratives provided one interpretation, whereas analyzing the same narratives in terms of social relationships provided another. The methodological question raised by Champion and her colleagues is when, how, and by whom such analyses are appropriate (see also McCabe, 1997).

So far the discussion of narrative discourse as structure has focused on the structure of narrative as text. However, structural questions can also be asked about narrative discourse as event and practice. For example, ethnomethodologists have provided structural approaches to storytelling, focusing on what participants need to do to accomplish a story—to coproduce it conversationally. As Jefferson (1978) argued, “Stories emerge from turn-by-turn talk, that is, are locally occasioned by it, and upon their completion, stories re-engage turn-by-turn talk, that is, are sequentially implicative for it” (p. 220; italics original). A story must be topically coherent with what is being talked about in a conversation and then inserted into the conversation—a task that requires speaker and coparticipant to align themselves as recipients of the story. One implication is that narrative discourse must be oriented both to what has occurred in the conversation and its projected reception. After a story has been told—that is, after a series of storytelling segments, the conversational coparticipants must reengage their turn-by-turn conversation. Various conversational exit devices may be used to mark the transition and jointly acknowledge its end. The occurrence of a told story implicates coparticipants to respond to the story.

Quasthoff (1997) provided a structural approach to the analysis of storytelling events building on ethnomethodological theory. Quasthoff described a series of discourse devices, forms, and jobs that need to be jointly accomplished within a storytelling event. The discourse jobs include: display of relevance, topicalization, elaboration, closing, and translation. One implication of this approach is to focus attention on (a) how people interactionally accomplish a storytelling event, (b) how they learn to do so, (c) what resources they have available for doing so, and (d) how the discourse forms, devices, and jobs frame cognitive processes and the realization of the narrative text and story.

Unity and Its Discontents. One of the issues involved in the structure of narrative discourse concerns the construct of unity. The construct of unity can be applied to both the view of narrative discourse as text and
narrative discourse as event and practice. In the former case, ques-
tions are asked about the unity of the text; in the latter case, questions
are asked about the unity of the event. Although the use of unity as a
heuristic cannot be assumed for all narrative discourse, it can none-
theless be helpful in analysis of a broad range of narrative discourse if
for no other reason than unity has been a major heuristic of Western
poetics since Aristotle (1997). Unity typically refers to the presence of a
beginning, middle, and end—to the relationship of components of a
narrative text or event to each other and to the nature of the aesthetic
experience. Even a novel that contains multiple voices and multiple
genres (e.g., Dos Passos, 1951) can nonetheless be considered struc-
tured by the construct of unity as the multiple voices and genres are or-
chesterated (like a choir or jazz ensemble) toward a unified story,
rhetorical and aesthetic experience. This is similarly so when consid-
ering narrative as event. A court trial in which multiple people recount
different parts of an event, multiple genres are used, and there are con-
tested recounts nonetheless has a beginning, middle, and end.

Although some literary critics and others have argued for unity as an
indication of quality and value (e.g., Brooks & Warren, 1960), others
have argued against unity, suggesting that it provides problematic
worldviews and reifies people and events. For example, Norton (2000)
argued that the construct of unity is associated with “an act of mas-
tery—one that establishes a necessary, triumphal relation between
mean and order” (p. 82) associated with “masculine anxiety over how
to understand and institute rhetorically the dynamics of domination”
(p. 82) and “in effect constitutes poetic value as the formalization of the
social supremacy of masculinity” (p. 82). Thus, following Norton’s ar-
gent, approaches to the analysis of narrative structure require
questioning underlying cultural and political assumptions about
unity. Indeed, literary works such as Shua’s (1998) *The Book of Mem-
ories* have sought to undercut the construct of unity, organized instead
around ambiguity and indeterminacy. To the extent that identity is tied
up with coherent and linear narratives about the self (cf. Currie, 1998),
the threat to narrative unity can be viewed as a threat to a coherent
identity both with regard to the characters in a narrative as well as to
tellers, readers, and listeners. This tension between unity and indeter-
minacy can be characterized as a kind of cultural schizophrenia. In his
history of narratology, Currie (1998) described this schizophrenia as
follows:

The schizophrenia of contemporary culture is partly about the super-
market shelving—literal and metaphorical—of history. . . . For the con-
suming classes, the supermarket offers a kind of compressed tourism
which erodes the traditional relationship between identity and place. It is
an experience dedicated to cultural diversity which offers the shopper an international spectrum of possible identifications, where the signs of other cultures compose the shopper’s identity through affiliation with various ethnicities, as if shopping itself were a process of identity construction. (p. 104)

In brief, the issue of narrative unity is less one of whether a particular narrative has or does not have unity, but of how unity and indeterminacy are defined within the local and institutional contexts in which people are consuming and producing narratives. Following Currie (1998), both unity and diversity of narrative discourse become commodities that, metaphorically, are available off a supermarket shelf. That is, one has available many different ways of constructing narrative unity, and the selection of how to construct narrative unity for one’s life is little different than the consumerism of selecting one item versus another from a supermarket shelf. Although Currie’s perspective may be harsh, one implication of his view is to raise questions about the social and political contexts of constructing unity.

Narrative Discourse as Substance. In a variety of textbooks and professional education books on the teaching of composition, attention is focused on composition processes. Similarly, pedagogical discussions of how teachers might respond to student writing often focus on composition processes and text structure. Similar discussions occur in scholarly discussions of the assessment of student writing. There are analogous discussions in the teaching and assessment of reading. Indeed, the teaching and assessment of reading are rarely linked to the content of what students read. One reaction to these pedagogical approaches to reading and writing that focus exclusively or heavily on process has been to raise questions about the potential danger and harm to students by ignoring the substance of what is written and said (Gilbert, 1993; Lensmire, 1994; Miller, 2000). For example, Gilbert (1993) described how a group of boys in an elementary classroom wrote stories that denigrated various girls in the classroom and some of the boys who were outside the in group. Their stories contained violent acts against the girls and the female teacher. Yet it is not just the content of a narrative text that needs to be considered, but also the meanings that the use of a narrative text might have within particular contents. For example, Möller and Allen’s (2000) study of a classroom discussion of The Friendship (Taylor, 1987) made clear that the substance of narrative discourse is located not solely within the narrative text, but in the historical and interactional events that surround the use of a particular narrative.
Articles such as Beach’s, Gilbert’s, Lensmire’s, and Möller and Allen’s raise questions about the adequacy of analyzing narrative discourse solely in terms of empty structures and processes. Among those questions raised are: (a) What constitutes the substance of narrative discourse? and (b) Where is the substance of narrative discourse located?

One way to address the question of what constitutes the substance of narrative discourse is through a distinction that Bruner (1986) made between the landscape of action and the landscape of consciousness. By *landscape of action*, Bruner basically referred to the plot—to the events and actions that occur in a narrative. The *landscape of consciousness* refers to the emotional states, intentionalities, goals, perceptions, motivations, and so on of the protagonists. Feldman et al. (1990) argued that attention to the landscape of consciousness is especially important as it is “our form of self-understanding” (p. 30) and the form of modern literature.

The distinction between the landscape of action and the landscape of consciousness has primarily been addressed as a structural or cognitive issue. It is a structural issue both by definition and by how the two landscapes are related to each other. As a cognitive issue, the distinction concerns how readers/listeners understand the mental states of characters, their goals, and their motivations. When the landscape of action and the landscape of consciousness have been described with regard to a specific narrative, the description has primarily been a description of content. However, here I want to use Bruner’s concepts of the landscape of action and the landscape of consciousness to raise questions about what constitutes the substance of narratives.

The landscape of consciousness can be embedded within the broader cultural enterprise of establishing personhood. That is, what constitutes a human being, how is a human being defined, and what attributes does a human being have? As Egan-Robertson (1999), Gergen and Davis (1985), and Kirkpatrick (1983) made clear, personhood varies from culture to culture, situation to situation, and over time. Apprehension of the landscape of consciousness is not just a matter of comprehension (receptive or constructive comprehension), but a political and cultural process through which definitions of personhood are proposed, established, and contested.

For example, consider “A Way of Talking” (Grace, 1991), a short story about two Maori sisters who challenge how members of the dominant New Zealand society refer to the Maori men who work for them. The story is an explicit challenge to who counts as a human being within that community and the way of talking that produces the bound-
aries of personhood. A similar contesting of who counts as a human being is taken up in “Deportee,” a song by Woody Guthrie based on a newspaper article reporting a plane crash. The newspaper article gave the names of the dead pilot and copilot and then described the other passengers as deportees. In the song, the landscape of action includes not only the plane crash, but also the story of migrant farm laborers. By changing the landscape of action from just the boundaries of the plane crash story to the broader story of the exploitation of poor people who migrate from Mexico to the United States, and by juxtaposing the two narratives (the one in the song with the one in the newspaper article), Guthrie foregrounds the social and power relationships between migrant workers from Mexico, the dominant society, the mass media, and the state. With regard to the landscape of consciousness, Guthrie shifts the focus from the families of the pilots (by naming the pilots, readers can infer social relationships with family members and friends who will mourn their loss) to the Mexican laborers and their families and friends who are the victims of the plane crash, labor exploitation, poverty, and namelessness. The juxtaposition of the landscape of consciousness in the newspaper article and that of the song foregrounds the issue of personhood. It foregrounds who counts as a person, who gets to have the attributes of family and friends who will mourn their death, and the loving and caring relations that are copresent in the concept of mourning.

Beyond questions about the substance of narrative texts, both “A Way of Talking” and “Deportee” also raise questions about how what happens in storytelling events constitutes narrative substance. Building on Foucault (1972) and Silverstein (1976, 1993), Briggs (1996) raised this issue explicitly. He wrote:

discourses consist less of messages generated within circumscribed linguistic and social systems than as a means of determining the norms and limits regarding who can say what to whom and the relations of power created in the production and reception of discourse. In proposing the term metadiscursive practices, I draw on Foucault (1977 [1975]) in identifying discourses that seek to shape, constrain, or appropriate other discourses. By presenting discourse in narrative form, narrators and their audiences gain access to a range of metadiscursive practices for shaping social interactions as well as the production and reception of discourse (p. 19; italics original)

With regard to narrative discourse as substance, the methodological issue is not whether one defines narrative as text or narrative as event
and practice nor the specific disciplinary frame one employs, but whether the frame through which one is examining narrative discourse provides mechanisms for understanding how narrative discourses define and structure relationships among people.

Part of what constitutes the substance of narrative discourse is its location. I use location here as different from the physical setting or scene in which a narrative is produced or used. Rather, I refer to the cultural systems within which narrative discourse is located, and, more specifically, I refer to the level of cultural system (e.g., macro, micro). Following Ochs (1997), I locate the substance of narrative discourse locally and in its particularity. Ochs wrote:

Given the variety of modes and genres that realize narrative activity, it is an enormous task to consider how narrative is rooted in cultural systems of knowledge, beliefs, values, ideologies, action, emotion, and other dimensions of social order. Typically, cultural analyses of narrative focus on a particular context of narrative activity, for example, spoken or sung narrative performances . . . .mythic tales, . . . . conversational narratives of personal experience, . . . . reading stories, . . . . writing stories, . . . . gossip, . . . . or classroom narrative events. Narrative in each of these contexts is rendered meaningful vis-à-vis some property of local ethos—for example, an orientation towards autonomy or intervention, explicit moralizing, sacredness of text, facticity of text, imagined selves, social asymmetries, and so on. (p. 185)²

One methodological implication of Och’s discussion is that an analysis of narrative discourse requires a thick description (cf. Geertz, 1973) of the location of the production and use of narrative discourse.

Part of a thick description requires attention to particularity. By particularity, I refer to the linguistic concept developed by Wittgenstein (1970) and elaborated by Becker (1988), Geertz (1983), and Bloome and Bailey (1992), among others. Particularity is often defined by contrast. For example, Geertz (1983) argued for an approach to anthropology that focuses on cases and interpretations as opposed to rules and generalizations. Becker argued for approaching language through the study of particular texts and conversations. Becker (1988) writes:

[A particularity approach to the study of language works] with a particular text grabs the theories as they come by and celebrates them by applying them and learning their lessons. But the discipline is not in the theories. The discipline is in the particularity of the text-in-context. . . . (p. 30)²

²The ellipses in the quotation replaced a substantial list of example research studies.
Thus, particularity is both an achievement of daily life (its practices, events, texts, and conversations) and a stance taken by those researching it (whether researchers or ordinary people attempting to understand what is happening around, with, to, and by them). However, Bloome and Bailey (1992) went further. They argued,

A concern for the particular inherently takes seriously what people do in their everyday lives, refusing to homogenize the events of people's lives into a nameless and faceless set of general tendencies or rules. . . . To do this is a political act. (p. 197)

With specific regard to the study of narrative discourse, one implication of viewing the location of narrative substance as local and particular is to shift the method and goal of research on narrative discourse from the application and generation of rules and generalities to the study of particular stories, narratives, and storytelling events within the local contexts in which they occur, as they reflect, refract, and bump up against other stories, narratives, and events. Such a shift in methods and goals is consistent with Geertz's (1983) discussion of local knowledge in anthropological inquiry. Further, such a shift does not mean overlooking broader social dynamics. As Bloome and Bailey (1992) wrote:

It is within everyday events that broad sociological processes are revealed and defined. It is not so much that in the event there is a microcosm of the world, but rather that any event, and its particularity, does not happen in isolation from other events, the past, or the future. Nor do events occur in isolation from material conditions, or, as Delamont (1983) points out, in isolation from power relationships. (p.197)

It is similarly so with narrative discourse.

Final Comments

This chapter is organized around two tensions inherent in the study of narrative discourse: the tension between viewing narrative discourse as text and viewing it as event and practice, and the tension between viewing narrative discourse as structure and viewing it as substance. These two tensions are heuristic. They (a) provide productive points of entry for the study of narrative discourse, (b) distinguish different approaches to the analysis of narrative discourse from each other, (c) frame complementary relationships among different approaches to the analysis of discourse, and (d) create a vantage point to examine theory–method linkages.
The discussion of the two tensions, which I also called problematics, was deliberately multidisciplinary. Further, issues in the discussion of one topic often reappeared in discussion of other topics. The boundaries between one topic and theoretical construct and another were often unclear. Such a situation is characteristic of exotropic theories. Unlike endotropic theories that help define, categorize, and systematize knowledge, exotropic theories force theoretical constructs to bump up against other theoretical constructs to generate insight and generative debate. Rather than characterize narrative discourse as a complex system waiting to be formalized within a set of endotropic theories, this chapter foregrounded and valorized the complexity, uncertainty, and problematic nature of the analysis of narrative discourse as reflective of those qualities of narrative discourse. In brief, although one can tidy up complexity, in so doing one no longer has complexity to study.

Given the problematics, tensions, and debates described in this chapter, it would be fair to ask, “What is (are) the narrative(s) being told in this chapter?” and “What social and ideological work does it seek to accomplish?” I view the narrative being told in this chapter as part of a broader narrative in the field of language study of the struggle between transparent and representational views of language on the one hand with views of language as complex social and cultural practices, events, and joint action on the other hand. The study of narrative discourse requires investigators to either anchor down in a method and theory or to try and ride the “waves, tides, and cross-currents” of multiple exotropic theories with all of their tensions, complexities, and indeterminacies.

ACKNOWLEDGMENT

I gratefully acknowledge editorial comments from Art Grasesser, Susan Goldman, and Laurie Katz. Whatever flaws remain are solely my responsibility.

REFERENCES


8. NARRATIVE DISCOURSE


Literary reading has been a topic of inquiry among scholars of literature and educationalists for nearly three quarters of a century (Richards, 1929; Rosenblatt, 1937). Among empirical researchers, in contrast, attention to literature is quite recent, with most studies having been carried out only over the last 20 years. The literary field is fraught with controversy, however. Basic differences over the object of study militate against the emergence of a single paradigm for empirical research. Disagreement over the nature of literature centers on whether literature is a fundamental category of discourse with distinctive properties or a cultural formation produced during the last 200 or 300 years (e.g., Terry, 1997) and sustained by specific conventions (possibly facing extinction in the face of new electronic media). In this chapter, therefore, although it is possible to elaborate a number of specific components of literary reading that have been studied empirically, at the present stage of research a coherent account of literary discourse remains out of reach.

One notable feature of the research to be discussed is its limited attention to questions of interpretation or literary meaning. In contrast, mainstream literary criticism has traditionally been dominated by a focus on interpretation carried out within one of two main traditions: either a hermeneutic approach centered on the text or a contextual approach that appeals to major cultural formations thought to impose certain requirements on literary production and reception (e.g., gen-
der issues or the economic and social concerns of the new historicism). When considering reading outside the academy, however, an emphasis on interpretation may be misleading, as Sontag (1964/1983) argued forcefully some years ago. Although readers are at times undoubtedly concerned with understanding what they read, this should not overshadow another and perhaps more primary mode of engagement, which is to experience literature—whether to appreciate its formal qualities, be aroused by a suspense filled plot, or suffer empathically the vicissitudes of its fictional characters. To be asked to generate an explanation of a literary work, as commonly occurs in the literature classroom or in many empirical studies, is perhaps atypical of most reading situations. Yet it is clear that the demand for techniques of explanation has tended to drive research on reading, which has been dominated by the prevailing cognitivist emphasis on the processes of comprehension (Kintsch, 1998). This is considered further in the studies described later, but the limitations of this approach to literary discourse are also suggested preparatory to outlining a range of other approaches to literary reading.

The predominant questions of this chapter are: What is literary discourse? Does it result in a type of reading different from that studied in mainstream discourse processing research? Among a number of possible markers of the distinction between literary and nonliterary processing to be discussed later, empirical research suggests that literary readers form specific anticipations while reading, that the interpretive frame may modify or transform while reading a literary text, and that markedly more personal memories are evoked during reading. There is evidence for a constructive role for feeling in the reading process—a process that may be driven in part by response to stylistic and other formal qualities. First, however, the relationship between empirical research and mainstream literary scholarship should be sketched because this continues to provide an important, if problematic, context for considering literary issues and framing empirical studies. Although the gulf between literary scholarship and empirical research remains wide, three issues in particular serve to illustrate the difficulties and prospects of this relatively new discipline: history of reader response theory, role of genre, and question of whether literature has distinctive qualities.

THE ROLE OF THE READER

Although reader response study had its inception with the work of I. A. Richards (1929) in his book Practical Criticism, Richards’s one foray into empirical study unfortunately suggested to the community of liter-
ary scholars that readers, as represented by the undergraduate students he studied at Cambridge University, were poor at discriminating between poems and badly in need of the guidance of the experienced literary critic (cf. Martindale & Dailey, 1995). The experience of the ordinary reader, in contrast to the professional reader, thus fell under a cloud—a fate confirmed 20 years later by the influential essay “The Affective Fallacy” of Wimsatt and Beardsley (1954/1946). This effectively placed an interdiction on attention to actual readers, whose responses were deemed impressionistic and relative. For example, the critic E. D. Hirsch (1967) referred to the ordinary reader’s “whimsical lawlessness of guessing” (p. 204) at literary meaning—an initial first step subject to correction in the light of what could be determined about the author’s intention.

When reader response criticism eventually emerged, with publications by Holland (1968), Fish (1970/1980), Iser (1978), and Jauss (1982), Holland’s work was confined to developing his own psychoanalytic approach, which concentrated almost exclusively on the stories of individual readers. In contrast, Fish, Iser, Jauss, and their followers remained at the level of theory. Fish proposed an affective stylistics of readers’ hesitations and errors that he considered an integral part of literary meaning. Iser, drawing on the phenomenology of Ingarden, placed reading within the reader’s “horizon of expectations,” in which the text’s gaps and indeterminacies called for constructive interpretive work. Jauss, who worked alongside Iser at the University of Constanz, developed a reception theory attentive to historical changes in literary reading. Although this work, offering new and suggestive theories of reading, has been influential in redirecting attention to questions about the reader, this generally consisted in postulated reader-based modes for interpreting literary texts. The study of actual readers was either neglected or actively discouraged. For example, Culler (1981) suggested that a study of actual readers would be fruitless because the critic’s focus of research should be on the conventions that he considered paramount in determining all reading, whether literary or nonliterary. These conventions could be examined in the numerous interpretations already available in the professional literature on a given text.

On the one hand, then, critics such as Fish or Iser hypothesized specific reading processes based on demonstrable features of literary texts and their purported effects. On the other hand, it turned out that attention to such features was constituted from the start by conventions of reading. Because readers were thought to acquire such conventions through a process of training, usually in the classroom, professional attention shifted away from considering what individual readers might actually be doing. Among the most influential formula-
tions of this view, Fish’s (1980) forceful and widely accepted assertion that the interpretive community to which a reader belonged determined any possible reading appeared to make reading a purely relativistic process. The comprehensiveness of this approach, which redirected attention away from the reader toward questions of culture and history, foreclosed attention to reading almost as soon as it had begun: The reader response project was described by one of its reviewers as “self-transcending” and “self-deconstructing,” suggesting “that it has a past rather than a future” (Freund, 1987, p. 10).

Yet the issues raised by the reader response theorists were of considerable interest, and in various ways continue to be reflected in empirical studies of literary reading. Thus, Schmidt (1982) and his colleagues, although motivated by a research tradition quite different from that of Fish, have placed the conventions of reading at the center of their approach. They have hypothesized that literary reading depends on two conventions: (a) the aesthetic convention (opposed to the fact convention that is held to apply in regular discourse); and (b) the polyvalence convention (opposed to the monovalence convention)—that is, the supposition that in a literary context readers recognize the possibility of multiple interpretations of the same text. If reading is held to depend on the acquisition of the appropriate conventions, we might also consider the impact of literary training on reading. This issue has motivated several studies of literary expertise, where the range of interpretive strategies shown by novice and expert readers has been investigated through empirical study (e.g., Graves & Frederiksen, 1991; Hanauer, 1995a). If readers differ according to their local interpretive community, cultural differences in reading the same text should also be discernible. This topic has been the focus of empirical studies by László and his colleagues (e.g., László & Larsen, 1991).

Other approaches that develop the suggestions of reader response theory are considered in more detail later. Meanwhile, two other specific issues raised by literary scholarship should also be outlined.

GENRES

Perhaps the most significant convention dominating contemporary literary scholarship is that of genre. Although interpretation has tended to dissolve the distinctiveness of the text by relating it to underlying structures of power and desire, genre focuses attention again on the specific qualities and structures of the text. The features of a text are determined by its particular generic form. Genre, which used to be considered descriptive, is thus now considered explanatory: Genre is held to embody certain social roles that govern the relation between
text and reader. As Bawarshi (2000) put it, “genres create a kind of literary culture or poetics in which textual activity becomes meaningful” (pp. 346–347); they “constitute the social reality in which the activities of all social participants are implicated.” Genre is said to provide an essential framework for reading. As Derrida (1980) insisted, “a text cannot belong to no genre, it cannot be without or less a genre. Every text participates in one or several genres, there is no genreless text” (p. 65). In this respect, genre can be understood as the defining context for all textual behavior, literary and nonliterary, constituting textual relations in such spheres as the university classroom (lectures, the production of assignments), the law court (speeches for the defense, instructions to the jury), as well as readers’ engagements with plays, novels, or sonnets. Moreover, genre governs what Halliday (1978) referred to as the register apparent in any given text—that is, the semantic and syntactic features that create the communicative situation, including the stance of the participants (cf. Viehoff, 1995).

Although Bawarshi (2000) argued that the primary theoretical question is whether genre is regulative or constitutive, he gave little consideration to the possibility of variance or play within a given genre. Our response to a sonnet, for example, is a product both of our relation to its obligatory formal features as well as the distinctive semantic or formal qualities that the writer has embodied within the constraints of the form. The rules of genre allow us to specify both what is conventional and unexpected (a view that would suggest that the laws of genre are regulative). As discourse structures, therefore, genres are characterized in part by the types of story grammars or schemata they call for; they specify situation models that characterize a given literary text and enable us to make predictions about how the text is likely to unfold. In this respect, genre theory provides a potentially rich resource for more precise empirical studies of literary reading, enabling us to build on the research (which so far is not specific to literary reading) on how readers construct and monitor situation models (Zwaan & Radvansky, 1998).

Although differences between discourse types have received some attention (e.g., Zwaan, 1993), the laws of genre have received little consideration in empirical research on literature. If they exist with the force attributed to them, we might expect to find traces of their presence everywhere in the data obtained from readers. For example, when the readers studied by Brewer and his colleagues (e.g., Brewer & Lichtenstein, 1981) reported suspense only in the case of stories with appropriately ordered components, this seems to reflect genre expectations: A story that arouses suspense must also satisfy it; otherwise it is considered ill formed. Miall’s (1989) study of responses to a short
story by Virginia Woolf can be seen as the hesitation of readers as they tried first one genre (a romantic story about a relationship) then another (a story about the impossibility of forming genuine relationships). In brief, empirical study, which so far has tended to take genre for granted, stands to benefit from the more detailed specifications of genre knowledge available from literary scholarship.

LITERARINESS

Empirical studies have also replicated, in foreshortened form, a debate central in literary theory until recently: whether literary texts enjoy some distinctive status or literariness. Although the term literariness was first coined by Jakobson in 1921 (Erlich, 1981), conceptions of literature that imply its distinctiveness from other types of text are apparent in discussions from Plato (who wished to banish literature from his Republic) and Aristotle up to Coleridge and beyond. However, recent theorists, in emphasizing the conventional nature of literary response, have dismissed the argument that literary texts are distinctive by virtue of specific features. In Eagleton’s (1983) view, “Anything can be literature” or “can cease to be literature” (p. 10) depending on the doctrine currently in force. More specifically, Fish (1989) objected to the notion that literary texts contain distinctive formal features. Because formal aspects of language cannot guarantee stable meaning, as students of stylistics had tried to claim, there can be no formal aspects of language; these are an illusion.

Similar arguments have been made by prominent scholars in the empirical domain. For example, Van Dijk (1979) proposed that the cognitive processes shown to underlie text comprehension applied to all discourse including literature: “Our cognitive mechanisms will simply not allow us to understand discourse or information in a fundamentally different way” (p. 151); “therefore, we strictly deny the completely ‘specific’ nature of so-called ‘literary interpretation’ ” (p. 151). The differences such as they are are said to lie primarily in the pragmatic and social functions of literature. Similarly, the constructivist approach of Schmidt (1982) led him to suggest that locating the attributes of literariness in the surface features of texts is an “ontological fallacy”: it is “the human processes performed on such features that define the attributes in question” (p. 90).

This controversy has been framed as a contrast between conventionalist and traditionalist approaches (Hanauer, 1996; cf. Zwaan, 1993). The argument is a problematic one, however, because the two positions focus on different aspects of the reading process. Whereas the conventionalist examines reading for the effects of prior cognitive
frames, whether prototypes, genres, or schemata (e.g., Schmidt, 1982; Viehoff, 1995), the traditionalist focuses on specific text features such as meter or personification and attributes changes in readers’ feelings or evaluative responses to this source (e.g., Hunt & Vipond, 1986; Van Peer, 1990). This might suggest that two different systems of response are at issue—one based on cognitive processes, the other on affective processes—and that the latter might be more appropriate for embodying what (if anything) is unique to literary processing. Perhaps the conventionalist has simply been looking in the wrong place. However, the question is not as simple as this might make it seem. Feelings are also subject to conditioning by convention, and readers’ evaluations are clearly bound up with the norms imposed by a specific local culture. Thus, convention may operate here too, although in a less apparent and measurable form.

Nevertheless, some of the features said to be distinctive to literary texts may have been dismissed too readily. For example, although Gerrig (1993) agreed with the conventionalist position in asserting that “The ‘look’ of the language . . . cannot differentiate factual and fictional assertions” (pp. 100–101), he cited a short fictional extract in which a character is described surveying the people in a department store, “studying the crowd of people for signs of bad taste in dress” (p. 99). The limited omniscient narrative mode, of which this is an example, provides access to a character’s mind in a way that is distinctive to fiction (one of its characteristic markers is free indirect discourse). In this respect, the feature cited by Gerrig is unmistakably literary and could only occur in another context in violation of that context’s genre rules (e.g., if a journalist were to impute thoughts to a person in a news article). Another example comes from the field of poetry, where sound effects are often distinctive and measurable. As Bailey (1971) showed, compared with ordinary language, higher frequencies of a particular phoneme group can be shown to occur in poetry; it seems possible that this can influence a reader’s response. Although empirical studies in this area are sparse (e.g., Tsur, 1992), phonetic features like free indirect discourse seem to be characteristic of literary discourse and deserve more careful examination. The principal issue is not their presence as such, but whether it can be shown that readers of literary texts are influenced by them in measurable ways.

THEORETICAL ISSUES

In comparison with the often divisive debates that have occurred in mainstream literary scholarship, empirical researchers have generally been preoccupied with other theoretical issues. Reviews of earlier em-
Empirical work in North America (e.g., Purves & Beach, 1972) show that workers in the field were primarily concerned with improving educational practices. In her review, Klemenz-Belgardt (1981) criticized extant studies for offering poorly theorized views of both literary response and the literary texts under examination. Although the studies helped inform classroom practice, the result was an undertheorization of the nature of literary response (cf. Galda, 1983). Over the last 20 years, however, researchers have drawn on a much wider and better developed set of theoretical contexts, ranging across discourse processing theories, psycholinguistics, social psychology, personality theory, emotion theory, and psychobiology, in addition, of course, to several branches of literary theory. However, a consensus on theoretical issues has yet to emerge.

A new paradigm for research has been claimed by Schmidt and his colleagues known as the Empirical Science of Literature (ESL). ESL undertakes to widen the focus to the actions within the literary system as a whole, including producers, readers, publishers, and critics. In this way, it is argued (Hauptmeier & Viehoff, 1983; Schmidt, 1983), ESL can bring about a Kuhnian paradigm shift, establishing a scientific program of research on literature, one independent of hermeneutic assumptions. Among the theoretical bases of ESL, literary study is based on a theory of human action. It also gives up all ontological commitments—notably the view that there is an essential literariness that distinguishes literary texts. In this perspective, literature is an outcome of linguistic socialization carried out in the interests of social groups; thus, “literariness cannot be regarded as a textual property but as a result of actions of analysis and evaluation performed by subjects within an action system” (Schmidt, 1983, p. 31). Perhaps the best-known aspect of Schmidt’s work has been to define literary response in terms of the aesthetic and polyvalence conventions—a claim that prompted some empirical study (e.g., Meutsch & Schmidt, 1985).

Outside Germany, however, the theoretical claims of ESL have not been generally accepted by other empirical researchers. Viehoff and Andringa (1990) asserted that the ESL model “is nearly without any psychological relevance in literary reading processes” (p. 223). Within Germany, too, the principles of ESL have been disputed: Groeben (Schmidt & Groeben, 1989) argued that Schmidt’s radical constructivism has no way to assess the subjectivity of responses: Without independent descriptions of a text, any response must be accepted. Groeben is concerned that appropriate and inappropriate responses should be distinguishable, implying a normative framework for understanding. As he went on to suggest (Groeben & Schreier, 1998), Schmidt’s polyvalence convention implies a norm. If readers fail to fol-
low the convention, “they are by definition not participating in the liter-
ary system,” which is only one step from saying that literary readers
“must strive for polyvalence” (p. 58).

In fact, normative conceptions of literary response seem inescap-
able, although they are rarely acknowledged in the studies to be dis-
cussed later. This is shown, first, by the choice of literary texts of ac-
knowledged quality in most studies, sometimes in comparison with
manifestly nonliterary texts (such as extracts from newspaper arti-
cles). Studies of response with subliterary genres, such as popular
fiction, have been less common (e.g., Hansson, 1990; Nell, 1988). Sec-
ond, certain qualities of literary response, such as polyvalence, per-
sonal relevance, or a sensitivity to poetic features, imply qualitatively
different and hence desirable dimensions of reading. As Beach and
Hynds (1991) put it in their review of the instructional research, re-
searchers may try to avoid privileging particular styles of reading, but
“an underlying assumption seems to be that developing a sophisti-
cated repertoire of response options to use in a variety of reading situa-
tions should be a major goal of literature instruction” (p. 459). Third,
although Halász (1995) warned, “We may be inclined to evaluate [the]
literary reading process itself as a superior kind of reading.” Investi-
gating the phenomenon for its own sake to understand its processes is but
one step from valuing such processes. This step, easily taken, also
leads to examining the conditions under which literary reading takes
place in education with a view to improving them (e.g., Miall, 1993).

In this respect, the argument over whether literary processes are
driven by socially determined conventions or innate proclivities takes
on a new significance. To put it simply, is the literature teacher trans-
mittting an agreed-on, socially sanctioned technique of reading or facili-
tating an inborn facility for literary experience? Although an interac-
tion of the two processes seems most likely, strong arguments have
been advanced that only the first view can be valid (e.g., Schmidt, Fish).
Because traditional literary scholarship has no way to assess such an
issue outside its own conflicting theoretical perspectives, the empirical
study of literature, with its access to a range of powerful theories and
methodologies, is well placed to play a critical role in this debate.

In the sections that follow, studies of literary discourse and literary
response are placed within several different frameworks. This sug-
gests a problem in theoretical contexts that can be said to exist on two
levels. First, it is suggested that the familiar paradigm of discourse
processes is necessary but insufficient as a basis for understanding lit-
erary questions. The phenomena involved in experiencing and under-
standing literary texts such as novels and poems constitute a larger,
but still little understood, system of psychological processes within
which discourse processes play an essential, but possibly not the most important, role. This is not a new problem: Limitations of discourse theory were suggested by Spiro (1982), who contrasted it with experiential understanding, by Hidi and Baird (1986), who called for the inclusion of motivation in reading studies; and most recently by Zwaan (1999), who outlined the limitations of amodal representations in discourse theory. Empirical literary studies provide a promising field for examining some of these alternative perspectives and working toward a better formulation of their role in discourse.

Second, an important theoretical problem is revealed by a difference inherent in two kinds of empirical method. This is evident, for example, when content analysis methods, such as those of Andringa (1990) and Kuiken and Miall (2001), are compared. Although the first predetermines categories in which think-aloud data are located, the second allows categories to emerge from the data. The second method may allow for those categories (such as certain kinds of feeling or personal memories) that participate in processes distinctive to literary response, whereas its dependence on readers' verbal facility and its closeness to the text being read may make theoretical generalizations about such evidence problematic. The first method may inadvertently filter out precisely that which, from another perspective, makes the responses distinctively literary, yet the second method with its idio-graphic focus risks limiting its conclusions to the specific case being analyzed. The larger problem this suggests is the difficulty of capturing empirically what is distinctive to literary response, given that a specific literary text may call for a mode of response that is more or less unique to that text. Perhaps responses to different literary texts are not entirely incommensurable, but it may be difficult to establish what (if any) literary response has in common that distinguishes it from other experiences of discourse. The range of approaches to be described herein is thus an important feature of the empirical study of literature, helping to maximize the opportunities for triangulating on the significant issues. Like the blind observers in the fable about the elephant, each empirical researcher has a somewhat different conception of the object of study, and certainly none of us has yet seen the whole animal, whereas others again deny that any such animal exists.

**DISCOURSE PROCESSING**

**Discourse Structures**

Although theories of discourse processing in general are not discussed here, several specific studies are mentioned in which literary processing is at issue. Its emphasis on comprehension rather than affect gives
discourse processes both theoretical power and methodological precision, but also limits its scope in capturing literary processes (Miall & Kuiken, 1994a), as Van Dijk’s discussion (cited earlier) suggests. In brief, discourse processing, with its emphasis on comprehension, forestalls attention to those features of literary response that might signal the presence of a different class of response processes. As Spiro (1982) pointed out, referring to a story by James Joyce and the comments made in it by an enigmatic priest, what is central to our experience of the story is that we cannot know what its situation is about. It could be argued that it is just those aspects that resist ordinary comprehension that trigger the response modes specific to literature, such as the emotive, evaluative, and attitudinal. Even this way of putting the issue is, perhaps, misleading if it implies that the alternative modes of response only come into play when the normal cognitive processes have broken down. It may be more plausible to postulate several systems able to operate in parallel as a review of two typical discourse studies suggest.

In one study, Van den Broek, Rohleder, and Narváez (1996) examined the role of causal connections in a story as revealed by a recall measure. Their model of story understanding proposes that at each statement the reader “attempts to establish sufficient explanation for each event or state that he or she encounters” (p. 187). Although 50% of the variance in recall could be explained by the causal connections they had modeled, several other factors also appeared to influence recall, such as statements describing setting, those with a strong impact due to emotion or imagery, or statements relating to the theme of the story. Thus, causal relationships were found to be a major, but not the only, component of story structure influencing recall. Their model of literary comprehension, as the authors suggested, cannot tell the whole story. In fact, although it makes use of a literary text, the findings illuminate that component of response that may be least significant to literary texts.

In another study, Zwaan, Magliano, and Graesser (1995) focused on literary short stories. To examine how far the theoretical components of the situation model were reflected in readers’ responses, the segments of the story (its successive sentences) were analyzed for shifts in causal explanation, shifts in time and space, and propositional complexity. Analysis shows that each of the factors of the situation model contributed to predicting the time taken to read a sentence. For example, there were longer reading times when the story shifted location, requiring readers to construct another spatial setting. Such situation model components appear to be a necessary dimension of response to any narrative. In a subsequent study based on a subset of the reader re-
sponse data of Zwaan et al., Miall and Kuiken (1999) showed that reading times were predicted as strongly by stylistic features in the sentences of a story (e.g., by alliterative patterns, syntactic deviations, or striking metaphors). It might be hypothesized that only literary texts are likely to influence readers through such effects. How response to stylistic features develops, however, remains a difficult issue to examine.

**Expert–Novice Differences**

To the extent that literary response is more fully developed in experienced readers, the difference between novices and experts may help illuminate what is distinctive to literary discourse. Are experts merely more elaborate in their responses or are there qualitative shifts in response as readers gain expertise? Graves and Frederiksen (1991) employed a form of semantic analysis of readers’ comments; Zeitz (1994) measured expertise with comprehension and memory for gist. These can be compared with two other studies by Hanauer (1995a) and Andringa (1996) based on ratings for literariness and degree of involvement.

In their study of think-aloud data from eight students and two experts, Graves and Frederiksen (1991) asked readers to describe a passage from a novel while reading. The protocols were analyzed with “discursive patterns grammar,” a form of semantic analysis whose categories were decided a priori. Comparison showed that experts provided more complex, inferential descriptions, generating more than two and one-half times the number of comments on linguistic structures than the students. Among these, experts’ comments were focused more on syntax, whereas students focused on individual word forms. In general, students tended to remain at the level of paraphrase; experts made more inferential, derived descriptions. Zeitz (1994) distinguished between what she called the basic representation of a literary text (e.g., what happens) and the derived representation (e.g., theme), and she proposed that, although novices and experts may be alike at the basic level, they would differ at the derived level. She found superior performance of skilled readers on gist level but not verbatim recall, better recognition of “multilevel sentences” (i.e., those rich in literary style, such as irony or metaphor), a greater facility in producing interpretive sentences, including references to language or themes, and more complex levels of argument in essay responses.

In the terms of Graves and Frederiksen (1991), the reading processes examined in both studies can be characterized as goal directed: In both cases, readers were instructed to comprehend and describe
the texts. It is not evident that the methods were appropriate for reflecting what, if anything, is distinctive to literary processing. In the first study, it is perhaps significant that up to 10% of readers’ comments were discarded as fitting none of the categories (e.g., quotations from the text, comments on thought processes, and general evaluative statements). In contrast, a think-aloud study by Olson, Mack, and Duffy (1981) found that readers’ comments about a story or their own understanding were infrequent, but diagnostic of those aspects of a story to which a reader was particularly attentive. Although expertise of the kind studied here significantly reflects how literary training enables readers to analyze texts and report on their features, it seems likely to differ significantly from the processes manifested by those reading for pleasure.

In two other studies, measures of appreciation and involvement by Andringa (1996) and measures of literariness by Hanauer (1995a) appear to indicate types of response in which little difference is found between expert and novice readers. Andringa manipulated story frames or narrator comments in two literary stories to vary hypothesized degrees of reader involvement. She reported that, although with greater literary expertise more complex aesthetic structures can be appreciated, emotional involvement seems to change little. Comparing undergraduate and graduate students of literature, Hanauer (1995a, 1995b) presented 12 texts for rating, of which 9 were poems. Experienced readers gave higher ratings overall, suggesting that they were able to recognize more texts as literary, and ratings for poeticity correlated rather highly across groups, whereas the ratings for literariness showed less correspondence. This suggests that training in literature has a greater influence on recognition of literariness than the ability to recognize poetry. Both studies suggest ways to discriminate expertise based on aesthetic and affective factors, but so far this line of inquiry has not been pursued further.

Beyond Discourse Processes

Other studies that go beyond the discourse processes model also suggest that affective factors may have a constructive role to play in literary response. For example, Long and Graesser (1991) found that memory for surface structure is stronger in the context of a literary story than theories of comprehension memory have suggested. In their study, readers of two short stories, who were tested for recognition of two versions of conversational sentences from the stories, gave 68% correct answers—well above chance level. Additional analyses of the sentences showed that memory was facilitated the closer the style of a sen-
tence to oral discourse, or when a sentence was high in expressive evaluation; sentences showing both oral and expressive qualities were best remembered. An affective component consistent with this was proposed by Kintsch (1980), who described cognitive interest as a factor supervening on the processes described in the standard text comprehension model. The reader of a story, in contrast to an expository text, “sets up his own control schema” based on his interests, which may differ from the event-based model of the conventional story schema. Following this suggestion, Zwaan (1993) proposed a “literary control system” (p. 31). Once a literary work such as a novel has been recognized, the control system activates a distinctive form of processing that regulates the basic comprehension processes. Kintsch suggested that “stylistic variations . . . serve as cues for invited inferences” (p. 94), and that these and other semantic surprises in literary texts cause the reader to make inferences about matters not stated in the text. In particular, the appearance of interesting items in the story that fall outside existing schemata invite the reader to activate or construct a schema to account for the story. Alternative processes that construct literary meaning may thus operate in conjunction with the standard comprehension processes (Miall & Kuiken, 1994a).

**ALTERNATIVE FRAMEWORKS OF LITERARY READING**

**The Literary Perspective**

The reader’s knowledge of a text’s genre, as noted earlier, is a control condition influencing the reading process. Readers’ responses are shaped quite differently according to whether a text is a poem, newspaper story, or joke. Yet the limits of genre as an explanatory factor are suggested by three studies in which genre was manipulated experimentally.

Zwaan (1991) compared the effects of placing reading in either a newspaper or literary perspective. With the help of a pilot study, six texts were chosen from newspapers and novels that could be read as either a newspaper or literary prose (texts were thus devoid of marked literary features). As expected, it was found that in the newspaper condition, readers read faster—on average about 12% faster than readers in the literary condition. In a recognition test, the literary condition readers were more accurate in identifying words that had occurred in the text just read. In a second study using the same paradigm (Zwaan, 1994), readers in the literary condition were found to have a poorer memory for the situational information in the texts than the newspaper condition readers. The study suggests that when readers think
they are engaged in literary reading, they read more slowly and form a better representation of the surface structure of the text.

Zwaan's study argues for the regulative functions of genre. Yet genre information is not invariably decisive. As Hoffstaedter (1987) showed, it cannot always take precedence over textual features. She presented 24 poems to readers in two conditions: a newspaper reading condition (modifying the layout of the poems appropriately) and a poetry reading condition. Readers were asked to make judgments of poeticity on a scale from nonpoetic to poetic. She reported that for only 10 out of 24 poems were judgments significantly different across the two conditions. For 14 of the poems, it was their poetic properties that appear to have determined the reading condition. This study shows how the text features encountered here, such as prominent metaphors, personifications, or unusual syntax, can override the supposed genre in a bottom-up fashion.

The reading framework manipulation is called into question by László (1988): A reader may be told a text is literary but not read it as such. Devising a different method, László implemented changes in the texture of an American and a Hungarian literary story to examine effects on both reading times and readers' deployment of schema knowledge. Three different versions of each story were presented: the Original version, the Insert version (where two or three key passages based on an action sequence had been rewritten in summary form), and a Script version (where the whole story was rewritten in summary form, keeping the plot structure clear). Readers of the Original and Insert versions read the stories on paper except for the critical passages (original or rewritten), which they read on a computer screen to collect reading times, after which they were asked to choose among three alternative plot continuations; they then continued reading on paper. Script readers read the whole story on computer. Both stories were about short-term romances that turned out to be unstable because of social differences, thus suggesting two themes—social and romance. László suggested that readers' expectations regarding the romance theme might stem from scriptlike cognitive structures (cf. Miall, 1989).

László reported that reading times were longest for the Original story, but shorter for the Script than for the Insert version in the case of the American story, whereas reading times for the Script and Insert versions were the same for the Hungarian. Although these results are equivocal in part, they show that the literary texture of the original stories prompted the longest reading times. However, the manipulation of the stories had no consistent influence on the plot-continuation choices—a finding that seems to show construal of the romance plot to
be a feature of the discourse level, not the literary level of processing. In this respect, the longer reading times found in both Zwaan’s and László’s studies can be attributed to the distinctive components of literary processing, at least part of which springs from closer attention to the surface features of texts. One may wonder what processes are occurring in literary readers during that extra time. This question is addressed next.

**Polyvalence**

One answer is suggested by Meutsch and Schmidt (1985). In their conception, literary reading centers on problem-based understanding. Polyvalent constructions are said to reflect the reader’s response to reading problems. In their study, readers’ think-aloud data were collected in response to either a poem or short literary narrative and analyzed in terms of four classes of response: (a) descriptions of, or (b) comments on the reading process, (c) references to literary conventions, and (d) anticipatory or retrospective comments on text meaning. Among other findings, Meutsch and Schmidt reported that a mean of 2.7 frame of reference changes was found per reader (more in the case of the poem than the narrative), and frame of reference changes were in general evaluated positively more strongly than they were negatively.

The conception of literary reading as problem based is questionable, however, especially the assumption of Meutsch and Schmidt that think-aloud responses are only elicited by problems during reading. Readers may be prompted to comment by interest, surprise, pleasure, and other immediate impulses while reading. In a number of instances, however, the presence of more than one meaning or a transition from one meaning to another is signaled by such comments. Thus, polyvalence appears to be a characteristic of literary reading, although not necessarily in a form that requires readers to entertain several meanings at the same time. However, the study only shows that readers deploy polyvalence during reading, not whether the polyvalence convention is a governing condition of literary reading, as Schmidt and his colleagues have proposed.

Other examples of empirical findings that tend to support the polyvalence conception are provided by Graesser, Kassler, Kreuz, and McLain-Allen (1998) and Miall and Kuiken (2001), where evidence for changes in meaning was obtained, although under rather different conditions. Graesser et al. (1998) studied changes from the normal schema for time prompted by reading Alan Lightman’s (1993) novel, *Einstein’s Dreams*, where each chapter offers a different deformation of time (e.g., time running backward or repeating each day). The find-
tings show that readers were able to develop conceptions of time inconsistent with their normal conception, not only polyvalently developing conceptions of time across several chapters, but also being able to assess deviation in the time schema in each chapter in relation to the normal model of time.

Miall and Kuiken (1995) asked readers of a literary short story to register shifts in story understanding explicitly while reading using a modified version of the Remindings paradigm (the paradigm is discussed in more detail later). This instructed readers to mark those segments of the story that they found striking or evocative and then describe what memories or shifts in story understanding might have occurred. The story had also been analyzed for stylistic effects. Miall and Kuiken found that the mean number of story shifts per segment reached a peak systematically at around 12 segments following a marked cluster of stylistic features. This sequence of events appeared to occur three times. This finding suggests that readers are likely to experience stylistically prompted changes in understanding several times during the course of reading depending on the style and structure of the story in question. Theoretical conceptions of shifts in understanding have also been proposed by Harker (1996) and Cook (1994). Harker portrayed literary meaning as emergent, appearing in repeated cycles of familiarization and disruption that require the reader to engage in what Harker called reattentional activity. Cook argued that literary texts are “a type of text which may perform the important function of breaking down existing schemata, reorganizing them, and building up new ones” (p. 10). The studies of Graesser et al. (1998) and Miall and Kuiken (2001) not only provide empirical support for this process, but also indicate some of the initiating causes for schema change.

Anticipation

The emergent nature of literary meaning suggests that readers’ understandings during reading may be provisional. At the point when a shift in meaning occurs or a reader polyvalently entertains two or more possible meanings, such meanings may be held somewhat like hypotheses—provisional interpretations to be tested against incoming evidence while reading continues. The polyvalence of literature may thus predispose readers to experience a greater degree of anticipation during reading. This phenomenon was demonstrated in two studies by Olson, Mack, and Duffy (1981) and Langer (1990).

In Langer’s (1990) study, think-aloud data were obtained from 7th- and 11th-grade students in response to literary and expository texts.
She reported that literary reading was characterized by a forward looking and shifting context of understanding, whereas expository reading was characterized by successive and cumulative relationships to a fixed point of reference established early in reading. A similar finding was reported by Olson et al. (1981), who also studied think-aloud responses elicited by two texts, expository and narrative. In comparison with readers of the expository text, the story readers produced a number of anticipatory comments, whereas the expository text readers generated almost none. Olson et al. also collected reading time data from other readers and were able to show through multiple regression analysis that at the points where reading was longer per syllable the think-aloud readers were generating more inferences and predictions. Along with polyvalence, anticipation may thus be one of the distinctive markers of literary response. A neuropsychological model of anticipation in reading, based on feeling, has been developed by Miall (1995).

Rereading Effects

Another method for examining emergent effects during literary reading is rereading. Measures of reading, such as evaluative judgments, ratings of feeling, or interpretive statements, are taken during two or more readings and then compared. For example, Dixon, Bortolussi, Twilley, and Leung (1993) argued that the literariness of reading is more likely to appear during a second reading and developed a simple measure of literary appreciation to evaluate it. Using a literary story by Borges and a subliterary detective story, they found a marked upward shift in appreciation following the second reading of Borges, but not the detective story. In a second study, postulating that the literariness of Borges was due in particular to passages that suggest the unreliability of communication, they prepared a second version of the story in which these passages were regularized and made unproblematic in relation to the issue of communication. The same test for appreciation showed that little or no shift occurred when the manipulated version of the Borges story was read a second time.

In a more complex experimental design, Cupchik, Leonard, Axelrad, and Kalin (1998) were also able to show systematic changes during a second reading. Cupchik et al. prepared passages from four stories by James Joyce, chosen to be either emotional (showing character’s actions and emotional responses) or descriptive (stylistically complex descriptions of settings and characters). After reading each passage, readers gave ratings on several text-focused judgments and several reader response qualities; they then either generated an interpretation or received one, after which they read the passages again and rerated
them. Although readers’ evaluative ratings showed that the two emotional passages were preferred and were read more quickly than the descriptive passages, this effect somewhat diminished during the second reading. The ratings for complexity and expressivity of the emotion episodes diminished at the second reading. Overall, there was some evidence that the descriptive passages, at first seen as less complex and interesting, increased in significance for readers at a second reading. The authors suggested that “the language which has passed by takes on new potency, finding a meaningful role that goes beyond mere description to serve as rich allusion” (pp. 843–844). In other words, readers tend to move beyond a story-based understanding toward one focused on stylistic and evaluative components.

In contrast, Halász, Carlsson, and Marton (1991) found no reliable differences in a rereading study. Halász et al. reported that little improvement in recall occurred across four trials; in addition, readers appear not to have changed their understanding of the stories during the rereadings. They suggested that readers form a frame based on their initial impressions, which then determines subsequent readings. Unlike the first two rereading studies, however, in which affective and evaluative measures were employed, the recall task in this study may have served to concretize readers’ understandings early in the process. The recall question, it should be noted, was, “Could you tell us what the text is about?” It seems possible that directing readers’ efforts to a verbal report on comprehension forestalled attention to the literary qualities of the texts.

**Literary Meaning**

Ensuring that readers in an empirical study are engaged in a literary reading may not be entirely straightforward. In the first of a sequence of reports on studies of reading, Vipond and Hunt (1984) outlined three different strategies for reading: information driven, story driven, and point driven. They proposed that the reading of literary narratives is best characterized as point driven—a process in which the reader considers what the narrator is getting at. Among other components, they suggested that point-driven reading involves a sense of an author seeking to make a point. At the same time, construction of a point is culturally relative and in part based on expectations derived from genre—that is, what kind of story is being told. Yet they found it surprisingly difficult to locate point-driven reading at least among the student readers they studied. A questionnaire survey of over 150 readers of Updike’s short story “A & P” showed that only 5% were engaged in a point-driven form of reading; most students found the story incom-
plete and without a point. Only after adopting different methods, including a nonexperimental method based on classroom discussion, did point-driven reading clearly emerge (Hunt & Vipond, 1991).

Reading for point may be an effective strategy for approaching literature, but it is not evident that literary meaning is necessarily to be captured by points. Vipond and Hunt (1984) hypothesized that the components of point-driven reading included (a) **coherence**, the attempt to connect apparently unrelated or unnecessary parts of a story; (b) greater attention to **surface** features, such as syntax or style, on the assumption that these features are motivated and contribute to meaning; and (c) the **transactional** stance, an awareness of an author in control of a narrative who has a point to make. Each of these strategies could be tested empirically, as they suggest, yet responsiveness to any of these components does not oblige the construction of an overall point. Authors may be understood to have succeeded in producing a compelling and thought-provoking narrative, such as Coleridge with “The Rime of the Ancient Mariner” or Beckett in his play *Waiting for Godot*, without readers being in a position to know what the point of either of these texts may be. Readers may engage with texts in other ways yet still be reading in a literary mode: Some of these reading modes are examined in the next section.

**PERSONAL READINGS AND FEELING**

**Implicating the Self**

Literary reading may in part be distinctive for interacting with a reader’s self-concept and personal goals. For example, Klinger (1978) reported a series of studies in which he found readers’ current concerns (their personal and unconsummated goals at the time of reading) were reliably evoked by a literary text. In one study, readers heard excerpts from two texts played simultaneously in the reader’s ears, where one of the texts was modified at specific locations to embody words and phrases known to be relevant to the individual’s concerns. When readers were interrupted with a tone during one of the locations and asked to report their thoughts, readers responded with concern-related thoughts in relation to the modified passages about twice as often as to the unrelated passages. What is not evident from this study is the nature of the relationship between the process of reading and the reader’s concerns. Although we might suppose that literary reading assists readers to conceptualize and evaluate their concerns, gaining evidence of this within the confines of an experimental situation is a difficult task. The systematic influence of literary texts in calling up
personal meanings, however, has been suggested by findings within several experimental conditions.

Halász (1996) examined the frequency with which personal meanings were invoked by a literary compared with a nonliterary text. After reading each of three short sections, participants were asked to generate its accepted meaning and personal meaning. In counting the frequency of accepted and personal meaning units, Halász found that, although the expository text produced three times as many general to personal meaning units, the literary text produced almost the same number in both categories, showing that the literary text enabled readers to generate a markedly higher proportion of personal meanings. Among the personal meanings, the predominant types were actions, feelings, evaluations, and cognitive qualities (images, daydreams, intuitions, etc.).

Several studies that have examined the role of personal meaning have made use of the Remindings method or "self-probed retrospection." Developed by Larsen and Seilman (1988) as a less disruptive procedure than think-aloud method, readers are asked to note with a marginal mark when a reminding occurs (i.e., when they think of something they have experienced). After reading, readers are asked to describe what they were reminded of at each marked passage. In a study by Sielman and Larsen (1989), the authors proposed that when comparing responses to a literary and an expository text, the literary text would involve more memories of the reader as an actor than as an observer. Using two texts—a short story and a text about population growth (each of about 3,000 words)—they found that, although a similar quantity of remindings was elicited by both texts, twice as many actor-perspective remindings were elicited by the literary text, whereas the expository text elicited more receiver remindings (memories of things read or heard about). Thus, they suggested, literary reading "seems to connect particularly with knowledge that is personal in the sense that one is an agent, a responsible subject interacting with one's environment" (p. 174). Remindings were also found to occur more frequently in the opening section of each text, but more markedly so in the case of the literary text. This suggests that readers call on specific, personal information to contextualize the world of the text.

Using the remindings method with literary and nonliterary texts, Halász (1991) also found actor role memories were more frequent in response to the literary text than either observer memories or memories of events heard about. Halász suggested that the inappropriate-ness of readers' existing schemata for a literary text impels the reconstrual of readers' knowledge to overcome the obstruction: Personal remindings may be a source for this reconstrual, and in turn
readers’ self-knowledge may be enriched as a result. A preliminary study reported by Miall and Kuiken (1999) pointed to this possibility. Readers in a remindings study responded to Coleridge’s “The Rime of the Ancient Mariner.” Think-aloud material from one reader is analyzed. Her comments show an evolving pattern of existential concerns about death, in which the world of the Mariner in the poem and that of the reader appear to merge in the final set of comments. This convergence of the situation of the (fictional) protagonist with the concerns of the reader suggests how the encounter with a literary text can significantly alter the reader’s self-understanding.

Another way to conceptualize variance in reading due to personal meaning is to consider readers’ proximity to the setting and themes of a particular literary text. This may be mediated by formal features in the text such as free indirect discourse, where a third-person narrator indirectly represents a character’s thoughts and feelings, bringing the reader close to the character. Dixon and Bortolussi (1996) found that by manipulating free indirect discourse they were able to bias readers toward favoring one or the other of two characters in a story. The degree of personal meaning may depend on the extent of the reader’s familiarity with the culture shown in the text. In several studies by László and his colleagues (e.g., Larsen, László, & Seilman, 1991), culture was the variable manipulated. It was found that readers presented with a story from their own culture generated more personal experiences. When point of view was manipulated, however (László & Larsen, 1991), inside point of view elicited more personal experiences (sensations, affects, and images) regardless of culture of reader, suggesting that point of view has universal effects on readers.

Another form of proximity was proposed by Dixon and Bortolussi (2001). They suggested that readers of a literary text are likely to process it as if it were a communication of the narrator; thus, readers form a model of the narrator and his or her stance toward the narrative. They pointed out that some aspects of the narrator’s stance are computed automatically because they are with a partner in a conversation. Other aspects, such as the narrator’s view of the characters or the theme of the story, may require more deliberate processing. The reader’s relation to the narrator may thus form an important component of literary reading.

Two Types of Feeling

The feelings mentioned in the previous section have been invoked by fictional representations. The feelings of readers resonate in various ways in response to characters and their settings, prompted by free in-
direct discourse or by reminders that arouse personal action memories. Yet as Kneepkens and Zwaan (1994) pointed out, building on a suggestion by Frijda (1986), readers also experience feelings about the form of the texts they read, such as appreciation for stylistic or structural features. Thus, they referred to two types of feelings: feelings aroused by fictional events (i.e., narrative emotions) and feelings in response to the artifact (i.e., aesthetic emotions).

One class of narrative feelings can be related to the textbase linked to characters and events. Thus, readers of a narrative may experience suspense or curiosity (Brewer & Lichtenstein, 1982; Brewer & Ohtsuka, 1988), feelings of empathy for a character (Bourg, 1996), or feelings aroused by major thematic concerns such as death, danger, power, or sex. However, narrative feelings can be divided into two subclasses: those directed to others in the story (altercentric) and those directed at the self (egocentric). For example, I may reexperience a feeling on behalf of a character or the events in a story may cause me to experience a feeling about myself.

Kneepkens and Zwaan (1994) also postulated a phasic model of response. They suggested that egocentric feelings are called into play when a particular literary passage seems unusual, abstract, or vague. As the story becomes clearer, however, the egocentric feelings fall away until a new episode begins. Although this proposal has not been tested directly, Miall (1988) and Miall and Kuiken (2001) reported an initial positive correlation between reading times and affect ratings at the beginning of story episodes, where the coincidence of longer reading times and higher affect suggested that readers’ feelings were implicated in helping contextualize the new episode. In the story studied by Miall and Kuiken (2001), a cluster of striking stylistic features appeared to signal the onset of a new episode (an example of the role played by aesthetic feeling).

This last finding questions the suggestion of Kneepkens and Zwaan (1994) that the prevailing cognitive model of text understanding provides an appropriate framework for situating effects due to feeling (cf. Miall, 1995). This approach is taken in a study reported by Dijkstra, Zwaan, Graesser, and Magliano (1994). When reading times were collected from readers of short stories, it was found that text difficulty was a significant variable in the case of narrative emotions. For passages portraying character emotions, the more difficult stories elicited slower reading times; for less difficult stories, reading times for such passages were faster. This may not only have been because the characters’ emotions were more difficult to understand as the authors suggested. It may have been that emotion felt by the reader was being used as a prompt for emergent story interpretations based on analogies in
the reader's experience. This possibility is examined more closely in
the next section.

**Feeling-Based Understanding**

An initiating role for feeling in readers' developing understanding is
shown by a study of Andringa (1990), who collected think-aloud proto-
cols. Her primary intention was to develop a systematic method for
classifying think-aloud comments. Thus, two levels of analysis were
elaborated, consisting of a range of speech and reception acts. The
method was applied to analyzing readers' comments on a Schiller
short story. The most notable finding was that for the less experienced
readers the most common sequence consisted of emotion references,
followed by evaluation, then argument. Andringa commented that in
most of the protocols this "seems to be a regular sequence," suggesting
that emotion "initiates, selects, and steers the way of arguing." An ex-
ample of such a sequence is provided by Andringa (slightly modified
here): "[Emotion] Oh (smiles, laughs), [Evaluation] yeah, that's a little
bit theatrical . . . [Argument] I can only say she was very stupid, that
woman, because . . ." (p. 247). Among her more skilled readers, the
more regular sequence was a reference to the text followed by a
metacomment elaborating the meaning of the reference, then interpre-
tive comments.

In a study of responses to a Virginia Woolf short story, Miall (1989)
also found evidence for the constructive role of feeling. Readers rated
segments from the opening section of the story either for affective in-
tensity or importance. The first group then attempted to recall the
phrases, and the second group read the whole story and then rated
the opening segments of the story again for importance. Both groups
provided written comments on the opening section before rating it.
However, the segments in the story were previously classified as fall-
ing into one of two types: phrases describing two characters who may
be about to form a romantic relationship or phrases that described
the setting (emphasizing the sky and the night). The first group re-
called significantly more phrases from the relationship group, al-
though high affect ratings had been given to both types of phrase. Yet
after reading the whole story, the second group shifted their impor-
tance ratings for the opening section away from the relationship
phrases toward the sky and setting phrases. Because the promised
relationship later turns out to be an illusion, it seems probable that
readers turned to the setting phrases to reconstrue their understand-
ing of the story. The high affect ratings given to these phrases appears
to anticipate the significance they accrue later in the reading. As one
reader put it, after reading only the first section, “They will probably continue to talk and either drift apart or recognize some kind of connection and become closer. Sky and moon will somehow influence how they relate to one another” (p. 72).

In this study, readers’ responses to the relationship schema can be seen as the application of familiar feelings. The feeling for setting, however, appears to have prompted new feelings for most of the readers—feelings that became more appropriate as a context for understanding the story. A similar contrast is made in a study by Cupchik, Oatley, and Vorderer (1998), who compared emotion memories and what they termed fresh emotions in a study of responses to short stories by Joyce. Passages in the stories were identified as either descriptive or focused on characters’ emotions. After reading each of four segments from a story, classified as either emotional or descriptive, readers answered questions about the emotions they experienced and whether they were fresh or remembered. Results show that in general fresh emotions were elicited more often than emotional memories and were less pleasant, whereas emotional memories were more powerful. As expected, descriptive passages tended to evoke fewer emotion memories than fresh emotions. Over the four segments of a given story, however, emotion memories were more frequent early in the story, whereas fresh emotions became more frequent later. Similar to the findings of Miall (1989), this seems to imply a schema-setting role for emotion memories, but an interpretive role for fresh emotions.

LITERARY COMPONENTS

Imagery

Denis (1984) pointed to several effects of imagery in a review of the role of imagery in prose. In general images appear to facilitate recall of texts, although individuals with a tendency to form imagery spontaneously showed better recall than those without. It has also been found that high-imagery subjects have longer reading times than low-imagery subjects. Denis suggested that imagery enhances recall because it facilitates encoding of information in a structured form where items interact rather than remain separate. In normal reading, however, it is questionable whether readers form detailed imagery except under special circumstances. Spatial imagery was studied by Zwaan and van Oostendorp (1993) and was found to be quite poor when readers were asked to report the position of items in a setting.

Yet the occurrence of imagery in response to narrative appears to be consistent and correlate significantly with other features of response.
Goetz and Sadoski (1996), for example, asked readers to rate the strength of imagery at the different sections of a literary story. They found that readers showed a surprisingly high degree of agreement, with alpha reliability coefficients of above .90; imagery ratings also correlated consistently with ratings for affect. As the authors pointed out, their studies suggest that “visual imagery and affective, or emotional imagery” play a key role in the reading experience. Visual imagery may thus provide one, concrete matrix for registering emotional responses during reading. Other image modalities that may function in this way, such as olfactory or kinaesthetic, remain to be studied.

**Foregrounderg**

As mentioned earlier, stylistic features have been considered a hallmark of literary language until recently. Evidence for their influence on reading is available in several studies, suggesting that contemporary critical disregard for this aspect of literariness may be premature.

Hunt and Vipond (1986) examined the effect of local features of a narrative that they termed *discourse evaluations*. Occurring at the sentence level, these are the means that narrators use to convey beliefs, values, and attitudes. It was proposed that readers notice evaluations because they stand out from the local norm of the text. Evaluations are of three kinds: discourse (unusual style), story (unusual plot elements), and telling (unusual comments by a narrator). Evaluations are held to signal the point of a narration, here taken to be the theme or meaning. Thus, evaluations are “deliberate invitations to share a meaning with the storyteller” (p. 58). Two versions of a story were prepared, in which evaluations in one version were replaced by semantically equivalent neutral statements (e.g., “they camped around the room” was replaced by “they sat around the room”). Asked to note what they found striking in the text, readers of the original version, as expected, noticed discourse evaluations more frequently than the equivalent normalized passages in the revised version.

Comparable studies by Van Peer (1986) and Miall and Kuiken (1994b) focused more closely on linguistic and phonetic features of style, termed *foregrounding* after Mukačovský (1964/1932). Van Peer (1986) examined phonetic, grammatical, and semantic features in six short poems and ranked the lines of each poem for the presence of foregrounding. For example, this opening line from a poem by Roethke, ranked high: “I have known the inexorable sadness of pencils.” The pencils are personified; it contains an unusual word, *inexorable*; it contains repeated phonemes such as /n/ and /e/. Foregrounded features can be classed as deviations from normal language
use (e.g., a metaphor) or they constitute an unusual parallelism (e.g., use of rhyme or a repeated stress pattern). Among other tasks, readers were asked to rate the lines for strikingness. Van Peer found that the mean ratings for strikingness were strongly predicted by the presence of foregrounding. This effect was obtained whether experienced or novice readers were involved.

Miall and Kuiken (1994b) carried out a similar study with three modernist short stories. They scored each segment of the stories (roughly one sentence) for foregrounding. Readers read the stories on computer while reading times per segment were recorded. In a second reading, they rated the segments on one of four judgments: feeling, strikingness, importance, or discussion value. For each of the stories, foregrounding was found to strongly predict reading times and ratings for feeling and strikingness. This effect was also found whether readers were experienced, senior students of literature, or relatively unskilled readers. Ratings for discussion value and importance, in contrast, were not reliably associated with foregrounding possibly because segments high in foregrounding did not always coincide with the most important sections for narrative events. Miall and Kuiken (1994a) argued that structures of foregrounding, in contrast to the semantic and narrative features usually at the center of discourse processing studies, offer readers an alternative, feeling-based mode of response. Temporally speaking, as readers encounter foregrounding, they first find the passage striking. It defamiliarizes customary or accepted meanings, arousing feeling in the process. This then leads readers to engage in a search, led by feeling, for a context in which to locate the unusual meanings suggested by foregrounding. This phasic sequence of response has as its outcome the registering of a shift in understanding downstream from the moment of foregrounding (Miall & Kuiken, 2001).

Evidence for the moment of defamiliarization is provided by Hoorn (1996), who studied the electrophysiological response of readers. Hoorn proposed that event-related potentials (ERPs) would mirror the response to deviations in semantic or phonological expectations in the last word of a four-line verse. In his study, the final word was made either consistent or inconsistent with the semantic content of the verse, or an expected rhyme word occurred or did not occur, or both inconsistencies occurred together. Hoorn found reliable differences in ERPs: Phonetic deviation produced a response with significant negative shifts, N200, N400, and N700; semantic deviation produced a N400 shift (i.e., a shift at 400 msecs following the appearance of the anomalous word). Although the experimental feature studied here violates poetic form (at the level of sound or meaning), it can be considered a test of stylistic deviation (Van Peer, 1986).
Phonetic Variation

The sound of poetry is perhaps one of the most distinctive features of literariness, offering unusual rhythmic features or striking clusters of phonemes such as alliteration or assonance. Van Peer (1990) argued that literary language does more than refer to a state of affairs; it also creates a sense of significance beyond the ordinary meanings of the sentences. This forms a part of that extra level of meaning that has usually been considered aesthetic. Meter is one such device, which appears to have a double function—aesthetic and mnemonic. It enables a text to be better remembered, and it helps create the significance of the text.

In his study, a modern humorous Dutch poem with a marked metrical structure was varied by creating a second version where the metrical features were largely removed while keeping other literary features intact as far as possible (e.g., alliteration, rhyme). Readers read either the original or altered version silently and then completed two tasks. First, they made judgments on a set of 16 semantic differential ratings designed to elicit aesthetic judgments. Second, they completed a questionnaire that tested recognition and recall. The overall ratings on the semantic differential showed a significant aesthetic advantage for the original poem; readers of the original poem were also more accurate in identifying lines from the poem. However, both groups of readers performed at the same level in recalling content from the poem. The results “show the form of literary texts to carry specific informational possibilities in its own right” (p. 270), thus responding to a literary text is not confined to understanding meaning.

Another form of phonetic variation was examined by Miall (2001). It has been a common intuition that the sound of language supports its meaning (e.g., a narrow, front vowel such as /i/ is used to signify something small or high). Miall found that rank orderings of vowels and phonemes according to their position of pronunciation in the oral tract (e.g., from high to low or front to back) were related to several features in nonliterary texts, such as differences between groups of male and female names. They also reliably distinguished different parts of literary texts, such as Milton’s descriptions of Hell and Eden in Paradise Lost. In a study of readers of a short story, phonemic contrasts were found to contribute to variations in reading speed and readers’ ratings of story segments, suggesting that readers were sensitive to variations in tonal patterns while reading the story.

Another approach was introduced by Bailey (1971). He developed a prominence index for phonemes, where frequency in a given poem was used to place phonemes in rank order; these ranks were com-
pared with the ranks in a sample of standard English. He argued that higher frequency of a given phoneme was more likely to be noticed by a reader especially when a phoneme was relatively rare in standard English. This test revealed several systematic effects in the poems selected for study. In Dylan Thomas’ poem “Fern Hill,” for example, he found a prevalence of voiced over unvoiced phonemes. Although this approach has not been verified with studies of readers, a sensitivity to phonetic effects has been shown in a nonliterary context by Zajonc, Murphy, and McIntosh (1993). Other accounts of phonetic patterns in poetry are offered by Tsur (1992). These results challenge the claim of Meutsch (1989) that “we can state the irrelevance of textual qualities for the management of problems during the cognitive process of comprehension” (p. 69). It seems probable, on the contrary, that literary readers are influenced by the sound structures of a text in their effort after meaning.

PROSPECTS

The components of a future theory of literary processing may emerge from the studies discussed herein. In most of the studies, the focus is confined to one or two features of the response process, thus the findings cannot be related to a larger theory of processing except in a preliminary way. Studies that triangulate on a postulated phenomenon from a number of positions are clearly required, given the complexity and considerable variance apparent in all the dimensions of the literary process—the range of texts of different genres, the wide variety of responses exhibited by readers (although focusing on differences in interpretation may not be a priority), and the range of response processes from inferences to feeling. Magliano and Graesser (1991) are undoubtedly correct in advocating a three-pronged approach to literary understanding. This requires us to (a) make predictions based on medium-level theories about literary response, (b) analyze think-aloud data from readers, and (c) use behavioral measures such as reading times. The studies reviewed here, however, suggest that the addition of several more prongs might be required before an appropriately complex and powerful methodology is available—prongs that would be applied and coordinated within a single research program.

As the discussion has already shown, methodological questions are decided within the perspective of the researcher. Broadly speaking, literary processing has often been approached as a branch of discourse processing, with preset categories for analyzing response. However, this may deselect those features most characteristic of literary reading. A specific example of this problem lies in the type of instruction often
given to readers. For example, in the study of Graves and Frederiksen (1991), readers were asked “to provide a verbal description of the passage while reading it, commenting on its content and style,” an instruction that appears directed to comprehending rather than experiencing the text. Similarly, the readers studied by Olson, Mack, and Duffy (1981) were told before reading that “later we would explore how well they understood each story.” In an informal study of my own (Miall, 1986), I suggested that when asked to write freely about their responses to a poem, the first responses of readers are predominantly affective. Only later, as readers come to define their experience of the poem, do more conceptual, analytic comments appear. Finding appropriate ways to elicit information about their response processes from literary readers, devising methods sensitive to the complex and fugitive aspects of the response, will continue to present a major challenge to future researchers.

Several key questions about the reading of literature remain. First, although the limitations of the cognitivist basis of research on reading has been pointed out (e.g., Miall & Kuiken, 1994a), a number of effective studies situated in alternative frameworks are now extant, ranging from studies of personal meaning (Halász, 1996) to electrophysiological measurement (Hoorn, 1996). Thus, it is evident that a rapprochement of cognitive and alternative approaches that examine affective, self-referential, and cultural issues must be sought. This may help distinguish the boundary (no doubt a blurred one) between literary and nonliterary processes. For example, do the situation model processes demonstrated in Zwaan, Magliano, and Graesser (1995) operate in parallel to foregrounding processes (Miall & Kuiken, 1999)? Do features distinctive to literary processing displace cognitive forms of processing in some instances? Van Dijk’s (1979) assertion, “we strictly deny the completely ‘specific’ nature of so-called ‘literary interpretation’,” seems open to challenge.

Second, the problem of understanding what may be innate in literary experience calls for careful and innovative kinds of study. Although it is clear that in major respects literary reading is premised on the cultural and educational context in which a reader learns to experience literary texts, it remains to be examined how far literariness is based on features of response that are a part of the evolutionary acquirement of human beings. Given that artistic production in the visual mode is known to date from at least 30,000 years before the present, it seems probable that literary experience (first oral, then in written form) has an equally long history that may be embedded in the human genotype.

Finally, although literary reading has a prestigious past, its future has been called into question given the advent and rapid spread of digi-
tal media. Advocates for electronic forms of literature, including proponents of innovative hypertext literature, have cast doubt on the validity of literature in the printed book, yet it is far from clear whether the repurposing of literature for the electronic medium can continue to offer the same experience (Miall, 1998). Thus, it becomes more urgent to understand what literary reading is—what role it plays in the ecology of human culture and the health of individuals—before it is reconfigured or disappears in the face of new forms of electronic literacy.

REFERENCES


course: Results and problems of interdisciplinary approaches (pp. 16–46). Berlin & New York: Walter de Gruyter.


Why do people sometimes mean something other than what they say? People often voice frustration when a speaker’s words fail to express what they appear to mean. Nonliteral speech acts seem to obscure people’s intentions as much as they clarify them. In some instances, listeners feel that indirect and figurative language is purposefully employed to evade, equivocate, and deceive.

There are cases, however, where speaking indirectly or figuratively appears to have a close link to both what people think and what they really mean to communicate. Consider the following brief exchange that occurred on CNN TV the day after the New Hampshire presidential primary. The discussion focused on what Bill Bradley would do to enhance his chances of winning later primaries just after he lost to Al Gore in New Hampshire.

**Judy Woodruff:** Bob Novack, how much mileage can Bill Bradley get out of raising questions again and again about how much people can trust Al Gore?

**Bob Novack:** It’s his only hope, Judy. And it worked very well in New Hampshire. As Jeff said, of course, you have a situation in New Hampshire where independents can vote. One of Bradley’s advisors described the situation in fight terms: that for the first seven rounds of the fight, Bill Bradley
played rope-a-dope. He laid against the ropes and got hit, and the vice president kept building up an edge. Then suddenly Bradley in the last week took two little left hooks, and two little left jabs, and suddenly there’s blood all over the place. Is Al Gore a bleeder? Is he the kind of person, if you really hit him on—much harder than Senator Bradley has so far—on the Buddhist temple, on the “no controlling legal,” on his imagining himself and his exaggerations, does he begin to bleed all over the place?

Novack is clearly employing different nonliteral expressions to convey the severity of the situation confronting Bill Bradley after losing to Al Gore in New Hampshire. Yet his metaphorical rhetoric is serious too in capturing something of how he, and the Bradley campaign, thinks about Bradley’s political situation. Most listeners and readers experience little difficulty understanding nonliteral speech acts, especially when these are used in appropriate social contexts.

Why do people speak indirectly and figuratively? What psychological processes underlie our understanding of such language? This chapter addresses these issues by describing some of the pragmatic and conceptual complexities associated with nonliteral speech acts. My central argument is that nonliteral speech acts are pervasive in all types of discourse, easy to comprehend, and in many cases directly reflect the ways people conceptualize their experiences and the world around them. Much of the recent empirical and theoretical work on nonliteral speech acts highlights how people’s easy facility in speaking and understanding nonliteral language is rooted in basic psychological processes that are central to many other aspects of human cognition. My discussion focuses on several recent controversies in the interdisciplinary study of nonliteral discourse and raises different methodological and theoretical challenges for future research.

**THE VARIETY OF NONLITERAL SPEECH ACTS**

There is no single, principled way to divide literal from nonliteral speech (Gibbs, 1984, 1994; Gibbs, Buchalter, Moise, & Farrar, 1993). People can sometimes judge a particular utterance or text as having literal or nonliteral meaning. Yet these meaning judgments are notoriously variable and do not necessarily reflect the outcome of different cognitive processes. Yet there are clearly a variety of linguistic strategies that speakers use to communicate in an indirect or a nonliteral
manner. A list of some of these is presented next (Brown & Levinson, 1987; Gibbs, 1999a):

1. **Hints:** Sally says to Peter *It's cold in here* to get him to close the window, or Jane says to Paul *The soup's a bit bland* to get him to pass the salt. Both utterances indirectly convey what the speaker desires without having to directly impose on the addressee. Hints are one of the frequent ways for making indirect requests (Gibbs, 1986a).

2. **Association cues:** Larry says to David *Oh God, I've got a headache again* as a request to borrow David's swimming suit if both Larry and David know they share an association between Larry's having a headache and his wanting to borrow David's swimsuit to swim off the headache. Associations work best when the speaker and listeners share specific mutual beliefs or common ground (Clark & Carlson, 1981).

3. **Understatements:** These indirectly imply messages by saying less than is required and providing less information than is required. The speaker thus invites the listener to consider the reason that he or she is providing insufficient information. For instance, Rick and Stacey are looking at a house they are considering buying. Rick says *It needs a touch of paint* when both can obviously see that the house is in poor shape. By stating less than is obviously the case, Rick indirectly conveys what he really feels without unduly criticizing the house or the person selling it (Colston, 1997a; Colston & Keller, 1998).

4. **Overstatements or hyperbole:** These involve a speaker saying more than is necessary. By exaggerating beyond the actual state of affairs, the speaker indirectly implies something beyond what is said. For instance, Greg says to Martha *There were a million people in the co-op tonight* as an excuse for his being late getting home. To take another case, Steve might say to a customer in his store, *Oh no, Mr. Smith, we never meant to cause you any trouble. Nothing could have been further from our minds. I can't imagine how you could come to that conclusion. It's out of the question.* Speaking this way indirectly conveys sarcasm, a message that might be understood by Steve and his coworkers but not by the addressee, Mr. Smith (Gibbs, 1994).

5. **Tautologies:** By uttering a tautology, a speaker encourages others to look for an interpretation of the noninformative, patently true, utterance. These may take the form of excuses such as *War is war* or *Boys will be boys*, or as a criticism as in *Your clothes belong where your clothes belong, my clothes belong where my clothes belong* (Gibbs & McCarrell, 1990).

6. **Contradiction:** By stating responses that contradict each other, speakers make it appear that they cannot be telling the truth. Speakers
thereby encourage listeners to look for interpretations that reconcile the two contradictory propositions. For instance, Chris says to Arthur, *Are you upset about that?* Arthur replies *Well, I am and I'm not* and thus conveys his ambivalence. Contradictions may convey criticism; for instance, Kerry might say of a drunken friend to a telephone caller, *Well, John is here and he isn’t here.*

7. **Irony:** Speakers can convey something indirectly by saying something that appears to contradict what is really meant. *John’s a real genius* (after John has done many stupid things in a row), *Lovely neighborhood* (in a slum), and *Beautiful weather, isn’t it?* (to a postman drenched in a rainstorm). Ironic remarks are one of the most powerful weapons speakers have at their disposal to indirectly convey ideas and attitudes.

8. **Metaphor:** The meanings of metaphorical statements are usually on the record, but exactly which meanings the speaker intends may be off the record. For instance, when Tom says to Richard *Harry’s a real fish,* Tom may convey any one of a number of ideas about Harry that Richard is free to choose from and which, if confronted with, Tom may deny as part of his meaning (e.g., that Harry swims, drinks, is slimy, or cold-blooded like a fish). Many metaphorical remarks reflect widely shared metaphorical concepts or thoughts (see below).

9. **Euphemisms:** Many difficult or taboo topics can be discussed indirectly by using well-known euphemisms. For instance, one can talk of a dog defecating on the living room rug by saying *The dog went to the bathroom on the living room rug.* On the surface, this statement seems nonsensical, but we use terms like *go to the bathroom* as a metonymy to refer to a more specific act.

10. **Rhetorical questions:** Speakers may ask questions that leave their answers hanging while implying various meanings, such as *How was I to know?* (an excuse) or *How many times do I have to tell you . . . ?* (a criticism).

11. **Proverbs:** Speakers often find proverbial expressions an excellent way to pass along suggestions in an indirect manner, but also have great cultural value. A parent may urge her son to be more aggressive in his attempt to find a job by saying *The early bird catches the worm* and convey her belief without directly telling him what he should do.

12. **Vagueness:** Jonathan says to Melissa *I’m going you-know-where* or *I’m going down the road for a bit* to imply that he’s going to the local pub. When vague remarks successfully communicate, they demand that speakers and listeners share specific common ground information. Of course speakers may be vague to evade or equivocate, such as in the politician’s narrative described above.
These indirect and nonliteral strategies for communicating are pervasive in both speech and text. Many of these various indirect speech acts allow people to deny their covert communicative intentions if questioned by someone else. Speaking indirectly or nonliterally protects individuals from face-threatening acts (Goffman, 1967). For instance, it is impolite in many cultures to make requests of others directly. Making indirect requests such as *Would you mind passing the salt?* forces listeners to draw the right inference about the speaker’s desire without feeling imposed on or directly obligated to fulfill the request.

**THE TRADITIONAL VIEW OF NONLITERAL SPEECH ACT UNDERSTANDING**

How do listeners infer the meanings of nonliteral speech acts? What psychological processes characterize how listeners go from what speakers say to what they intend to communicate in discourse? Scholars from many disciplines have attempted to answer these questions. The most influential and comprehensive view comes from the philosopher H. Paul Grice’s (1989) theory of conversational implicature. Grice argued that the inferences needed to understand indirect and figurative remarks are derived from certain general principles or maxims of conversation that participants in talk exchange are mutually expected to observe (Grice, 1975, 1978, 1989). Among these are the expectation that speakers are to be informative, truthful, relevant, and clear in what they say. When an utterance appears to violate any of these maxims, as do many of those described in the previous section, listeners are expected to derive an appropriate *conversational implicature* about what the speaker intended to communicate in context given the assumption that he or she is trying to be cooperative.

For instance, consider a situation where John says to Mary, just as she lit up a cigarette, *Cigarettes are time bombs.* Understanding this metaphorical statement requires that listeners must (a) analyze what is said literally, (b) recognize that the literal meaning (i.e., that cigarettes are literally bombs that are set to explode at a specific future time) or what the speaker says is contextually inappropriate, and only then (c) infer some meaning consistent with the context and idea that the speaker must be acting cooperatively and rationally (i.e., smoking cigarettes can have devastating effects for a person at a later time). Understanding literal utterances does not demand the extra step of figuring out how a speaker’s intended meaning differs from his or her literal statement.
The standard pragmatic view, as this approach is often dubbed (Gibbs, 1994), suggests that indirect and nonliteral language should always be more difficult to process than roughly equivalent literal speech. More generally, this traditional view assumes that understanding what speakers literally say requires accessing of semantic information, whereas understanding what speakers implicate in context demands pragmatic information that is more difficult to access than semantic knowledge. A significant assumption of the Gricean view is that what speakers say is isomorphic to the literal, semantic, truth-conditional meaning of these sentences. This assumption about literal meaning turns out to be one of the most problematic questions in theories of nonliteral speech acts.

EARLY EMPIRICAL TESTS OF THE TRADITIONAL VIEW

Grice’s ideas about indirect, figurative language use (see also Searle, 1979) has been the subject of much debate, again in virtually all disciplines with an interest in discourse processes. Perhaps the greatest attention to the traditional Gricean view is seen in experimental psycholinguistics. Psycholinguistics ask two related questions. Is indirect and nonliteral speech necessarily more difficult to comprehend than literal language? Must people first analyze the literal meanings of utterances before deriving their intended nonliteral messages?

Psychologists have experimentally examined these issues (see Gibbs, 1994; Glucksberg, 1998). Numerous reading-time and phrase classification studies demonstrate, for example, that listeners/readers can often understand the figurative interpretations of metaphors, irony/sarcasm, idioms, proverbs, and indirect speech acts without having to first analyze and reject their literal meanings when these expressions are seen in realistic social contexts. People can read figurative utterances (i.e., You’re a fine friend meaning “You’re a bad friend”) as quickly, sometimes even more quickly, as literal uses of the same expressions in different contexts or equivalent nonfigurative expressions. These experimental findings demonstrate that the traditional view of indirect and figurative language as always requiring additional cognitive effort to be understood has little psychological validity. It appears that with sufficient context people can understand the intended, pragmatic meanings of indirect and figurative language without first analyzing the literal, semantic analysis of what a speaker says.

The idea that people can immediately understand the intended meanings of figurative discourse has often been called the direct access view (Gibbs, 1984). This alternative to the Gricean view has become the focus of considerable research in experimental psycholin-
The direct access view suggests that people need not automatically analyze the *complete* literal meanings of linguistic expressions before accessing pragmatic knowledge to figure out what speakers mean to communicate. People clearly analyze the words spoken or written. However, they do not necessarily combine these individual word meanings to form a complete literal representation of an utterance as part of their immediate comprehension of nonliteral speech.

**DEFENDING THE PRIMACY OF LITERAL MEANING**

In recent years, several empirical studies have attempted to reestablish the importance of literal meaning in figurative language understanding. For instance, one study examined irony comprehension by asking people to read stories ending with statements (e.g., *You are just in time*) that could have either literal or ironic meaning depending on the context (Giora, Fein, & Schwartz, 1998). Participants read these stories one line at a time on a computer screen. The results show that people took more time to read the final statements in ironic contexts than when these same expressions were seen in literal contexts (see also Dews & Winner, 1999). This finding is consistent with the traditional standard pragmatic model and contrary to the findings reported by Gibbs (1986b, 1986c), where people actually took less time to read ironies than literal uses of the same expressions.

These recent data raise several methodological and theoretical issues. First, the fact that people take longer to read a sentence in one context than in another does not directly imply anything about whether people first process the putative literal meanings of these sentences. Second, listeners may take longer to understand a novel expression because of the difficulty in integrating the figurative meaning with the context and not because listeners are first analyzing and then rejecting the expression’s literal meaning. Various empirical evidence supports this idea (Schaw, 1995; Shinjo & Myers, 1987). Moreover, the story contexts examined in different studies of nonliteral speech understanding often vary considerably. Some story contexts, like those employed by Gibbs (1986a, 1986b) in studies of irony comprehension, set up ironic situations that likely facilitated processing of speakers’ ironic statements. People conceptualize many situations ironically (Gibbs, 1994; Lucariello, 1994) and can understand someone’s ironic or sarcastic comment without having to engage in the additional computation. Additional processing may be needed when people understand ironic remarks in situations that are inherently less ironic. People may still need to draw complex inferences when understanding some ironic statements, but part of these inferences can occur before
one actually encounters an ironic utterance. Various contextual factors influence linguistic processing so irony need not always stake longer to process than literal language (see Katz & Lee, 1993; Katz & Pexman, 1997, for studies on the influence of different discourse factors in irony comprehension).

The most vigorous defense of the traditional Gricean view is seen in recent research on proverb understanding. Temple and Honeck (1999) criticized the direct access view by arguing, “In our view, the traditional problem is that the direct access position places essentially all the basis for understanding on the context and none whatsoever on the proverb. That is, people are presumed to already have an understanding of some topic, with the proverb serving as a mere redundant confirmation of it” (p. 47).

Temple and Honeck’s hypothesis, similar to the standard pragmatic model, assumes that a complete construction of a proverb’s literal meaning is an essential step in a theory of how proverbs are understood. To test their hypothesis, Temple and Honeck asked participants to read two 2-sentence contexts. In the first condition, one context was relevant to the literal meaning of a proverb and the other was literally irrelevant to the proverb (i.e., the literal condition). In the second condition, one context was relevant to the figurative meaning of a proverb and the other was figuratively irrelevant to the proverb (i.e., the figurative condition). Participants in the first study either saw all literal or figurative contexts. For both conditions, after reading both contexts, participants read a novel proverb (e.g., *The cow gives good milk but kicks over the pail*) and then quickly judged which context the expression was most meaningfully related to. For both conditions, then, the literally relevant or figuratively relevant contexts were the ones that participants were expected to select. The latencies to make these choices were measured. The data showed that reaction times were longer in the figurative condition than in the literal one—a finding taken as supporting the multistage view of proverb comprehension. Follow-up studies essentially replicated this main finding.

The main difficulty with these findings is that the choice reaction-time task conflates the process of comprehension with the process of judging the appropriateness of a proverb against two different situations. This metalinguistic task is different from what people ordinarily do in everyday conversations or when reading. People interpret statements by inferring speakers’ communicative intentions given the context at hand (i.e., the speaker and listener’s common ground), not by judging which of two contexts a statement relates to.

Moreover, the fact that there is a difference in choice RT to figurative and literal contexts does not imply that a proverb’s complete literal
meaning was analyzed as part of people’s understanding of the non-literal interpretations. Temple and Honeck admitted as much when they stated.

. . . this study has not directly shown that literal meaning was in fact used to construct nonliteral meaning. Rather, it was shown that it takes less time to understand that a proverb, such as *The used key is always bright*, is about keys, brightness, and a general key-using schema, than about frequently used instruments retaining their functional value. (p. 67)

Unlike the Temple and Honeck work, most recent studies examining the role that literal meaning plays in figurative language understanding employ online methodologies that are sensitive to rapid activations of meanings. These online studies are presumed to be better indicators of literal meaning activation than are more global measures of utterance comprehension, such as reading time and phrase classification techniques.

For instance, one research project examined comprehension of familiar and less familiar metaphorical expressions (Blasko & Connine, 1993). Participants in these experiments heard different sentences and made lexical decisions at various times to visually presented word strings. For instance, as participants heard the sentence *The belief that hard work is a ladder is common to this generation*, they were visually presented a letter string immediately after hearing the word *ladder*. The letter string visually presented was related to some aspect of the sentence’s literal meaning (e.g., *rungs*), a letter string related to the sentence’s metaphorical meaning (e.g., *advance*), or a control word unrelated to the sentence (e.g., *pastry*). The results reveal that participants were equally fast in responding to the literal and metaphorical targets, which were both faster than the latencies to the controls. This was true both when participants made their lexical decisions immediately after hearing the critical word (e.g., *ladder*) and when the same decisions were made 300 msecs after hearing the critical word. However, the findings differed when participants made these same types of lexical decisions to literal and metaphorical targets having heard less familiar expressions, such as *The thought that a good professor is an oasis was clung to by the entire class*. In this case, only literal targets were primed immediately after hearing the critical word (e.g., *oasis*), whereas responses to the metaphorical targets were facilitated only 750 msecs after the critical word.

How do these experimental results relate to the traditional Gricean and direct access views? My argument is that this type of experiment examines different aspects of meaning (word vs. phrasal) when it com-
pares response times to literal (word) and metaphoric (phrasal) targets. The literal target _rung_ is a simple semantic associate of the word _ladder_, whereas the metaphoric target _advance_ only relates to the general meaning of the entire expression. This makes it difficult to conclude anything about the time course under which literal meanings of entire sentences are activated compared to figurative meanings of these expressions (i.e., the hypotheses proposed by the traditional Gricean and the direct access models). Furthermore, the words used as literal and metaphoric targets in these studies do not seem to reflect distinctive literal and figurative meanings. The literal target _rung_, for instance, is related to the idea of advancing (i.e., the figurative target) given that climbing ladders, even literally speaking, is one kind advancing along some physical path.

These criticisms of the work on online metaphor processing apply to several of the other online studies of nonliteral speech understanding, including those looking at idiom (Cacciari & Tabossi, 1988) and irony (Giora & Fein, 1999a) comprehension. These online studies each mistakenly assume that there is only one kind of literal meaning (i.e., word meaning). These studies fail to distinguish between the literal meanings of words with the literal meanings of entire phrases or expressions. The fact that some aspects of word meanings are quickly activated and consequently responded to faster than the figurative meanings of entire phrases is not at all unexpected. These studies do not show that people combine word meanings to form literal meanings for an entire expression as an obligatory part of figurative language interpretation (i.e., the traditional Gricean model). For this reason, the results of these online studies do not directly bear on either the Gricean or direct access views.

More recent studies have employed different online methods to examine the role that literal meaning has in figurative language processing. One set of experiments investigated processing of sentences containing place-for-institution metonymies such as _That blasphemous woman had to answer to the convent_ by measuring participants’ eye movements as they read (Frisson & Pickering, 1999). Results show that people were as fast to understand these familiar metonymies as to read literal sentences, and that processing unfamiliar metaphors took more time than did reading compatible literal sentences. Thus, figurative language processing need not be delayed for familiar metonymies. A second study showed similar findings for familiar place-for-event metonymies such as _A lot of Americans protested during Vietnam_ and unfamiliar ones such as _A lot of Americans protested during Finland_. Frisson and Pickering argued that the overall findings do not support either a literal-first or figurative-first model, but fit best with a parallel
model where a single underspecified representation is compatible with both literal and metonymic senses.

A related set of studies examined the time course for understanding literal and figurative interpretations of simple sentences using a signal, speed-accuracy trade-off procedure (McElree & Nordlie, 1999). Participants were presented strings of words, one at a time, at a rate that approximated fast reading (250 msec/word). The final word in each string forced either a literal (e.g., some tunnels are sewers), figurative (e.g., some mouths are sewers), or nonsensical (e.g., some cattle are sewers) interpretation. Participants judged whether each word string was meaningful when a tone appeared at varying times after the critical last word. No differences were found in the comprehension speed for literal and figurative strings. McElree and Nordlie argued that the lack of time-course differences is inconsistent with the claim that figurative interpretations are computed after a literal meaning had been analyzed. In general, the time-course data presumably support the idea that literal and figurative interpretations are computed in parallel.

This pair of recent online studies presents results that are clearly contrary to the traditional Gricean view where people are assumed to analyze the complete literal meanings of utterances before deriving their figurative meanings. Yet I question whether the null results (e.g., no difference in processing literal and figurative sentences) obtained in these experiments necessarily provides evidence in favor of a parallel processing model. The activation of a particular meaning (i.e., literal or idiomatic) is assumed to reflect the output of entirely different linguistic processes. The possibility remains, however, that activation of different kinds of meaning (i.e., literal or idiomatic) may arise from a single linguistic process. The fact that scholars label one kind of meaning as literal and another figurative does not necessarily indicate that different processes operate (i.e., a literal processing mode and an idiomatic or figurative processing mode) to access these meanings (either in a serial or parallel manner).

My suggestion is that scholars resist interpreting the findings of different online studies of sentence processing, including those looking at literal meaning in figurative language understanding as necessarily demonstrating different linguistic processes. An important consequence of this idea is that differences in the activation of literal and figurative meanings should not be viewed as evidence for the primacy of literal processing in utterance interpretation. One need not postulate different literal and figurative processing modes to account for any of the data obtained in these studies.

Another reason to question whether different linguistic meanings reflect different linguistic processes is the fact that there are numerous
types of figurative meaning, including metaphoric, idiomatic, metonymic, ironic, satirical, proverbial, hyperbolic, oxymoronic, and so on (Gibbs, 1994). Scholars often assume within the context of a single set of studies that there are two processes at work during figurative language understanding, such as literal versus idiomatic, literal versus metaphoric, literal versus metonymic, and so on. Yet if there are numerous types of meaning, must there be dozens of types of linguistic processes all at work or potentially at work when language is understood? Psycholinguists have not addressed this question primarily because they focus too narrowly on one kind of figurative meaning against a simple view of literal meaning.

My main argument has been that recent research in psycholinguistics provides little support for the traditional Gricean view and is most consistent with the direct access view of figurative language understanding. The occasional findings that some figurative utterances may take longer to process than literal expressions is insufficient to support the traditional model or reject the direct access view. Similarly, showing that people access aspects of what might be referred to as the literal meanings of words in figurative statements also does not imply that the direct access view is wrong (see Gibbs, 1994, 2002, for discussion of how to falsify the direct access view).

**WHAT IS LITERAL MEANING?**

The original proposals on literal meaning in nonliteral speech act understanding assumed that an analysis of what speakers literally say is in many cases an important part of interpreting what speakers implicate in context. The Gricean view, for example, assumes that what a speaker says is equivalent with an utterance’s semantic or literal meaning. Most linguists, philosophers, and psycholinguists maintain a similar belief about the close, if not isomorphic, relationship between literal meaning and what speakers say. Many scholars maintain that pragmatics plays at best only a small part in determining what speakers literally say. The most recent psycholinguistic studies described earlier also implicitly adhered to the distinction between context-free literal meaning and pragmatically determined indirect or figurative meaning.

There are several problems with this linkage of literal meaning to what speakers say that have hurt empirical work on nonliteral speech act understanding. First, many aspects of what might be termed literal speech have their roots in figurative thought and language. Cognitive linguists have demonstrated that different figurative schemes of thought—most notably metaphor and metonymy—underlie system-
atic patterns of conventional linguistic expressions, historical development of word and phrase meanings, and novel linguistic expressions patterns of linguistic expressions (Johnson, 1987; Lakoff, 1987; Lakoff & Johnson, 1980, 1999; Sweetser, 1990). A wide variety of experimental research in psycholinguistics shows that people have tacit knowledge of conceptual metaphors—for instance, that partly motivates how people learn, make sense of, and comprehend different linguistic expressions (Gibbs, 1994). These studies collectively call into question any simple distinction between literal and figurative meanings and emphasize the figurative nature of much language that people at times may refer to as literal. Simply referring to some pieces of language as literal and others as figurative, or whatever other tropes may be of interest, does not empirically establish that literal meanings are entirely different than are figurative meanings or are produced and understood by different cognitive mechanisms.

In fact, it is not clear what the operational definition of literal meaning is in most psycholinguistic experiments. These studies individually compare metaphor versus literal meaning, irony versus literal meaning, idiomatic versus literal meaning, metonymy versus literal meaning, and so on. Yet across the vast number of empirical studies that compared literal and nonliteral meaning, the variety of forms for literal utterances is as great as the differences among metaphors, metonymies, ironies, and so on. Yet scholars continue to assume that the literal meaning they examine is somehow empirically the same variable that other researchers investigate in their experiments. My continued plea is that scholars not refer to any linguistic expression as literal unless theoretical reasons can be clearly stated as to what makes this type of meaning different from all other kinds of meaning (e.g., metaphoric, metonymic, poetic, etc.). I raise this issue here as an important challenge for researchers looking at the role of so-called literal meanings in understanding what speakers say and implicate.

Finally, one of the greatest puzzles in the traditional Gricean view is that it is not clear what role literal meaning actually plays in indirect speech act understanding. For the most part, literal meanings are only rejected along the way to understanding what speakers actually imply. Psycholinguistic studies also say nothing about how literal meaning functions to derive speaking meaning regardless of whether literal meaning is processed first or in parallel to understanding indirect or figurative meaning. My claim is that what speakers say actually plays a major role in understanding what people actually intend to communicate. This is best understood when we reconceive of what speakers say as itself having pragmatic meaning. Significant aspects of what speakers say, and not just what they totally communicate, are deeply de-
Under the new theory, significant aspects of what speakers say, and not just what they totally communicate, are deeply dependent on enriched pragmatic knowledge. Under this revised theory, people may indeed analyze aspects of what speakers pragmatically say, and not just semantically or literally say, as part of understanding what speakers conversationally implicate when using indirect and nonliteral speech acts.

In recent years, several linguists, philosophers, and psychologists have persuasively argued that the traditional, Gricean view ignores that essentially the same sorts of inferential processes used to determine what speakers imply (i.e., conversational implicatures) also enter into determining what speakers say (Bezuidenhout, 1997; Carston, 1988, 1993; Gibbs, in press-a; Gibbs & Moise, 1997; Recanati, 1989, 1993; Sperber & Wilson, 1986; Wilson & Sperber, 1993). One set of psycholinguistic studies supports the idea that pragmatics play a major role in people's intuitions of not only what speakers indirectly implicate, but also what speakers say (Gibbs & Moise, 1997; see also Gibbs, 1999b; Nicolle & Clark, 1999).

Consider the expression *Jane has three children*. According to the Gricean view, the interpretation that Jane has exactly three children comes from applying specific pragmatic information to the minimally pragmatic (or semantic) proposition of what is said—a process that results in what Grice referred to as a generalized conversational implicature (i.e., implicatures that are normally drawn regardless of the context). Yet Gibbs and Moise showed in series of experiments that people do not equate a minimal meaning (i.e., Jane has at least three children and may have more than three) with what a speaker says. Instead, when asked to choose the best paraphrase of what a speaker says, people chose significantly more enriched pragmatic paraphrases (e.g., Jane has exactly three children) than they did paraphrases that were minimally pragmatic (e.g., Jane has at least three children and may have more than three). Even when people were alerted to the Gricean position (i.e., what is said is equivalent to the minimal proposition expressed), experimental participants still replied that enriched pragmatics is part of their interpretation of what a speaker says and not just what the speaker implicates in context.
These findings demonstrate that people's ordinary intuitions about what speakers say are informed by enriched pragmatics—a result that is clearly contrary to the Gricean view. Yet can people even distinguish between what speakers say and what they implicate? Many psycholinguistic studies on figurative language comprehension seem to indicate that listeners use contextual information to determine what speakers implicate without analyzing what speakers literally say (Gibbs, 1984, 1994). Perhaps people only view what speakers say as being informed by enriched pragmatics because they are unable to distinguish what speakers say from what they implicate.

However, other studies in Gibbs and Moise (1997) cast doubt on this possibility. For instance, consider the following story:

Bill wanted to date his coworker Jane.
Being rather shy and not knowing Jane very well,
Bill asked his friend, Steve, about Jane.
Bill didn’t even know if Jane was married or not.
When Bill asked Steve about this, Steve replied,
Jane has three children.

What does Steve say and what does he implicate by his utterance? Steve pragmatically says that Jane has exactly three children (an enriched pragmatic meaning), but implicates in the context that Jane is already married. To the extent that people can understand what Steve says, but not implicates, by Jane has three children, they should distinguish between the enriched pragmatic meaning and the implicature. The results of one study show this to be true. When participants were asked to choose the best paraphrase of what a speaker says in contexts like the earlier one, they chose the enriched pragmatic meanings (i.e., Jane has exactly three children) and not implicatures (i.e., Jane is married).

These findings show that pragmatics strongly influence people's understanding of what speakers say and implicate. More generally, the Gibbs and Moise (1997) results suggest that the distinction between saying and implicating is orthogonal to the traditional divisions between semantics and pragmatics, and literal and nonliteral meanings. These empirical results illustrate the error in equating what speakers say with context-free literal meaning.

Recent studies examined the implications of this revised theory on the pragmatics of what is said for processing conversational implicatures that expand upon what speakers pragmatically say (Hamblin & Gibbs, 2000). Although there is significant psycholinguistic work to claim that people do not automatically analyze the semantic, literal meanings of speaker’s statements during figurative language under-
standing, people may actually analyze what speakers pragmatically say as part of their interpreting what speakers implicate in context.

We tested this idea in a series of reading-time experiments. Results show that people take less time to read statements (e.g., *Jane has three children*) when what speakers say is identical to what they imply (e.g., that Jane has exactly three children) than to read the same statement in discourse when what speakers say underdetermines what they imply (e.g., that Jane is married). A second study showed that people take longer to comprehend statements that express conversational implicatures in context (e.g., that *Jane has three children* implies that Jane is married) than to read direct assertions (e.g., *Jane is married*) when what speakers say is identical to what they imply. A third study indicated that people take longer to read statements in contexts conveying minimally pragmatic meanings (e.g., *Jane has three children* meaning that Jane has at least three children but may have more than three), similar to what Grice and others refer to a speaker’s literal, said meanings, than to read statements that convey what speakers pragmatically say or pragmatically implicate. This latter result demonstrates that people are *not* first analyzing the minimally pragmatic, putative literal meanings of what speakers say before deriving what they imply in context.

Hamblin and Gibbs’ data indicate that, although pragmatics is a significant determinant of what speakers say, people still take longer to interpret what speakers implicate when these meanings extend beyond what speakers say. It appears that people’s complex pragmatic knowledge is applied differently when understanding what speakers pragmatically say and when interpreting what speakers implicate in context. One possibility is that there are two kinds of pragmatic information or knowledge—primary and secondary—that become activated during normal language understanding (Gibbs & Moise, 1997; Recanati, 1993). Primary pragmatic knowledge applies deep, default background knowledge to provide an interpretation of what speakers say. Secondary pragmatic knowledge refers to information from context to provide an interpretation of what speakers implicate in discourse.

The Hamblin and Gibbs (2000) data suggest that people’s deep background knowledge, as part of primary pragmatics, initially dominates the application of secondary pragmatic information to reveal what speakers say as distinct from what they implicate. This idea is consistent with the findings in Gibbs and Moise (1997) that people are capable of distinguishing what speakers pragmatically say from what they pragmatically implicate in discourse. When speakers implicate something beyond what they say, listeners may require addi-
tional cognitive effort to process these meanings over than needed to understand speakers’ utterances when what they say is identical to what they imply.

The recent empirical findings on processing speakers’ indirect meanings raise several important possibilities about how people understanding nonliteral speech acts. Perhaps the most intriguing idea is that what speakers say may indeed have a major role in understanding what they figuratively communicate. What speakers say is not a semantic or literal meaning, but an enriched pragmatic one. In some cases, listeners understand speakers’ figurative meanings as part of their comprehension of what they pragmatically say. For example, linguists (Groesfema, 1992; Papafragou, 1996) have argued that the indirect meanings of many indirect speech acts such as Can you pass the salt? and certain ironies and metonymies may arise from an analysis of what speakers say without any need to draw more complex conversational implicatures. This possibility would be consistent with psycholinguistic studies showing, for instance, that people can process some indirect and figurative utterances quite quickly in context (Gibbs, 1994). In other cases, listeners must expand on what speakers pragmatically say to infer what they figuratively imply. Thus, people may need to derive conversational implicatures to understand indirect and figurative meanings of some linguistic expressions, such as those studied by Hamblin and Gibbs (2000), and perhaps some unfamiliar metaphors and ironies that seem to take longer to process than corresponding direct expressions.

One challenge for future studies is to recognize the difference between how people understand indirect and figurative language that demand some expansion on what speakers say (i.e., understanding what speakers imply by virtue of what they say) from interpreting nonliteral meanings that require a rejection of what speakers say (i.e., understanding what speakers imply despite what they say). There is clearly much room for empirical work, beginning perhaps with a detailed analysis of past studies to see exactly what kinds of figurative utterances are processed quickly and whether these are understood as part of what speakers pragmatically say and not just part of what they pragmatically imply.

**METAPHOR IN THOUGHT AND LANGUAGE**

The most widely discussed and studied type of nonliteral speech act is metaphor. Consider the following conversation between a man and his psychotherapist where the two participants negotiate the meaning and personal implications of several metaphors (Ferrara, 1994). Howard
is a client in his 30s who is talking with Judy, his therapist, in their third session. One month earlier, Howard had been fired from his job as an orderly in a hospital because of suspicion over some missing drugs. Howard maintained that he did not steal the drugs and was eventually reinstated.

Judy: When you have a problem, what do you do with it?
Howard: I usually let it be a problem. I don’t usually do anything much or I. . . I was thinking about that the other day.
Judy: Does the problem go away if you don’t do anything about it?
Howard: No, it gets worse . . or it just complicates things as you go further down the road.
Judy: Can you look at your own life, kind of on a continuum? Look down the road of that line and see what that’s gonna do. . . in your life?
Howard: Look on down the road.
Judy: Yeah, kinda visualize what un. . your own life will be like if you don’t deal with some of it. . . your problems. . . Can you see how it might complicate. . . your life?
Howard: It will just continue the way it is.
Judy: Kind of like a snowball. . . effect
Howard: No no not a snowball. Just kinda floating, floating down the river.
Judy: Floating down the river.
Howard: That’s what I’m doing now. That’s what I was afraid I was gonna go back into all this. I said something the first time I talked to you about
Judy: Yeah.
Howard: floating and being afraid of going back into floating. That’s just you know, floating, drifting . .
Judy: So you’re adrift right now?
Howard: Yeah. And feel deal and I feel like I’m- I drink to feel a little bit deader. No, that’s not true.
Judy: Feel depressed. . . or numb?
Howard: Yeah.
Judy: Numb, you feel?
Howard: Yeah. Yeah.
Judy: What’s it like to be floating down the river? Tell me more.
Howard: It’s comfortable. It’s safe. . . Everything just keeps on an even keel, you know.
Judy: Mmmhmm.
Howard: You're just kinda floating . . .
Judy: Kind of in a canoe? . . . going down the river, or
Howard: No, more like a great ole big barge . . . on a great old big river.
Judy: Barge, very stable, kinda.
Howard: Yeah, plenty of room to spread out and . . . sit in the sun. Yeah, and you don’t have to worry about falling off the edge.
Judy: Mmhmm.
Howard: And sun, you know, it’s kinda hazy. It’s not really clear sun. It’s kinda hazy.
Judy: Mm-hmm.
Howard: Kinda half asleep, that what’s it’s like.
Judy: What happens when you kind of come to the . . . falls, the falls that are down there, about two miles down the river?
Howard: Get the hell off the river.
Judy: That’s certainly one way to handle it. Get out.
Howard: I feel a lot of discomfort. That’s what happened just last month. I hit those falls last month. (noise)
Judy: I don’t know why it did that. So that’s what happened . . . um this . . . last time there was kind of um . . . an external situation that sort of forced you out of your boat.
Howard: It was uncomfortable but I was, I was pretty, I was enjoying it too. And I didn’t want to go back to just floating. It was uncomfortable and I was out, I don’t know, I been floating a long time.
Judy: Mmhmm . . . Well, you’ve found that it works for you . . . in a sense.
Howard: What works for me?
Judy: Floating.
Howard: Because I’m . . . stay . . . comfortable and
Judy: In a sense, but it may . . . now be inappropriate. It may not be working as well . . . as it did in the past.
Howard: Mnnn . . . Yeah, I’d like to have a little excitement now and then.
Judy: Some rapids.
Howard: Yeah (laughs) Something I can keep in control of maybe and not drown. But . . . yeah, I think I am bored.

This conversation exhibits how people are capable of understanding one another’s messages even when what is said does not fully convey their intended meanings. Yet the conversation is notable because of
the way metaphor structures the discourse where participants mutually negotiate the meanings of the related metaphorical expressions. For instance, the client, Howard, introduces the metaphorical idea of his future life as being down the road. The therapist picks up on this idea and asks Howard to elaborate on the metaphor in her questions about visualizing the road ahead and whether he feels adrift at present. Howard provides more detail when he rejects the therapist’s question about life being like a snowball effect and advances how his experience feels like Just kinda floating, floating, down the river. Client and therapist further extend their metaphorical characterization of how Howard feels about his life by talking of keeping on an even keel, experiencing some rapids, and not having to worry about falling off the edge. The metaphorical meanings here seem quite appropriate and perhaps highly desirable in this therapeutic context. Howard and Judy’s conversation illustrates the value in talking metaphorically about one’s experiences in life.

One of the main issues for psychologists studying metaphor comprehension focuses on how people correctly map meanings from a source domain onto a target domain when understanding verbal metaphors. For instance, when understanding Howard’s comments about his life being like floating down a river, not all of our knowledge about river journeys (e.g., that one can get very wet) get applied to create a new mental understanding of Howard’s recent experiences in life. Only part of what we know about river journeys gets applied to restructure the target domain such that we understand Howard’s recent experiences as one where he just lets things happen and does not exert much control over events. Various major theories have been proposed to account for metaphorical mappings, and the empirical literature on metaphor comprehension is immense (see Gibbs, 1994). I briefly talk here about two new ideas about metaphor comprehension and then discuss how metaphors in language arise from metaphors in thought.

The first new proposal suggests that interpreting metaphors involves the basic cognitive mechanism of suppression (Gernsbacher & Robertson, 1999). Suppression is defined as a general cognitive mechanism that functions to attenuate the interference caused by the activation of extraneous, inappropriate information. A wide variety of empirical evidence shows that suppression plays an important role in attenuating extraneous information during numerous phases of linguistic processing (see Gernsbacher & Robertson, 1999, for a review).

Suppression may provide a mechanism to attenuate irrelevant aspects of source domain information during metaphor interpretation. Thus, understanding the metaphor Lawyers are sharks requires that shark attributes such as swimming, living in the ocean, having
fins, and so on should be suppressed, which enhances relevant attributes such as tenacity, ferocity, and aggressiveness. Evidence in support of this claim was shown in studies where participants were faster to verify the statement *Sharks are tenacious* having first read the metaphor *Lawyers are sharks* than when they first saw *Hammerheads are sharks*. Reading the metaphor *Lawyers are sharks* also slowed people’s verification of the statement *Sharks are good swimmers* compared with when they first read *Hammerheads are sharks*. This pattern of findings shows that metaphor understanding requires both the suppression of irrelevant attributes in the source domain and enhancement of the appropriate ones. Gernsbacher and Robertson (1999) described the importance of suppression for understanding other nonliteral speech acts including idioms, ironies, and proverbs.

A second proposal that merits attention is the *graded salience hypothesis* (Giora, 1997; Giora & Fein, 1999b; Giora et al., 1998). According to this idea, highly salient meanings of both words and phrases are automatically processed during the initial stages of figurative language comprehension. One consequence of this idea is that processing familiar metaphors (e.g., *step on someone’s toes*) should activate both of their literal (e.g., foot) and metaphoric (e.g., offend) meanings even when these metaphors are seen in appropriate discourse contexts. Processing unfamiliar metaphors (e.g., *Their bone density is not like ours*) may only initially activate their literal meanings as these are most salient. Different empirical studies, ranging from reading-time to word-fragment completion experiments, provide evidence in support of this general idea (but again note that the literal meaning activated always corresponds to word meaning, whereas the figurative meaning always corresponds to phrase meaning).

Both the suppression and graded salience hypotheses are important to consider because they attempt to explain a wide range of figurative language phenomena and involve cognitive mechanisms that also appear to function during many others aspects of linguistic processing. Of course these hypotheses are not entirely independent and possibly relate to other theories of metaphor comprehension proposed earlier. For instance, the suppression hypothesis is explicitly seen as having an important role within Glucksberg and Keysar’s (1990; Glucksberg & McGlone, 1999) class inclusion model for metaphor comprehension. A significant trend in metaphor comprehension research is for scholars to propose hybrid models to account for the diversity of linguistic examples that most single theories seem incapable of explaining (e.g., see Shen, 1999). The two new proposal discussed here are consistent with this trend.
One question that has interested many discourse scholars concerns the relationship between metaphor in language and metaphor in thought (i.e., conceptual metaphors). The therapist–client conversation presented earlier clearly shows how two people can conceptualize one’s life in metaphorical terms. Over the past 20 years, cognitive linguists have argued from their exploration of systematic relationships among individual linguistic statements of how many aspects of metaphor in language are grounded in conceptual metaphors, including ideas like those seen in the dialogue between the therapist and her client where LIFE IS A JOURNEY structured much of the talk (Johnson, 1987; Lakoff, 1987; Lakoff & Johnson, 1980, 1999; Sweetser, 1990).

Many psychological studies demonstrate that nonliteral speech acts are easy to produce and understand because people actually conceptualize their experiences in figurative (e.g., metaphorical, metonymic) terms (Gibbs, 1994). These include experimental studies that have looked at people’s mental imagery for idioms (Gibbs & O’Brien, 1990), people’s context-sensitive use of idioms (Nayak & Gibbs, 1990; Gibbs & Nayak, 1991) and euphemistic phrases (Pfaff, Gibbs, & Johnson, 1997), people’s folk understanding of how the source domains in conceptual metaphors constrain what idioms (Gibbs, 1992) and proverbs mean (Gibbs, Strom, & Spivey-Knowlton, 1997), people’s use of conceptual metaphors in organizing information in text processing (Allbritton, McKoon, & Gerrig, 1995), and people’s use of conceptual metaphors in drawing inferences when reading poetic metaphors (Gibbs & Nascimento, 1996).

These psycholinguistic studies employ appropriate indirect methods for assessing people’s cognitive processes used when speaking and listening. Thus, these studies do not rely on the linguistic intuitions of language theorists to distinguish between varying theoretical accounts of metaphorical language use. Instead, these studies employ techniques that ask participants, who are ignorant of the hypotheses tested, to imagine expressions, read them in a speeded manner, make various speeded and nonspeeded judgments about metaphors, or match verbal expressions according to their similarity of meaning. These indirect methods from cognitive psychology are essential if we are to infer anything about rapid, mostly unconscious mental processes.

The claim that preexisting conceptual metaphors influence significant aspects of how people understand conventional and nonliteral speech acts has been criticized by various psychologists (Glucksberg & Keysar, 1990; Glucksberg et al., 1993; Glucksberg & McGlone, 1999; Kreuz & Graesser, 1991; McGlone, 1996). Much of the debate over whether metaphors of thought influence verbal metaphor under-
standing centers on the most appropriate methodology for examining linguistic understanding. One criticism about some of the prior empirical research is that asking people about their intuitions as to why figurative expressions mean what they do is an unreliable way to examine the conceptual foundations for figurative meaning.

For example, Keysar and Bly (1995, 1999) showed that if people were told that the meaning of an idiom like *The goose hangs high* is “things look bad,” when in fact its original meaning was “things look good,” they believed that the meaning presented to them originally made more sense and best captured what the phrase really means. Keysar and Bly (1995) interpreted these findings to suggest that methodologies that assess people’s intuitions about why idioms mean what they do should not be trusted. Unfortunately, the vast majority of the idioms they studied are based on metonymy and not metaphor. Thus, the phrase *The goose hangs high* means “things look good” because the act of hanging a dead goose up for all to see metonymically stands for an entire sequence of events leading up to the successful slaughter of the goose for food. Contemporary speakers often have great difficulty explaining why metonymically based idioms mean what they do even for widely used expressions (e.g., *kick the bucket*).

A related criticism of the conceptual metaphor view was examined by asking participants to paraphrase verbal metaphors, such as *The lecture was a three-course meal* (McGlone, 1996). Only 24% of these paraphrases contained any references consistent with underlying conceptual metaphors, such as IDEAS ARE FOOD. Even when participants were asked to give figurative paraphrases of the verbal metaphors, they still most frequently produced paraphrases inconsistent with related conceptual metaphors. Thus, when given the verbal metaphor *Dr. Moreland’s lecture was a three-course meal for the mind*, only one third of the paraphrased mentioned source domain terms (e.g., *food*) related to the conceptual metaphor IDEAS ARE FOOD. The findings from this and other studies were taken to imply that people’s interpretations of verbal metaphors are not necessarily related to their putative, underlying conceptual metaphors.

These data do not clarify that having people verbally paraphrase a metaphor is the best method for tapping into different types of possibly metaphorical knowledge that might be used when people interpret or make sense of verbal metaphors. The debate here has shifted somewhat in the last decade. Critics acknowledge that conceptual metaphors may be used when people judge the appropriateness of figurative expressions in context (Glucksberg, Brown, & McGlone, 1993; Nayak & Gibbs, 1990). Yet conceptual metaphors may not be automatically activated during online processing of metaphorical expressions.
(Glucksberg et al., 1993). However, the results of several online priming studies suggest that people are immediately computing or accessing at least something related to the conceptual metaphors (e.g., ANGER IS HEATED FLUID IN A CONTAINER) when they read idioms (e.g., blow your stack, hit the ceiling; Gibbs, Bogdonovich, Sykes, & Barr, 1997). More online studies may likely examine other aspects of whether conceptual metaphors are automatically accessed during metaphoric language understanding, as well as exactly what role these metaphors of thought play in people's immediate processing of figurative meaning.

Another question that linguists and psychologists have debated addresses why some conceptual metaphors, but not others, are used by people in speaking about abstract concepts. There is a large body of work in cognitive linguistics suggesting that much metaphorical thinking arises from recurring patterns of embodied experience (Johnson, 1987; Lakoff & Johnson, 1980, 1999). Consider the following narrative from a 40-year-old woman who was summarizing her life after the death of her parents and her own long struggle with eating disorders (Bullitt-Jones, 1999):

As I look over our lives—my father’s, my mother’s, my own—I can see plainly that this is a story of desire: desire sometimes gone awry, sometimes stifled and repressed, sometimes fulfilled, sometimes set free. It’s no easy business, learning to befriend our desires, to get past our lesser appetites in order to find what we really long for. It’s hard to set a course and sail straight in the face of high winds, heavy storms, and rough seas. And no one can chart the course for us. Sure, we sail together; hands on deck, every hand we know that will reach out to help us. And we should. But we’re not so different from the ancients when it comes to the business of life and death, and the meaning of the journey. In my case, I hungered, I yearned for something—or Someone—that would really fill me up, fill up my life, give me something to live for, something larger than the ordinary of everyday but found there, nevertheless, in the turning of the days and the seasons, the rising and the setting of the sun, in the sheer gift of being alive.

By the time my father died, I knew I was on my way. I had set my course. I knew that whatever my life was about, it was about desire, the desire beyond all desire, the desire for God. It was about learning to listen to my deepest hunger and to let this hunger guide me, as a ship steers at night by the stars.

This beautiful narrative focuses around the metaphorical ideas of LIFE IS A JOURNEY and DESIRE IS HUNGER. The narrator does not simply list different individual metaphors, each of which is based on
an entirely new source to target domain conceptual mappings. Instead, she talks of her experiences by elaborating on rich sets of inferences or entailments that arise when one conceptualizes of life as a kind of physical journey or desire as a kind of embodied hunger. Both of these conceptual metaphors are fundamentally grounded in embodied experiences that most people share, which facilitates the structuring of abstract target domains. Cognitive linguists have recently described how many conceptual metaphors are fundamentally grounded in enduring patterns of metaphorical correlations in our ordinary experience (e.g., the positive correlation between feeling different desires and feeling hungry; Lakoff & Johnson, 1999).

There are several studies from cognitive psychology that support the idea that people’s embodied experiences give rise to the metaphorical structuring of abstract concepts, which in turn constrains speakers’ use and understanding of figurative language. These studies first independently investigated people’s understanding of the source domains for these embodied metaphors. This was done independently of the way these source domains were employed in linguistic expressions. We then used the data from these independent analyses of embodied source domains to make predictions about what gets mapped onto different target domains in conceptual metaphors (Gibbs, Lima, & Francuzo, in press).

For instance, Gibbs (1992) examined people’s intuitions of the bodily experiences of containment and several other image schemes. These image schemas serve as the source domains for several important conceptual metaphors underlying speakers’ use and understanding of idioms, such as *blow your stack*, *flip your lid*, and *hit the ceiling* (Gibbs, 1992). These studies were designed to show that the specific entailments of idioms reflect the source to target domain mappings of their underlying conceptual metaphors. Most important, these metaphorical mappings preserve the cognitive topology of these embodied, image-schematic source domains. This is why people recognize that idioms such as *blow your stack* have specific meanings not equivalent to literal paraphrases, such as *to get very angry*.

Other experimental studies showed that people’s understandings of metaphorical expressions about human desires (e.g., *I am starved for his affection, I am hungry for power and fame*) are motivated by people’s embodied experiences related to feeling hunger, which form the source domain for the conceptual metaphor DESIRE IS HUNGER (Gibbs et al., in press). This is true for native speakers of both English and Portuguese. In particular, people’s embodied understanding of the variety of ways they experience hunger gets mapped onto their understanding of the abstract domain of human desires.
No empirical research within psychology has examined how people’s metaphorical and embodied conceptions of abstract ideas help structure and provide coherence for texts (as opposed to individual linguistic expressions). Yet this is clearly an important future direction (note the metaphor here) for discourse studies that seek possible links between cognition and language.

**IRONY IN LANGUAGE AND THOUGHT**

Irony is another of the most widely used and frequently studied forms of nonliteral speech acts. As is true for metaphor, much ironic talk appears to reflect people’s ironic understandings of situations. Consider the following exchange between two college students (Gibbs, 2000). This conversation occurred in their apartment and focused on some visitors who were staying with them at the invitation of another roommate:

Anne: *By the way, were our wonderful guests still here when you came out and ate lunch?*
Dana: *I had a sandwich and . .
Anne: *Isn’t it so nice to have guests here?*
Dana: *Totally!*
Anne: *I just love it, you know, our housemates. They bring in the most wonderful guests in the world and they can totally relate to us.*
Dana: *Yes, they do.*
Anne: *(laughs) Like I would just love to have them here more often (laughs) so I can cook for them, I can prepare (laughs) to make them feel welcome*
Dana: *unhuh*
Anne: *I just welcome them so much, you know, ask them if they want anything to drink or eat* (laughs)

Anne and Dana’s conversation reveals the intense joy that speakers of irony often share. Each person employs different forms of ironic language (e.g., sarcasm, jocularity, rhetorical questions, hyperbole) to indirectly convey their mutual displeasure about the people staying as guests in their apartment. Much of the irony here is humorous despite
its implied criticism of the visitors (and the roommate who invited them).

In recent years, psychologists, linguists, and philosophers have proposed various theories to explain how people use and understand irony. These theories focus on widely different cognitive, linguistic, and social aspects of ironic language use, although each theory claims to provide a single umbrella for capturing the essence of irony. For instance, some scholars from many disciplines maintain that irony is a type of echoic mention, in which speakers echo or repeat a previously stated utterance or belief, which in context is recognized as conveying ironic meaning (Sperber & Wilson, 1995; Wilson & Sperber, 1992). Research shows that people, in fact, find it easier to process and judge the ironic meanings of utterances when they echo or paraphrase some earlier statement or commonly held belief (Gibbs, 1986b; Jorgensen, Miller, & Sperber, 1984).

By contrast, pretense, not echoic mention, is seen as the key to irony by other researchers (Clark, 1996; Clark & Gerrig, 1984). Under this view, speakers of irony pretend to be some other person or persona and also pretend to be speaking, in some cases, to some person other than the listener. For instance, a person may pretend to be a weather forecaster when saying *What lovely weather!* in the midst of a rainstorm. By adopting this pretense, the speaker aims to alert the listener to this other persona and what that person may have said, which in the present context takes on an ironic meaning. Many forms of indirect, figurative language, including hyperbole, understatement, and certain indirect requests, may communicate their meanings effectively because listeners recognize the nature of a speaker's pretense (Gibbs, 1994).

Other researchers argue that ironic utterances mostly accomplish their communicative intent by reminding listeners of some antecedent event, even if not all such reminders are echoic or refer to actual or implied utterances (Kreuz & Glucksberg, 1989). Many ironic remarks merely remind listeners of the attitudes and expectations they might share with speakers. The allusional pretense theory combines features of both the echoic mention and pretense view by proposing that ironic utterances convey pragmatic meaning by alluding to failed expectations, which is usually achieved by violating the maxim that speakers should be sincere in what they say (Kumon-Nakamura, Glucksberg, & Brown, 1995).

It is difficult at present to say which of the prior views of irony best characterize how people interpret ironic messages. Empirical tests of these proposals rarely compare alternative views within the same set of experiments. Most studies tend to employ forms of irony as stimuli
that best fit one theory. Most experiments tend to assume that irony is best characterized by sarcasm. No single theory of irony comprehension appears to capture the variety of forms of irony noted in interdisciplinary research.

The psycholinguistic studies examining irony comprehension only indirectly talk about the reasons that people speak ironically. Some scholars propose that irony is special because it mutes the usually negative meaning communicated by an ironic statement (Dews, Kaplan, & Winner, 1995) and appears less rude, especially when expressing trivial criticism (Jorgensen, 1996). Other research suggests that irony can be more critical in many situations than literal statements (Colston, 1997a, 1997b). Various scholars also suggest that different forms of irony often have different communicative functions (Kreuz, Long, & Church, 1991; Lee & Katz, 1998; Roberts & Kreuz, 1994) and evoke different emotional responses in listeners (Leggitt & Gibbs, 2000).

For example, one set of studies (Colston, 1997a, 1997b) demonstrated that when events turn out unexpectedly for the participants involved, and these expectations were an explicit part of the speaker/listener’s common ground, ironic statements (e.g., This is a great situation) were rated as more appropriate than were either hyperboles (e.g., This is the best situation that anyone could be in) or understatement (e.g., This is a somewhat bad situation). Yet when the expectations were not explicit and the situation had a negative outcome, people judged overstatements to be more effective than either ironic remarks or understatements. The appropriateness of a particular form of irony depends partly in the contrast between what was expected and what ensued. Additional experimental findings suggest that irony creates more contrast between expectations and resulting events than does understatement, whereas both form of irony create more contrast than do literal statements (Colston & O’Brien, 2000). Furthermore, the degree of contrast between expectation and reality also affects listeners’ judgments of how humorous, condescending, expressive of surprise, and protective of a speaker any ironic remark appears to be (Colston & Keller, 1998; Colston & O’Brien, 2000). Other studies show that when a speaker restates another person’s inaccurate remark (e.g., Sure, Ronald Reagan was the president during the 1970s), the ironic message works better than either refutation or correction to communicate the idea that the first speaker should have known better (Colston & O’Brien, 2000).

One detailed analysis of over 60 ten-minute conversations between college students in public and private settings showed the diversity of ironic language (Gibbs, 2000). Overall this corpus revealed 789 different ironic statements as students made ironic remarks 8% of the time.
This fact alone shows that irony is not a specific rhetorical device only to be used in unusual circumstances. Yet this analysis of students' ironic utterances demonstrates that irony is, again, not a single category of figurative language, but includes a variety of types, each of which is motivated by slightly different cognitive, linguistic, and social factors and conveys somewhat different pragmatic meaning. People comfortably use various forms of irony to convey a wide range of blatant and subtle interpersonal meanings.

For instance, students used jocularity (i.e., *I would just love to have them here more often*) when speakers addressed one another in humorous ways. They used sarcasm (e.g., *I would just love to have them here more often*) when speakers spoke positively to convey a more negative intent—28% of the time. They used hyperbole (e.g., *They bring in the most wonderful guests in the world and they can totally relate to us*) when speakers expressed their nonliteral meaning by exaggerating the reality of the situation—12% of the time. They used rhetorical questions (e.g., *Isn't it so nice to have guests here?*) when speakers literally asked a question that implied either a humorous or critical assertion—8% of the time. Finally, they used understatements (e.g., *James was just a bit late with his rent*) when speakers conveyed their ironic meanings by stating far less than was obviously the case—2% of the time.

An analysis of the pragmatics underlying the different forms of irony indicated that there were differences in the degree to which the speakers' ironic utterances involved echo, pretense, or a special tone of voice. Thus, people using sarcasm and hyperbole adopted pretense much more so than they echoed a previous statement, whereas speakers of jocularity employed pretense and echo mention with near equal frequency. An important difference across the various types of irony concerned asymmetry. For jocular utterances, speakers more frequently presented a negative statement to convey a positive message than they spoke positively to express negative meaning. In contrast, a far greater number of sarcastic utterances were stated positively to convey negative messages than the reverse. Speakers often used various special tones of voices with each type ironic utterance, especially with sarcasm, jocularity, and rhetorical questions. Moreover, speakers of sarcasm were significantly more critical and mocking of others than were speakers of jocularity, hyperbole, and rhetorical questions. Almost all the ironic utterances, with the exception of understatements, were viewed as humorous. Most notably here, there appears to be a strong association between an ironic utterance mocking someone or something and it being viewed as humorous.

Another interesting finding was that addressees frequently responded to speakers' ironic statements by saying something ironic in
return (as shown nicely in the previous conversation between roommates). This averaged from 21% to 33% across the five types of irony. As was the case with the conversation between the therapist and her client, where metaphor was mutually shared and negotiated, speakers of irony share ironic views that are jointly extended and celebrated on as the conversation unfolds. These data are consistent with the earlier studies demonstrating the importance of contrast between what was expected and what ensued are consistent with the idea that people use different forms of irony in various discourse situation because they conceptualize situations in ironic terms (Gibbs, 1994; Lucariello, 1994). Under this perspective, the reasons for speaking ironically are not solely located in trying to convey specific nuances of meaning, but because people view situations ironically and their ironic talk reflects this figurative mode of thought. To the extent that speakers and listeners share a similar ironic construal of events, their understanding of what speakers ironically imply by what they say is greatly facilitated.

NEW THEORETICAL APPROACHES TO NONLITERAL DISCOURSE

My discussion of nonliteral speech acts in text and discourse has centered on psychological studies and theory. Yet there are several recent theoretical developments in neighboring disciplines that warrant attention. One exciting advance in the study of indirect speech acts is the recent work on conceptual blending theory. Blending theory attempts to analyze the dynamic nature of mappings in thought and language (Fauconnier, 1997; Fauconnier & Sweetser, 1996; Fauconnier & Turner, 1998). According to this approach, mental spaces are invoked during thought and communication as partially specified constructs (frames or mental models). In blending theory, multiple mental spaces can participate in a mapping. These input spaces project onto a separate blended space, yielding a new emergent meaning structure, which is to some extent novel or distinct from meanings provided by each input space.

Consider the familiar metaphor *Surgeons are butchers* (see Grady et al., 1999). One may argue that this metaphor, like all others, is explained in terms of the projection of information from the source domain of butchery to the target domain of surgery. Yet this mapping alone does not provide a crucial element of our interpretation of this metaphorical statement—namely, that the surgeon is incompetent. After all, butchers can indeed be as skilled at their job as surgeons are at theirs. Under a blending theory account, metaphors meaning is cap-
tured by a blended space that inherits some structure from each of the input spaces. Thus, from the target input space for surgery, it inherits elements such as of a person being operated on, the identity of the person who is doing the operation, and the place where this all happens. The source domain butchery input space inherits information such as what a butcher does and his relevant activities such as using sharp instruments to slice up meat. Besides inheriting partial structure from each input space, the blend develops emergent content of its own, which arises from the juxtaposition of elements from the inputs. Specifically, the butchery space projects a means–end relationship that is incompatible with the means–end relationship in the surgery space. For instance, the goal of butchery is to kill the animal and sever the flesh from its bones, but surgeons aim to heal their patients. This incongruity of the butcher’s means with the surgeon’s end leads to an emergent inference that the surgeon is incompetent.

Proponents of blending theory see it as a highly generalizable tool that can explain a broad range of linguistic and cognitive phenomena (Grady, Oakley, & Coulson, 1999). The theory is capable of explaining not only metaphor, but other types of cognitive activity, including inference and emergence of many kinds of linguistic meaning. Blending theory extends conceptual metaphor theory by allowing for mappings that are not directional between multiple domains. Furthermore, blending theory may capture aspects of online meaning construction better than entrenched conceptual structure (i.e., conceptual metaphor theory). Thus, metaphoric interpretations of novel poetic figures are constructed on the fly, emerging from blended spaces and not from the input spaces alone, nor from some additive space of what two or more domains have in common (i.e., the generic space).

A different development from linguistics is the work on relevance theory (Sperber & Wilson, 1986). The relevance-theoretic account of utterance interpretation proposes that a fundamental assumption about human cognition is that people pay attention to information that seems most relevant to them. Every utterance starts out as a request for someone else’s attention, and this creates an expectation of relevance. This expectation of relevance provides the criterion for evaluating possible interpretations of a speaker’s utterance. Relevance is defined in terms of contextual effects and processing effort. Contextual effects are achieved when a speaker’s utterance strengthens, contradicts, or denies an existing assumption or by combining an existing assumption to yield some new contextual implications. Sperber and Wilson claimed that newly presented information is relevant in a context only when it achieves contextual effects in that context, and the greater contextual effects, the greater the relevance.
Relevance theory suggests that metaphors and other figures of speech are examples of loose talk (Sperber & Wilson, 1985/1986). Consider the utterance *My neighbor is a dragon* (Blakemore, 1992). Speaking loosely like this requires that speakers have in mind some further idea or contextual implication beyond the single thought “My neighbor is fierce and unfriendly.” For instance, the speaker might wish to convey an image of fierceness or unfriendliness that is beyond most people’s experience and will expect the listener to put some effort toward exploring a wide range of contextual implications (e.g., having to do with the nature of the neighbor’s unfriendliness, the behavior it manifests, and perhaps the neighbor’s appearance). These implications are relatively weak, but they best resemble the speaker’s thought about the neighbor. The indirect nature of metaphor calls for extra processing effort on the listener’s part, but this is offset, according to the principle of optimal relevance, by extra effects not achievable by saying directly that *My neighbor is fierce and unfriendly*.

In general, nonliteral speech acts are simply means of optimizing relevance in verbal communication. The metaphor as loose talk view does not assume that metaphor requires special cognitive processes to be understood. Relevance theory offers a significant advance over the Gricean view of utterance interpretation, especially to the degree that it seeks to tie together pragmatic language use and cognition. Only a few experimental studies have directly tested the psychological plausibility of relevance theory as an account of nonliteral speech act understanding (Gibbs, 1986b; Jorgensen et al., 1984). Yet relevance theory holds much promise and should clearly be the focus on additional empirical research.

The intuitive promise of these new developments on indirect and figurative discourse is somewhat tempered by the difficulty in formulating potentially falsifiable hypotheses from these linguistic/philosophical theories. Yet empirically minded psychologists, as well as the scholars responsible for these new ideas, would do well to carefully examine the possible implications of both blending and relevance theories for their respective suggestions for how people use and understand nonliteral speech acts. For instance, both conceptual blending and relevance theories propose that certain information is likely to be more relevant or appropriate in different metaphorical mappings. Both theories implicitly assume something about information that is most salient and what information must be suppressed when people construct meaningful interpretations of poetic figures. The recent suppression and graded salience hypotheses provide possible constraints on these other theories. Blending and relevance theory provide larger theoretical frameworks for evaluating the plausibility of the newer psy-
chological proposals. Clearly, there is much room for the cross-fertilization of theory and data between the newer work in psychology, linguistics, and philosophy.

CONCLUDING REMARKS

The study of nonliteral speech acts in text and discourse provides several different opportunities to explore the intimate relationship between thought and language. My review of some recent debates over nonliteral speech acts illustrates the importance of pragmatic and conceptual knowledge in how people produce and understand what speakers say and what they imply in discourse. Furthermore, characterizing the psychological processes that underlie nonliteral speech act use requires scholars to recognize how different research methods tap into different aspects of what occurs when people produce, make sense of, immediately comprehend, and consciously interpret what speakers and writers aim to communicate when they imply something that varies from what they say. The time is right now for scholars in all fields interested in discourse processes to accept some of the main challenges identified in this chapter, especially those emphasizing the link of pragmatics and conceptual knowledge to nonliteral discourse. One likely result of the new scholarship that results from future empirical studies is the conclusion that nonliteral speech acts are fundamental reflections of typical cognitive processes.

REFERENCES


This page intentionally left blank
Language defines our lives. It heralds our existence, it formulates our thoughts, it enables all we are and have.
—Steven R. Fischer (1999)

But what is language divorced from discourse?

This chapter addresses developmental issues with respect to discourse. Hence, in addition to the definition of discourse, the issue of what entails development assumes importance here. Discourse is broadly taken to mean the use of language beyond that of a single sentence. Therefore, it encompasses conversation, the different genres of extended discourse (e.g., narrative, argumentation, and rhetoric), as well as monologues in which young children engage. Of course it is beyond the scope of this chapter to deal with the development of all these different genres. Therefore, only the principal trends that have a bearing on developmental concerns are explicated.

The definition of development is somewhat more controversial. In line with the process orientation of this volume, our focus on development is less on the mapping of the stages of the development of discourse abilities and their temporal sequence than on the processes that enable these abilities to emerge and develop. In addition, we also highlight the idea that development is always a theoretical lens and, as such, to a large extent, in the eye of the researcher (Bamberg, 2000a, 2000b). As Kaplan (1983) put it, “development does not lurk directly
in the population(s) studied but resides fundamentally in the perspective used" (p. 196). That is, any development is but a particular lens used to make sense of constancies and changes that can be observed in historical time and, with a change of that lens, the version of development undergoes a concomitant change. The question then arises as to which of these different developments should be taken as the paradigmatic one. The simple answer is that primacy changes with the tenor of the research environment. However, through a consideration of a number of different alternatives, we attempt to show that at present the most productive way to view discourse development is to consider it as a process with an active individual at the center, but one that is significantly influenced and stimulated within the linguistic and cultural environment and whose development is advanced through participation in language practices.

We first begin with a brief overview of established research findings concerning the development of conversational skills and extended discourse. Next, and consonant with the narrative turn in psychology and psycholinguistics, the main current paradigms of research into narrative are considered. The ground is thus prepared for a dissolution of the artificial dichotomy that has held between conversation and extended discourse (e.g., Ninio & Snow, 1996; Pan & Snow, 1999), here principally represented by narrative, leading to a perspective in which they each feed into and support the development of each other, an innovation based mainly on the work of Levy (1989, 1999; Levy & Nelson, 1994). Following this synthesis, discourse development is considered as a holistic entity that contributes to the development of the identities of individuals who participate in discourse practices in situated contexts. The chapter concludes with a summary of the main concerns of contemporary research on discourse development and future avenues of inquiry.

CONCISE DEVELOPMENTAL OVERVIEW

Conversation

To participate in conversation at the level of a mature adult conversationalist, the child has to master a myriad of interactional skills that go far beyond knowledge of phonology, semantics, and syntax (Ninio & Snow, 1996). These include learning how to take appropriate turns, avoiding culturally defined levels of overlaps and interruptions, being able to respond appropriately, discerning the role of the listener, and cooperating in the maintenance and flow of topics. In addition, children have to eventually learn some cultural variation of the Gricean
maxims of conversation (Grice, 1975). For example, they must discover that in English-speaking, Western cultures, participants in a conversation are usually cooperating with each other, providing information that they believe to be true, limiting the information that they provide to the needs of the situation, ensuring that what they say is relevant to the topic at hand, and avoiding obscure, ambiguous, and verbose utterances. On another level, children must learn and respond to the different communicative demands that characterize interactional conversations as opposed to transactional conversations. As defined by Brown and Yule (1983), interactional conversations are those in which the prime objective is to maintain friendly social relations, whereas transactional conversations are primarily concerned with a transfer of information flow between the two (or more) participating parties. It is only recently that the different communicative demands inherent to these different conversational types have begun to be researched (Anderson, Clark, & Mullin, 1994). This brief overview thus first describes children’s developmental processes toward being competent interactional conversationalists and then compares this to their efforts at transactional conversations.

**Turn Taking.** When children are observed in interactional conversations with adults, it appears that they are expert turn takers from an early age. They are able, or seem to be able, to respond to the adults’ utterances with a smooth and effortless automaticity. The reason for this is that children are participants in the interactional mechanics of conversation from infancy forward. Whiten (1977) found evidence of turn taking and reciprocity in infant–mother dyadic interaction in which the infant’s role is compensated for and supported by the mother. In such dyadic interactions, it has also been found that the infant either tries to take the lead (Trevarthen, 1979) or is periodically able to set the pace while the caregiver provides the replies (Schaffer, Collis, & Parsons, 1977). Indeed, by the age of 10 months, infants have been found to produce protoimperatives as well as protodeclaratives, which presage requests and statements, respectively, and mothers respond to these as if they are already legitimate communicative acts (Bates, Camaioni, & Volterra, 1975). Tomasello (1999) theorized that it is the protodeclaratives that uniquely characterize human communicative abilities because these are arguably incipient attempts to make statements through which the infant initiates joint attention episodes with her interlocutor. It is from these joint attention situations that all further learning and communication are considered to develop.

As the child acquires more linguistic skills, it remains clear that caregivers continuously scaffold their charges into the turn-taking
conventions of normal conversation (Bruner, 1977, 1983). For example, Snow (1977) found that caregivers structure conversational responses around the contributions of infants, that parents of 21-month-olds provide frames into which children can insert suitable responses (Söderbergh, 1974), and that when necessary caregivers speak for the child or even answer their own questions to keep the interaction going (Martinez, 1987). The importance of the role of adults in the socialization of children into normal conversational turn-taking practices is highlighted when such scaffolding becomes unavailable (i.e., when young children talk to their peers). Garvey and Berninger (1981) found that in such peer conversations, there were longer between-turn gaps and fewer anticipatory overlaps, whereas Ervin-Tripp (1974) reported that children’s peer conversations were characterized by disorderly turn taking in comparison with adult conventions. However, by age 4, children have been found to be able to use the sentence initial and as well as and then as floor holders to signal that their turn is not yet complete.

Development of Initiations. Schober-Peterson and Johnson (1991) found that children are rather poor at initiating conversations with peers, their attempts often failing so that monologues instead of dialogues result. Although initiating a conversation may seem extremely difficult for children, there is evidence that children begin to develop this ability from early on. Ochs and Schieffelin (1976) identified that the two most important abilities required to initiate a conversation are attention-getting and attention-directing devices. There is evidence to show that children use attention-getting devices, both verbal and nonverbal, from the age of 2 (Wellman & Lempers, 1977), and that this usage parallels the development of a child’s communicative abilities (Mueller, 1972). The development of attention-directing devices also follows a similar developmental pathway, originating in nonverbal gestures and vocalizations and leading to the use of single words such as look in the second year (Carter, 1978), proceeding through the use of common nouns (Atkinson, 1979) and progressing to the use of relative clauses to refer to nonpresent objects by age 4 (McTear, 1985). In summary, it is in the preschool period that children begin to develop the ability to get attention to initiate a conversation, identify discourse referents, take into account their interlocutor’s state of knowledge, and secure an appropriate response (McTear, 1985). Nevertheless, these skills continue to develop well after this age range and into adulthood.

Development of Responses. To respond appropriately to an interlocutor’s utterances, a child must learn to contingently relate her own utterance to what has just been said to her. This ability does not seem
apparent at 21 months (Bloom, Rocissano, & Hood, 1976), although Horgan (1978) has reported that 17-month-olds are able to differentiate between yes/no and wh-questions. The latter notwithstanding, Steffenson (1978) found that the responses of 2-year-olds are often inappropriate from the perspective of a mature conversationalist. Evidence of children's developing knowledge of different response contingencies was provided by Olsen-Fulero and Conforti (1983), who found that between the ages of 29 and 36 months children provided appropriate responses to maternal questions 56.7% of the time compared with 16.7% for declarative statements, and they were more likely to respond to repair questions than to tag questions. Indeed, the ability to respond to wh-questions has been found to develop between the ages of 1 and 6 (Cairns & Hsu, 1978). It appears that children first learn to respond to where and what and then only later to why, how, and when apparently due to the different conceptual demands of the questions. Where and what can conceivably be answered gesturally, whereas the same is not possible with when, how, and why. A skeletal, average, developmental mapping of children's increasing abilities to respond contingently to their interlocutors' utterances as a function of increasing age would be: by 3:0,1 imitative early responses (which are predominant until 2;9) are added to or modified with clause constituents, and anaphoric pronouns become increasingly used (Bloom et al., 1976; Keenan & Klein 1975); between 2;9 and 3;6, children increasingly use auxiliary ellipses and discourse connectives like then, but, and because (Ervin-Tripp, 1978).

**Conversational Repair.** As would be expected, young children are less adept at initiating repairs to conversational irregularities than adults. Observations that children initiate conversational repair start at 18 months (Johnson, 1980), although such attempts remain below the frequency of adult-initiated requests until at least 3 years (Gallagher, 1981; Ninio & Snow, 1996). With respect to peer interactional conversations, children rarely request clarification (McTear, 1985), but when they do they favor nonspecific clarification forms (Garvey, 1977). However, it has been found that children respond readily to adult queries (85% of the time) and can already differentiate between general and specific queries when as young as 20 to 42 months (Anselmi, Tomasello, & Acunzo, 1986).

**Sustaining Coherent Dialogue.** Related closely to both initiating and responding to conversational repair initiatives is the general ability to

---

1Children's ages are referred to by years, followed by months after the semicolon.
sustain a coherent dialogue. In a comprehensive review of children’s conversations, McTear (1985) distilled three important components of increasing complexity that contribute to this ability: to be able to (a) initiate and respond to simple topics, keeping exchanges fairly closed and avoiding the combination of these into larger topically or interactionally related sequences; (b) structure longer sequences of dialogue by linking exchanges through the ability to carry forward material and continuing on a linked topic, having implicitly acknowledged the response in the preceding exchange; and (c) respond to the preceding turn while setting up expectations for or providing the possibility of further response often by fitting more than one utterance into one turn.

From an empirical standpoint, Dorval and Eckerman (1984) described a series of differences in conversational skills in a wide age range from second graders through adults. They found that second graders produce more unrelated conversational turns in comparison with older children and adults. Children of this age also tend to respond as much to nonverbal actions as to what was said. By contrast, the conversations of fifth and ninth graders consist of a substantial proportion of factually related turns, and it is in this age range that they start to trade personal narratives. Further developments are evident in the more mature skills of 12th graders and adults, who engage in much more perspective taking and exchange, use more cohesive devices like conjuncts (for example, so, anyway), and employ attitudinal expressions (e.g., really, perhaps) to signal stances. Lest the younger child be viewed as an extremely inept conversationalist, there is also research to show that, by the age of 4, children have control of about one dozen speech acts including requests for action, refusals and agreements to comply to others’ requests, and statements of propositions and statements of intent, being able to display these within the context of negotiating an ongoing activity or discussing an object in the immediate environment (Snow, Pan, Imbens-Bailey, & Herman, 1996).

Using an experimental paradigm to investigate individual differences in children’s conversational skills, Schley and Snow (1992) put children in a talk-show situation with the child as interviewer and an adult as the interviewee who was instructed to be politely responsive but not helpful by not elaborating on responses. Results indicate that the task was highly difficult for children, and many children ages 7 to 12 could not complete the allocated time of 4 minutes. More important, children who were independently rated more highly as conversational partners used more open-ended questions and questions that were contingent on previous utterances, avoided long lapses in conversation, and elicited elaborated responses from their partners.
**Interactional Versus Transactional Dialogue.** The repeatedly observed discrepancy between children’s performance within adult–child conversations and children’s unaided conversational and/or communicative skills can be generally explained by appealing to the scaffolding activities that adults provide to children in the effort to socialize them into an important means of social intercourse as conversation. There are several forms of scaffolding. Adults persistently try to keep the conversation going in the face of children’s inadequate responses (Shatz, 1983; Snow, 1977). They also use devices such as questions and requests for clarification, which mainly function to keep up the interaction rather than effect an exchange of information (Demetras, Post, & Snow, 1986; Olsen-Fulero, 1982). Furthermore, adults have a tolerance of underspecified or ambiguous messages as long as the conversational flow is not obstructed; they also have a dispreference for overt corrections of the child by a more expert listener not to obstruct communication (McTear, 1987). Clearly, then, where the transfer of information is not the main goal, the child is trained to be a skilled conversationalist from early on, mastering the relevant conversational conventions of her culture.

However, there is a form of dialogue that focuses on the transfer of information between the parties engaging in it. This has been termed transactional dialogue or discourse, to use Brown and Yule’s (1983) original term, which contrasts with interactional conversation or discourse, the main objective of which is the maintenance of friendly social relationships between its participants. Anderson, Clark, and Mullin (1994) investigated children’s communication skills within the framework of transactional dialogue. They instructed children between the ages of 7 to 13 to communicate about a route on a map that had slight variations in the versions used by each child in the conversational pairs. They allowed free verbal interaction between the pairs of children, excluded visual comparisons, and studied how young speakers collaborated during dialogues. On average, the 7-year-olds communicated less effectively than the 9- and 13-year-olds, whose mean performances were similar. However, there was considerable variation within age groups, with some 13-year-olds communicating at the level of the youngest participants. Questions that were answered by an interlocutor were shown to be the most successful topic introducers and characterized the more successful communications. Within the latter, more problem points were discovered and more explanations made. Less successful dialogues were ones in which instruction givers as well as listeners concentrated rigidly on their assigned roles, the former not being sensitive to the feedback of the latter, whereas the listeners did not assume joint responsibility for the establishment of a common
knowledge base by questioning disparate information and actively seeking compromises.

Interestingly, success was not found to be closely related to age. In interpreting this apparently nonintuitive finding, the researchers appealed to Robinson's (1981) suggestion that parental styles of conversational feedback may influence children's effectiveness as communicators. Children whose parents provide feedback that focuses on message quality may become more effective communicators. Anderson et al. further asserted that their research highlights the situational and contextual dependency of conversational learning. Children first learn the conversational mechanics of conversational maintenance because of parental input before going on to master the informational-transference possibilities of conversations. As is consistently demonstrated in this chapter, this is consonant with a viewpoint that the development of discourse skills is a contextually situated one, albeit with the developing novice as an active processor of the linguistic and cultural input impinging on her. That this applies to the realm of conversational development should already be apparent from the evidence reviewed.

Children's Conversations Beyond the Primary Care Dyad. As previously described for Western cultures, primary caregivers, especially mothers, scaffold their children into the conversational conventions of their culture. How then do children learn to talk to other interlocutors, especially those outside their families who might not be so helpful in structuring conversations supportively around inexperienced and less skilled novices? It was Berko-Gleason (1975) who first suggested that fathers play an important role in this respect by providing the child with a greater challenge as a conversation partner because, by and large, fathers spend less time with their children and so accumulate less knowledge about their children's experiences, competencies, and idiosyncrasies. The result of this is that fathers may be less accommodating as conversation partners, thus forcing their children to take into account a less supportive point of view. However, fathers are probably more familiar with their children than persons who do not belong to the immediate family and so they represent an intermediate step in children's development of skills for conversing with strangers.

Subsequent research with children up to the age of 3;6 has found support for Berko-Gleason's hypothesis. For example, Weist and Kruppe (1977) found that fathers are more prone to misunderstand what their children say to them than mothers and also seem to interrupt their children more (Grief, 1980). In addition, fathers have been found to use words that their children have encountered more infre-
quent or not at all, as well as words that their children seem not to understand (Fash & Madison, 1981). Another indicator that fathers use more complex words when speaking to their children is the finding that fathers use more *wh*-questions than mothers (McLaughlin, White, McDevitt, & Raskin, 1983). There has also been evidence that the speech of fathers to children, compared with that of mothers, is more controlling by containing more direct requests and imperatives (Malone & Guy, 1982). Nevertheless, a more recent study by Davidson and Snow (1996) did not find the same differential characteristics of fathers’ speech to their 5-year-old children. However, they acknowledge that their sample contained children who were atypically linguistically advanced, and that specific paternal contributions to the development of the children’s linguistic abilities could already have been made at an earlier age. They also recommend that a longitudinal study that starts at age 2 and continues through the kindergarten years of 4 and 5 be the tool of choice to disambiguate the effects of paternal speech on children’s discourse development.

Building on the Father Bridge Hypothesis, Mannle and Tomasello (1987) advanced the idea that siblings also have a role to play in this process of a child’s growing ability to skillfully converse with a wider range of people. Termed the *Sibling Bridge Hypothesis*, this would be due to siblings’ own limited abilities as conversationalists as well as their conscious or unconscious reluctance to provide the level of scaffolding that mothers provide. In the case of siblings, however, the bridge for the child is conceivably to peer conversations. In the latter case, it has been consistently found that incipient attempts at peer conversational interaction are aided by the presence of physical contexts that allow the sharing of knowledge and familiar events (French, Boynton, & Hodges, 1991). In this way, relative novices at conversation are presented with a reduced burden of intersubjective perspective taking as well as a lower need for shared presuppositions.

More recent research on children’s developing conversational abilities has focused on ever more specific issues. These include, for example, the different conversational patterns between children and their disabled and nondisabled siblings (Summers, Hahs, & Summers, 1997) and comparisons of the linguistic skills of twins to those of singletons within triadic contexts (Tremblay-Leveau, LeClerc, & Nadel, 1999). In short, it was found that children were more sensitive when talking to their nondisabled siblings, and the type of disability played a role in the level of sensitivity displayed (Summers et al., 1997). In the case of twins, Tremblay et al. (1999) found evidence that their linguistic skills were superior to singletons, especially with respect to re-initiating participation after being excluded from an interaction.
Although narrative is but one of the different forms of extended discourse, for several reasons there has been a conscious choice to use an exposition of the development of narrative skills to exemplify the development of the ability to construct extended discourse. First, it should be obvious that the principles inherent in the development of narrative ability are, in large part, common to the development of other forms of extended discourse such as scripts, explanations, and definitions (Ninio & Snow, 1996). Second, it is arguable that narrative is the most important genre of extended discourse because it is instrumental in the socialization of novices into the ways of being in their respective cultures (Bruner, 1990; Ochs, 1997; Polkinghorne, 1988), and it continues to function as a major medium for social meaning making and self- or other representation (Coupland & Nussbaum, 1993; Harré, 1998; Harré & Gillett, 1994; Kerby, 1991; Michael, 1996; Ochs, 1994). Third, mastering narrative is a particularly important achievement for children because not only must they relate events that are detached from the here and now, but also in doing so they must entwine two landscapes of behavior: (a) circumstances of the situation and actions of the protagonists therein, and (b) mental states of the protagonists including affect, beliefs, attitudes, and evaluations (Bruner, 1986). Finally, narrative is a form of discourse that is being found to be increasingly important in the building of children’s literacy skills (Michaels, 1981, 1991; Peterson, Jesso, & McCabe, 1999).

Although Labov and Waletzky (1967/1997) first defined narrative as extended discourse with at least two clauses encapsulating a temporal juncture between them (i.e., an irreversible sequence of events in time), since then this definition has been relaxed to encompass children’s narratives as discourse with at least two functionally related forms (e.g., Levy, 1999; Nelson, 1989). In addition, in work on narratives in different cultures, it became apparent that some of them do not emphasize the inclusion of sequential events in time to the same degree as others, as in narratives in the Latino tradition (Rodino, Gimbert, Perez, Craddock-Willis, & McCabe, 1991), thereby further contributing to a modification of Labov and Waletzky’s original definition. Indeed McCabe (1991, 1996) relaxed the definition of narrative so far as to take it to be a linguistic cross-road of culture, cognition, and emotion that serves the dual function of sense making and self-representation.

There have been many studies of the developmental trajectories of children’s narrative abilities since Labov and Waletzky (1967/1997) highlighted their utility as a sociolinguistic tool. Among them are
Peterson and McCabe’s (1983) extensive study of the personal experience narratives of children, Bamberg’s (1987) psycholinguistic analysis of children’s developing abilities of maintaining coherence in narrative, Bamberg and Damrad-Frye’s (1991) study of the evaluative dimension of children’s narratives, Reilly’s (1992) focus on the para-linguistic dimensions of children’s narratives, Hickman’s (1995) investigation into the development of the expression of reference, and Berman and Slobin’s (1994) large-scale cross-linguistic study of the development of linguistic forms together with their form–function relations in the narratives of children compared with adults across English, German, Hebrew, Spanish, and Turkish. In this brief overview, we focus on two of the developmental frameworks of children’s burgeoning narrative abilities and defer a consideration of other important paradigms of narrative development research until a later section. One framework is the temporal sequencing of children’s abilities in relating personal narratives (McCabe & Peterson, 1991; Peterson & McCabe, 1983). The second is a synthesis of Bamberg’s and Reilly’s work on children’s developing abilities to employ evaluative comments and affective expressions in narratives (Bamberg & Reilly, 1996).

**Development of Children’s Personal Narratives.** For the purpose of developing a temporal sequence of the development of children’s personal narratives, McCabe and Peterson (1991; Peterson & McCabe, 1983) drew on two data sets of middle- and working-class North American children of European origin. One data set consisted of 96 children ages 4 to 9 years who were primarily of working-class background studied cross-sectionally. This was supplemented by data from 3½-year-old children from a longitudinal study of children ages 2 to 4 years. In this research, McCabe and Peterson made use of Labovian highpoint analysis. This analysis assesses how well a narrative compares to the classic structure of having orienting information, a complicating factor(s), a climax or high point, a resolution, and possibly a coda (Labov & Waletzky, 1967/1997).

Their data indicated that most children at 3½ years produced two-event narratives. In a previous study, Sachs (1979) had shown that children begin to narrate at about 2 years of age, but are confined to the narration of one event at a time. Two-event narratives consist of either two propositions about an event or a recounting of two events. At the age of 4 years, children were found to most commonly produce a narrative type that the researchers termed *leapfrog* narratives. In such narratives, children narrate more than two events, although they often jumble up the order of the events as well as omit events or details that put themselves in a negative light. These leapfrog narratives become
much less common by the age of 5, when children mostly tell personal narratives that end at the high point of the story—so-called *end-at-high-point* narratives. In these narratives, children are able to adequately sequence the events temporally, but they do not bring the story to a resolution, ending it at the climax.

By the age of 6, children were able to relate narratives in the classic way that seems to be preferred by the culture of Euro-American origin. They begin by setting the scene with a statement of what the narrative is about, relate a series of events that lead to a high point, provide some evaluative comments on the proceedings, and continue on to bring the story to a resolution. Sometimes they even provide a coda, which relates the story to their present circumstances. Developments between the ages of 6 and 9 center on the lengthening of narratives and the provision of more orienting information at the beginning of the narratives.

As can be expected, there are individual differences in this general developmental trajectory: A small minority (3%) of children as young as 3½ can generate classic narratives, whereas some of the older children between the ages of 7 and 9 keep on telling end-at-high-point narratives. Nevertheless, the developmental trend is clear. Classic narratives are the most common narratives told after the age of 6, and there are no more leapfrog narratives at 7 years and above.

**Evaluation and Affective Expression in Elicited Narratives.** In an attempt to explore how young narrators learn to tie the referential plane of story content and the activity of story performance into a bounded (integrated) structural unit that considers the perspective of a generalized audience, Bamberg and Reilly (1996) combined the data and analyses of two studies in which children from the ages of 3 to 11 were asked to relate the story “Frog, Where are You” (Mayer, 1969) from a 24-page wordless picture book. Of prime interest was how young children learn to integrate the emotional significance of a sequence of events for a (third-person) story character and the way this content is presented (i.e., the actual emotional/affective performance of the telling). From their combined findings, they proposed a three-phase developmental picture of narrative ability.

In the *Early phase*, which spans the ages of 3 to 6 years, children focus on the communication of the affective message of the story they are telling; in doing so, they liberally draw on paralinguistic resources to make up for their lack of linguistic sophistication. They treat the storytelling activity as a holistic unit and are unable to provide much horizontal cohesion nor vertical, hierarchical coherence in their organization of the story. A *Middle phase* can be discerned beginning around
the age of 6, in which the children start to concentrate on the establishment of horizontal intersentential cohesion. The global, affective whole of the Early phase gives way to a focus on more local-level details of content organization. In this way, children start to emphasize the connections between the story events and the internal states of the protagonists (i.e., their emotions and motivations). They lose sight of an overarching perspective of the story as a whole, instead concentrating on their newer abilities to use language to link emotions and events on a local level.

Finally, at about age 9, a Final phase of narrative ability develops in which the concerns of the two earlier phases become integrated into one story-telling activity/performance. At this point, the use of references to emotions, as well as the expression of affect, are beginning to orient the audience in a more integrated fashion toward the story and its narrator as a unit. As such, both forms of presentation—namely, establishing a referential plane of what is depicted as emotionally and motivationally relevant, as well as the activity of the story performance by means of affective expression—are functionally united. In a more general sense, children gradually learn how to choose linguistic forms that instruct their listeners to take the perspective of the mind of the protagonist whose life events they are depicting. As a consequence of this process, something like a protagonist is called into existence. The units of narration that constitute this existence do not preexist outside or prior to the narrative or interactive situation, but emerge within the interaction as part of the narrator–audience involvement.

Specifically with regard to the expression and performance of emotions within the story-telling activity, it was found that references to emotions were strongly tied to their feelings and display. That is, when emotions were discursively referred to, the narrator also seemed to be experiencing and outwardly manifesting them. Consonant with the analysis of Harré and Gillett (1994), both modes were used to discursively express judgments and perform social acts.

After our relatively brief and general consideration of how children develop the abilities to engage in both conversation and narrative, we now move on to some specific contemporary paradigms of narrative research.

CONTEMPORARY RESEARCH PARADIGMS ON NARRATIVE

Contemporary research on narrative is so rich and varied that an effort has to be made to select and group the many research efforts into a manageable number of foci, each having a particular direction in common and characterized by a number of researchers who can be consid-
pered typical of that particular approach. In such an effort, there are naturally some omissions, and for these we must claim the defense of unintentionality. We have been able to identify six main contemporary approaches to narrative research and have decided to order them according to the closeness with which they adhere to strictly linguistic material in their analyses. Therefore, in the order of the use of increasing amounts of nonlinguistic factors for interpretive processes, the different approaches are story grammars, psycholinguistic perspectives, constructivist analyses, an interactive approach, a combinatorial/eclectic viewpoint, and symbolic/cultural considerations.

**Story Grammars**

This approach most closely relates to formal approaches to linguistic analysis in that its focus is on the specification of the most elementary structural components that make up narratives and the grammar of the organization of these units that contributes to the coherence of the narratives. The relation of these to the cognitive processes of the narrator, whether conscious or unconscious, is also of prime importance. This has led to the search for different frameworks onto which the structural configurations of narratives can be mapped, a few examples of which are hierarchically organized goals and events (Black & Bower, 1980), units of plot sequence (Lehnert, 1981), and networks of causal connections (Trabasso & van den Broek, 1985).

Although this approach maintains its traditional focus on both children's and adults' performance on the comprehension, recall and judgments of texts (Graesser, Singer, & Trabasso, 1994; Langston & Trabasso, 1999; Singer, Graesser, & Trabasso, 1994; Trabasso & Magliano, 1996; Trabasso, Suh, Payton, & Jain, 1995), more recent research has been oriented toward the determinants that contribute to coherence in the production of narratives (Stein & Albro, 1997; Trabasso & Stein, 1994, 1997). Common threads that connect these recent research efforts include the use of the knowledge of goals and plans to structure and interpret narratives, role of human agency and intentionality, assessment of evaluations of events in terms of a person's well-being, comparisons of evaluations occurring within narratives and conversations, role of future-oriented thinking and planning in the construction of an emotional event, and memory and inference processes in the construction of a mental representation of texts.

The typical methodology used for the generation of narratives is the employment of story stems that have to be completed to the best of each participant's abilities. Developmental analyses indicate that children as young as 3 years can be credited with rudimentary story-telling
abilities, with subsequent development revolving around the integration of more complex aspects of narrative—essentially a process of quantitative refinement. The only qualitative development is considered to occur in early adulthood when the ability to judge a narrative’s *goodness* emerges.

Thus, the generation of narratives according to this approach is akin to the learning of rules that are to be followed so that narratives, and later good narratives, can be understood, produced, and appreciated. The person in general, and the child in particular, are viewed as basically active and rational—as logical seekers and organizers of information who use cognitive schemata in a goal-oriented way to make decisions in the behavioral domain. Such processes are deemed to be universally valid before they become specified by cultural, social, and historical contexts.

**Psycholinguistic Perspectives**

Taking its impetus from functional linguistics, the psycholinguistic perspective in narrative research is primarily represented by the work of Slobin and his associates (Aksu-Koç, 1991; Bamberg, 1987; Berman, 1995, 1996; Berman & Slobin, 1994; Marchman, 1989; Slobin, 1990). The central concern of this approach is to discover the linguistic forms and their corresponding functions that serve to structure narratives. These include achievement of overall coherence, establishment of the causal and temporal sequence of events, and management of foreground–background relations.

Considering the large cross-linguistic research effort in five languages (English, German, Spanish, Hebrew, Turkish) by Slobin and his colleagues (and reported in Berman & Slobin, 1994) as archetypal, this perspective is guided in large part by three general themes (Berman & Slobin, 1994):

1) **Filtering**: The world does not present “events” to be encoded in language. Rather, experiences are filtered (a) through choice of perspective, and (b) through the set of options provided by the particular language into verbalized events.

2) **Packaging**: A skillful narrative does not simply consist of a linear chain of successive events located in time and space. Rather, events must be packaged into hierarchical constructions.

3) **Development**: Younger children take fewer expressive options because (a) cognitively, they cannot conceive of the full range of encodable perspectives; (b) communicatively, they cannot fully assess the listener’s
viewpoint; and (c) linguistically, they do not command the full range of formal devices. It is thus the task of the researcher to discover through a fine-grained analysis how linguistic forms link up with functions in the child’s construction of narratives, which is also taken to be representative of her general acquisition of language. (pp. 9–15)

Researchers working within this paradigm have used the wordless picture book “Frog, Where are You?” (Mayer, 1969) across all the studies for the elicitation of narratives. Typical developmental findings include dynamic form–function relationships where old forms acquire new functions before new functions stimulate the acquisition of new forms. Development toward being a proficient speaker is protracted, a proficient speaker being defined as one who has mastered the demands of achieving textual cohesion. As a child acquires a native language, she simultaneously learns particular ways of thinking for speaking, which relate to the fit between the conceptualization of events and the linguistic means afforded by the particular language to encode these.

In emphasizing why the process of integrating linguistic form and narrative/discourse is so lengthy, Berman (1997) appealed to the heavy cognitive demands placed on children in the mapping of appropriate linguistic forms onto the three main narrative functions of eventivity, interpretation, and informativeness. It is such complexity that continues to provide researchers in this domain with challenges.

Constructivist Analyses

A constructivist approach to narrative development is essentially a linguistic one, but one in which the referential function of language is sidelined in favor of its performative or practice function. In essence, such analyses investigate how the child learns to recruit different language forms to build narratives that provide her with knowledge of experience, self, and other. This is achieved not in the mind, but in language practices within communicative situations. In this sense, the constructivist approach can be considered an expansion of social constructionism (e.g., Gergen, 1995; Harré & Gillett, 1994), historically performed, with language delivering the building blocks and narrating as one of the central activities in which these blocks are put together so that experiences and selves can come into existence. The child is thus viewed as intensely active, participating in language practices that enable her to appropriate language forms for the construction of agency and perspective taking. These form the presuppositions that affect the order of characters in space and time and that relate
them to one another in the form of a moral order, albeit one locally constructed in a communicative setting. As such, the person is constrained by the linguistic habits and practices in which she participates and is not able to create life individually. The changes across childhood and different situations are described in terms of changes in linguistic construction types and how they are put to use for different discursive functions.

These aspects of language development are equally important as coordinations of forms and functions for the expression of perspective and agency, which seem to be more closely related to the construction of subjectivity and intersubjectivity. The impetus for the child to engage in narrating is taken to be her participation in any culturally accepted language practices. Almost any form of verbal data is considered appropriate for analysis, spanning the gamut from elicited stories built on a sequence of pictures, personal experience narratives, and explanatory discourse to one-clause answers to questions that ask for happenings or events. Consequently, this approach cannot properly address issues of when (i.e., at what age) and under what conditions children typically perform particular narrative tasks. Rather the analyses map the process of how linguistic forms are appropriated for discursive functions and within this process reveal the emergence of a moral self. The analysis is situated within what is commonly called variation analysis (Schiffrin, 1994).

This particular discourse analytic procedure starts from the assumption that speakers have choices in construction types and that their actual choices are signposts to how they want to be understood. Good examples of the application of this approach can be found in the work of Bamberg and his associates (Bamberg, 1997a, 1997b, 2000a; Bamberg & Damrad-Frye, 1991; Bamberg & Reilly, 1996) and also in the analysis of autobiographical narratives by Quigley (2000).

An Interactive Approach

An interactive approach to the study of narrative development is exemplified by the work of Quasthoff and her associates (e.g., Hausendorf & Quasthoff, 1992, 1996; Nikolaus, Quasthoff, & Repp, 1984; Quasthoff, 1986, 1995, 1997) in which the communicative dyad in which narration takes place is taken as the principal unit of analysis. The child and her more experienced interlocutor are subsumed as part of the interactive situation as a whole. Thus, although development can be viewed as taking place in the child as she gains communicative competence, it is more properly viewed as one part of a holistic development of the story telling in conversation and not as change in the per-
son per se. This approach microanalytically maps out how participants in an interaction contribute their different narrative competencies to the generation of a story-telling situation. This would include how a participant gains the floor, maintains the \textit{tellability} of her narrative and reciprocally, and how listeners utilize comments and interjections to either contribute to the maintenance of the activity or signal its close.

Originally based within a conversation analytic tradition, this approach has expanded the boundaries of analysis by relegating the cognitive and linguistic processes involved in narrating to the service of the overarching structure of how the activity of narrating is interactively organized. Of prime importance here is the sequential arrangement of the interactions.

Although the interactive situation in early parent–child interactions is characterized as unbalanced (due to their different communicative competencies), with development this imbalance is gradually leveled. This approach culminates in the developmental claim that the same mechanisms that are constitutive of the situational achievement of narrating also hold as developmental mechanisms. The following patterns of an adult listener’s contributions to the successful interactive communicative task is illustrative of this approach (Hausendorf & Quasthoff, 1992; Quasthoff, 1997):

1. \textit{Demands}. The adult directs the child into the appropriate moves by the use of local and global sequential implications in an age-specific way.
2. \textit{Localization of global demands}. The adult establishes local implications that steer the child into the fulfillment of the global tasks.
3. \textit{Explication of global demands}. The adult explicates global implications that are normally obeyed without explication.
4. \textit{Demonstration}. The adult obeys his or her own sequential implications, thus providing a model for the child to make the appropriate move.
5. \textit{As-if treatment}. The adult treats a locally initiated move of the child as if it had the global relevance that it should have according to the rules of job fulfillment.
6. \textit{Attribution}. If the child does not behave according to the age-specific demand, the adult offers an age-specific account for the child’s refusal, thus assigning a specific developmental niveau to the child.
The analysis is thoroughly empirical, emphasizing the functions of local linguistic forms in the generation of global structures of narrating.

An Eclectic/Combinatorial Viewpoint

Returning to the work of Peterson and McCabe and their associates (e.g., Ely, Berko-Gleason, & McCabe, 1996; McCabe, 1991, 1996, 1997; McCabe & Peterson 1991; Minami & McCabe, 1991, 1995; Peterson, 1994; Peterson & Biggs, 1998; Peterson, Jesso, & McCabe, 1999; Peterson & McCabe, 1983), it is apparent that they ground their research in a variety of analytical systems and apply themselves to a wide range of research objectives. There are, however, some common strands to all the subdomains of their research.

First, they prefer factual, personal narratives because these are considered more original and, in addition, allow the study of the subjective meanings of reportable events. Next they focus on production instead of comprehension and eschew the use of either story stems or serial pictures to standardize elicitation of narratives across participants. They make the point that it is presumptuous to assume that the same stimulus will have the same meaning for different participants because each individual has had a different developmental history. Therefore, they use a method of eliciting personal narratives from children that is based on the establishment of a comfortable conversational situation between the researcher and the participant after which the researcher introduces events believed to possibly have been experienced by the participant. When the participant takes up the narration of a personally relevant event, the researcher then proceeds to provide nonspecific social support to try and keep the narrative going. In general, they try to collect at least three narratives and then take the longest one as indicative of the best performance of the child.

In their approach, they view the child as a budding novelist, learning to piece together past experiences in meaningful ways. They emphasize that they do not consider narratives as accurate memories, but rather linguistic representations in story form. It is with the analysis of their data that the eclectic nature of their approach shows through. They use a range of analytical methods, striving to find the best fit between methods and particular data. These methods include high-point analysis (Labov & Waletzky, 1967/1997), story grammars (e.g., Stein & Glenn, 1979), stanza analysis (e.g., Gee, 1991), and dependency analysis (Deese, 1984). Their work on the temporal development of narrative ability was described earlier. They have also been concerned
with the role of differential parental input in the development of children’s narrative styles (Peterson & McCabe, 1992, 1994), the results of which indicate that different parental styles, such as a focus on emplotment versus descriptive detail, are later mirrored in their children’s spontaneous narratives.

Their more recent work (Peterson et al., 1999) has shown that it is possible to alter parental models of narrative input to facilitate children’s learning of the canonical narrative styles highly valued in the American school system, with a consequent facilitation in the achievement of literacy by these children. Other concerns have been gender differences in narrative styles—specifically the use of larger amounts of reported speech in the narratives of girls (Ely, Berko-Gleason, & McCabe, 1996) and cross-cultural differences between the narratives of Japanese and American children (Minami & McCabe, 1991, 1995). This outline of their work demonstrates the breadth with which they have utilized and combined different theoretical paradigms in their investigation of narrative development.

**Symbolic/Cultural Considerations**

In this category, we include both the symbolic approach to narrative development as exemplified by the work of Nicolopoulou (1993, 1996, 1997; Nicolopoulou & Weintraub, 1998) as well as some cultural influences that shape preferred narrative styles. Basing her work on a sociocultural tradition, Nicolopoulou asserted that narratives must be treated as symbolic forms and vehicles of meaning that help children shape reality and interpret their experiences. Children’s narratives must be analyzed with respect to both their form and symbolic content. Narrative is considered to be a form of action and serves as a vehicle to determine what children use narrative for. Typically, children use stories to represent the world to themselves and each other. Thus, narratives figure importantly in children’s formation and maintenance of their identities. Nicolopoulou emphasized that in story telling, children appropriate and utilize cultural resources in their environment in an active and selective way. This represents a situated culture-dependent activity that would be best understood through sociocultural analyses. Finally, through drawing important parallels between play and narrative as symbolic activities, Nicolopoulou called for a closer investigation of the developmental relationships between the two.

In line with Nicolopoulou’s definition of *story telling* as a culture-dependent activity, there has been a realization that the classic Labovian high-point narrative structure is not universally preferred
(McCabe, 1996). For example, Minami and McCabe (1991, 1995) found that Japanese children’s narratives were generally more terse and had shorter, although more regular sized, turns than the classic narrative structure that is preferred by Euro-Americans. Yet McCabe (1997) highlighted how the narratives of African-American children tend to be longer than the classic model and also incorporate more than one experience in a single telling.

Along the same lines, Melzi (1997; Berko-Gleason & Melzi, 1997) compared American Anglophone and Latino mothers with respect to the narrative styles they encouraged their children to develop. She found that Latino mothers encouraged children to move freely within the narrative, expand across experiences, include others in their stories, and talk about more abstract details of experiences, such as the motivations of the participants. By contrast, American Anglophone mothers focused more on the model of the classic narrative, encouraging their children to develop a topic by building on a single main event.

Such research into cultural differences in the form of preferred narrative styles should shed light on why children from certain cultural groups find difficulty in adapting their story telling to the classic model preferred in schools with consequent negative effects on their learning of literacy skills (McCabe, 1996).

**CONVERSATION AND NARRATIVE: SYNTHESES AND DEVELOPMENTAL SYNERGY**

So far we have considered the development of conversation and narrative separately from each other. More recent work by Levy (1999) has begun to show how these two abilities develop in an intertwined fashion. This synthesis is based on a long-term investigation of the crib talk of a young girl, Emily, who was audiotaped during a 16-month period between the ages of 21 to 36 months (see also Levy, 1989; Levy & Nelson, 1994; Nelson, 1989, 1996; Nelson & Levy, 1987). Levy essentially depicted how Emily appropriates cohesive devices from her father’s speech (such as deictic pronouns) and practices them in her monologues before she goes to sleep. Through mastering the use of these deictic devices, her ability to sustain coherent discourse increases, and this feeds into subsequent conversational dialogue with her father. Hence, Levy demonstrates how the development of discourse skills, both in conversation and narrative, is a culturally situated activity that progresses through participation in language practices. Planning the units and segments of discourse does not necessarily come from a cognitive activity that takes place before the act of speaking (although it
can). Rather, discourse planning is a situated language practice in which language forms are arranged for discursive purposes.

Although coherence in discourse has been extensively studied (e.g., Clark, 1994; Gernsbacher & Givón, 1995), Levy’s account is new in being explicitly developmental and getting closer to the source of these abilities in the young child. Levy (1999) charted out three stages in the process of the development:

1. **Acquisition of the cohesive function of language:** Prior to the ability to make thematic statements, children first have to appropriate the referential indexical properties of deictic devices (e.g., pronouns and temporal adverbs). Children achieve this by borrowing these devices from adult speech heard in face-to-face interactions and then practicing them in their own complexively linked speech. These reproduced and rehearsed utterances then become context for subsequent speech and serve a function similar to the original social environment from which they were taken. In this way, the roots of cohesion can originate in interactions even substantially removed in time from when the child uses them in her own speech. The mechanism for such appropriation is one of contrastive alternation, in which the child substitutes cohesive/deictic devices for already known language forms in chains of speech linked by phonological or rhythmic patterns. The following are some examples:

   i don’t know which kind of lady bought the crib
   i don’t know what lady bought it
   the tree fell down
   tree fell down
   what tree fell down
   that one fell down
   that must fell down (Levy, 1989)

It is also made clear that the child borrows these devices not only from the spontaneous utterances of the adults in her environment, but also from culturally packaged sources that she hears such as stories, fairy tales, and nursery rhymes read to her.

2. **The emergence of thematic statements that gloss preceding discourse (anaphora):** When the child is able to link pairs of utterances in a tightly cohesive way, these are then available to be used as contexts for creation of even tighter cross-clausal links. Further development of such cross-clausal links results in the emergence of thematic statements that summarize a series or sequence of utterances that precedes them.
3. The emergence of thematic statements that anticipate subsequent discourse (cataphora): In developing thematic statements used to guide the further construction of discourse, the child begins to apply a goal orientation to her discourse planning. This ability is a later development and is plausibly emphasized when the child enters the school system, where teachers provide guidance in the construction of a summary title or statement that encapsulates a succeeding oral presentation.

The social aspect of this process is closely tied to parental routines around bedtime when Emily’s parents would engage in talk about happenings of the day just past and also anticipate future events. Through her monologues following such talk, Emily was observed to be comparing the forms in her own speech with the forms she had heard in the speech of more experienced speakers. This comparison resulted in a progressive resystematization of forms that can serve cohesive functions. This process is a slow and gradual one: Development is protracted, and the acquisition of form-function pairings can be said to be at first localized to specific distributional patterns and activity types (cf. Budwig, 1995; Moissinac & Budwig, 2000; Nelson, 1996; Tomasello, 1992; Werner & Kaplan, 1963/1984). Throughout this process, Levy emphasized that the child is actively operating on and experimenting with form-function relationships she has heard in particular situated activities. Her practice of these units in turn transforms them into context for her subsequent speech. Hence, development of discourse abilities over time and within culturally situated practices takes place through “the mediation of discourse themes by the linguistic function of cohesion, the cohesive system, itself constructed through cultural learning, is used to create large-scale units of meaning” (Levy, 1999, p. 244).

Although it cannot be the case that all children engage in the same monologic activities as Emily, it is nonetheless plausible that Emily’s case makes transparent what is usually opaque to language researchers by being akin to Vygotsky’s inner speech (Vygotsky, 1962). Levy’s contribution is thus to demystify an important aspect of text building. In the framework of constructive grounding, the use of pronouns assists in establishing a ground so that new figures can emerge against a slowly changing and growing ground. They function as close-to-optimal indicators for the analysis of the activities of texting. These activities are shown to be not the achievements of the isolated mind, but rather are interactively achieved even in monologic situations such as Emily’s bedtime monologues. Therefore, the contents of language practices (i.e., what the activities are about) do not rest necessarily out-
side and before the activity, such as in the mind or in an independent author behind the text. Contents and aboutness are interactive achievements from where they can be extracted and appropriated as individual achievements, and, in turn, they can be brought back to new interactions and function as resources for new practices.

The perspective that Levy has brought to light is thus a dyadic and communicative perspective. It is expanded into monologic activities as well as thought and purposive behavior in general. By analyzing language forms in the purposive context of language practices, it can be documented how children develop themselves and are developed by others through their participation in communicative (sociocultural) practices. Children, like adults, are able to continuously transform themselves as guided but active participants (Budwig, Uzgiris, & Wertsch, 2000; Rogoff, 1990, 1998) in and through their participation in social communicative contexts. Such a perspective echoes that of linguistic socialization, which has long emphasized the role of cultural factors in the acquisition of language and that the child is socialized to both use language as well as through language practices that index ways of being in the culture at large (e.g., Ochs & Schieffelin, 1984, 1995; Schieffelin & Ochs, 1986).

Although most studies in this area have focused on the development of specific language forms across cultures, there have also been studies on the socialization of children through discourse practices in American families (see Blum-Kulka, 1997; Fivush, 1994; Miller, 1994; Ochs, Smith, & Taylor, 1996; Ochs, Taylor, Rudolph, & Smith, 1992). These studies lend general support to Levy’s point that adult discourse practices with and around children form a potent input for the socialization of children into the discourse patterns as well as the cultural expectations of their speech communities. For example, Ochs et al. (1992) found that discourse around the dinner table socializes children into taking particular perspectives of the world and introduces them to certain ways of solving problems that resemble scientific theory building—a mode of problem solving valued in the education system in which they need to succeed. This similarity between social problem solving and scientific theory building has stood up to closer scrutiny in a comparison between the data of Ochs et al. and problem-based learning in a group medical setting (Hall, 1999). Thus, it is becoming clear that socialization through discourse practices within the family not only endows children with sociolinguistic intelligence (Shatz, 1994), but also prepares them for the cognitive demands of their prospective educational experiences. That such discourse socialization patterns consonant with later practices in the school system have so far been found only for White, Euro-American families appears
to confirm the differential effect of familial language socialization experiences on the school performance of different racial groups in America (e.g., Labov, 1972).

On a more general level, Ochs (1996) theorized that discourse practices in all language communities share common socialization effects that lead novices to take particular affective and epistemic stances in different social situations. It is through discursive practices that members of speech communities learn to use particular language forms to index their emotions and attitudes in different situations as well as to signal the sources and levels of certainty of the information they possess. Further, it is through the patterning of such stances that social identities are created (i.e., “displays of relative certainty and displays of positive and negative affect are the building blocks of identities”; Ochs & Capps, 1997, p. 86). Of course there are important local differences from one culture to the next, but the general function of discursive practices in this respect is theorized to hold across all cultures.

Related to the role of perspective and stance is Bamberg's (1997b, 2000a, 2000b) research into the development of positioning strategies through discourse practices. Based on an extended research program, Bamberg identified three ways in which speakers construct relations among participants when engaging in discourse. At the first level (Positioning Level 1), speakers construct agency roles for the characters being talked about. For instance, they use linguistic means to mark one person as the agent who is in control, whereas the action is inflicted on another; they may portray the central character as being helplessly at the mercy of outside (quasinatural) forces or as being rewarded by luck and fate, or personal qualities (such as bravery, nobility, or simply character). At the second level (Positioning Level 2), it is the pragmatics of the speaker–audience relationship that become the focus of attention. At issue here is whether the speaker attempts to instruct her listeners in terms of what to do in the face of adversarial conditions, for example, or whether she engages in making excuses for her actions, attempting to attribute blame to others. At the third level (Positioning Level 3), it can be shown that speakers use discourse practices to position themselves vis-à-vis themselves in their talk. That is, they employ lexical and grammatical devices to make possible a position from which a sense of self can come to existence that subsequently can be opened up for purposes of reflection. This is the level of our epistemic selves, where beliefs, feelings, and wants are displayed, where claims vis-à-vis one's own actions and internal states are held to be true and relevant beyond the local conversational/discourse situation.

In other words, it can be demonstrated that linguistic devices employed in talk point to more than its content (or what the talk is about)
and the interlocutor. In constructing the content (Positioning Level 1) and one’s audience (Positioning Level 2), speakers (potentially) transcend the question of “How do I want to be understood by you, the audience?” and construct a local answer to the question “How do I want to understand myself?”, which is not any different from the answer to the question “Who am I?” Through the examination of how children from preschool through third grade express the emotions of anger and sadness from the perspectives of the self, the other, and the generalized other and in the role of either perpetrator or recipient, Bamberg has been able to map out a developmental progression of the ability to utilize discursive devices to employ these levels of positioning. In summary, the evidence again points to a protracted course of development that is almost fully functionally in place by the time children reach third grade.

DISCOURSE, IDENTITY, AND GENDER

This section develops further the theme that discursive practices contribute to the formation and maintenance of identities. This direction is an important concern in recent work on discourse development. It is also demonstrated that discourse practices, being inextricably tied to identity formation and change, continue to develop past the childhood years. Among the topics to be discussed are how discourse contributes to the formation of a self-concept, considerations of class and gender, and discursive contributions to group identity in the teenage years.

Discourse and the Autobiographical Self

It should be no surprise that the role of language in the formation of a representation of self seems no longer in dispute (see e.g., Harter, 1999; Neisser & Fivush, 1994; Nelson, 1996). However, how language exactly interacts with memory is still a matter of research (Brockmeier, 2000; Freeman, 1998; Nelson, 1997, 1998; Tessler & Nelson, 1994). Nelson posited that children begin to form autobiographical memories only when they are able to share memories with adults through conversational narratives and also to rework experiences on their own through the use of self-narratives. In this way, children build up a sense of self that endures through time and space. Within this process, the co-construction of past experiences with adults serves to highlight aspects of events that are important to remember (see also Hudson, 1990; Hudson & Fivush, 1991). However, there have been some misgivings about how authentic such reconstructed autobiographical memories can be (e.g., Wolf, 1990). There is the danger of children’s
memories being biased by the input of caregivers (Crittenden, 1994; Pipp, 1990).

More recently, however, it has been asserted that the tension between subjectivity and truth in memory cannot be completely resolved (Ochs & Capps, 1997) and that we constantly reinterpret our past in terms of our present and projected future, make meaning of our present in terms of the remembered past and the prospective future, and project our future in terms of our interpreted present and our subjective past (Phoenix, 1999). Thus, the discursive construction of an autobiographical self is constantly in the process of revision, and any particular construction at a localized point in time and space is an authenticating event for the self (Ochs & Capps, 1997). The importance of the ability to transform the self through participation in situated discursive practices (Budwig, Užgiris, & Wertsch, 2000) is once again brought to mind.

Interestingly enough, the role of narrative in the representation of an autobiographical self has been given support in recent neuropsychological theorizing (Gazzaniga, 1998). Gazzaniga considered that the left brain actively weaves a narrative that confers agentivity and intentionality onto processes that are, to all extents and purposes, automatic. What has evolved to give us a sense of control and selfhood is an interpreter in the left hemisphere that generates a fictional self. It does this by fabricating an illusory narrative that assigns credit to the self for orders already executed (e.g., moving a leg or uttering a sentence). Why the brain does this is not as yet clear, but its adaptive value in a world fraught with uncertainty should be immediately obvious. We do want to have some control over our actions to say the least. The linguistic and discursive means for the brain to produce this adaptation, assuming that it has a basis in reality, must be appropriated through situated language practices that are representative of the mores of a particular culture. From such a perspective, the development of self is clearly intertwined with the development of discursive abilities.

Gender and Class Considerations

Considerations of gender with respect to discourse development can be broadly divided into two questions: (a) Do adults talk to boys and girls in different ways? and (b) Do girls and boys talk to each other in different ways? On the first point, Fivush (1994) reported consistent differences in the conversations between parents and children of different genders from Euro-American middle-class families. Her data indicate that when talking about past events to girls, parents tend to use a more elaborative style, providing more detail and embellishing
the information more than when talking to boys. By contrast, recounting of past events with boys was predominantly short in duration, with questions that were repeated and with minimal amounts of embellishment. Similarly, Ely, Berko-Gleason, and McCabe (1996) found that girls were encouraged to talk about the past more frequently than boys, and that females used more reported speech in their narratives than males. Fivush also found differences in the way these parents interpreted emotional events when talking to their sons and daughters. She reported that mothers focused more on social relations when talking to their daughters about emotions, whereas sons were directed toward autonomous resolutions of emotional situations that, more often than not, involved retaliation.

Research into gendered differences in the discourse interactional styles of boys and girls was given impetus by the finding of such differences in the talk of adults (e.g., Gilligan, 1987). Female discourse was predominantly focused on maintaining relational ties (i.e., affiliative), whereas males were supposed to be concerned with autonomy and dominance (i.e., adversarial). The main vehicle for the investigation of whether such differences were already to be found in the talk of children was discourse around episodes of conflict between children. Sheldon (1990) did indeed find differences along gender lines in the ways children handled conflict. She investigated how preschoolers competed for ownership of an object during pretend play in a triad. Compared with boys, the girls were adept at formulating arguments to defend their own positions as well as to oppose that of their antagonists, but they were more predisposed to negotiating a resolution, maintaining the continuity of the play episode, and keeping up the relations among members of the triad. However, conflict episodes that boys engaged in were more drawn out, involved more coercive physical tactics together with larger amounts of directive speech, and depended on resolution by sheer dominance rather than through joint negotiation. Later research by Sheldon (1992, 1996) has shown that 3- to 5-year-old Euro-American middle-class girls are sociolinguistically skilled at managing conflict through the employment of double-voiced discourse in which the protagonist persists in pursuing her discordant competitive goals while remaining civil and socially facilitative. Additionally, Kyritzis (1994) found that girls and boys use persuasive justifications in markedly different ways. Kyritzis’ analysis showed how boys used justifications to exert control, which was adaptive in the maintenance of power hierarchies. In contrast, girls utilized justifications to validate jointly constructed themes, this being a productive strategy when common ground was to be established.
However, the story is not so clear cut. Proceeding from a conversation analytic orientation, Goodwin and Goodwin (1987) investigated the conflict talk of African-American girls that encompassed a wide range of speech events, including stories, requests, commands, insults, explanations, excuses, threats, and warnings. From this study, they concluded that girls’ conflict talk can be as confrontational as that in boys’ conflicts, includes episodes that can be described as playful, needs to be analyzed as locally constructed and situationally dependent. More recently, Goodwin and Goodwin (2000) extended their findings to the conflict discourse of Central American girls in Los Angeles, showing that the confrontational tactics of these girls are culturally defined practices performed with a full range of linguistic and para-linguistic resources and can be understood as embodied emotions within situated activities.

The disparate findings on conflict management that have been found in different cohorts of girls indicate that gender intersects with class and race in this domain. It is clear that the participants in the Goodwin’s research were not Euro-American nor middle class. Yet how would Euro-American working-class girls compare with their middle-class equivalents who participated in the research of Fivush, Kyratzis, and Sheldon? The work of Peggy Miller and her colleagues (e.g., Miller, 1994; Miller & Sperry, 1987; Wiley, Rose, Burger, & Miller, 1998) should shed some light on this issue. Drawing on samples of working-class children and their families in South Baltimore and Chicago, as well as a comparison sample of middle-class participants in Chicago, Miller’s work clearly depicts class differences in the socialization of anger, aggression, and autonomy across the Euro-American community. Essentially, this work indicates that Euro-American girls of working-class background are socialized from as early as their third year to use verbal resources to communicate anger and aggression and to defend their rights when these are infringed on. In a similar vein, working-class children are also socialized to earn and defend their autonomy, rather than assume it to be conferred on them as a birth right, the latter being characteristic of the middle class. Taking these differences into consideration, it should be reasonable to predict that working-class Euro-American children would be socialized into discourse practices for conflict management more similar to those of their African-American and Hispanic-American working-class counterparts than those of the Euro-American middle-class cohort.

Moving away from conflict talk, Kyratzis (1999) recently demonstrated that dramatic play is also a useful medium through which to investigate gender differences in children’s construction of narrative identities. Observing middle-class preschool children of mixed racial
heritage in same-sexed play groups, Kyritzis reported that girls’ construction of protagonists in dramatic play implied that they valued qualities of attractiveness, graciousness, and lovingness, whereas boys oriented themselves more toward physical power. Further, it was found that the girls were more reliant on the story for self-construction than the boys. Kyritzis’ analysis reveals that the children’s narratives in general were suffused with rounds of duetting voices that allowed them to position themselves with respect to societal norms, values, and cultural identities as well as to position themselves in relation to each other in such a way as to allow them to collaboratively explore the possibilities of different self-constructions. Kyritzis theorized that such experiments in self-construction in dramatic play by children parallels the function of personal experience narratives for adults in intimate social groups. Such work points the way toward future research in both the domain of gender issues as well as the construction of identities through narrative activities.

Discursive Identities in Adolescence

It is generally recognized in Western cultures that adolescence is a time of life when an individual is reassessing her self-concept and the way others view her as well as reevaluating his values and morals in comparison to those of his family and society at large (Damon & Hart, 1982). Some researchers would go as far as to say that the changes in behavior observed in adolescence can be traced to changing hormone levels (see Buchanan, Eccles, & Becker, 1992, for a review), but the issue is still much a matter of debate. Whatever the case, it is clear that adolescence is a time in which teenagers are undergoing large changes and experimentation with issues of identity. Because discourse practices are inextricably intertwined with the processes of identity formation and maintenance, we could also expect that changes in one would effect changes and development in the other.

In an ethnographic study of group identities in a midwestern American high school, Eckert (1989) discovered that identity was polarized into two main groups. The Jocks subscribed to the values of middle-class American culture at large, consequently did well in school, participated wholeheartedly in the extracurricular activities of the school, dressed and groomed themselves in a conservative manner, and generally refrained from the use of illegal drugs. In contrast, the Burnouts formed the group that openly rebelled against mainstream middle-class culture and the rules of the school, dressed and presented themselves in a way that portrayed a counterculture, partook of illegal drugs, and did poorly in school. Eckert (1989) underscored the role of
language practices that separated these two groups by claiming that "perhaps the strongest evidence of the depth of the difference between Jocks and Burnouts lies in their use of language" (p. 67). Although she did not refer explicitly to discourse practices, she listed differences that conceivably are components of discourse practices, such as the Burnouts' more frequent and public use of obscenities, drug-related slang, preference for the use of nonstandard grammar (e.g., the multiple negative \(I \text{ don't know nothing}\)), and phonological patterns (Eckert, 2000) that were reflective of a progressive historical change toward working-class patterns in the area in general. Eckert (1989) emphasized that "these kinds of linguistic features differentiate the Jocks and the Burnouts and serve as powerful symbols of category membership" (p. 68). Based on a theory of variation as social practice, Eckert (2000) viewed discourse practices not simply as reflections of particular memberships or identity claims vis-à-vis particular social categories, but rather "as related to the practices that give rise to and maintain those categories, and that make membership in them meaningful" (p. 4).

Deppermann (1999) reported a similar phenomenon in the use of idiosyncratic language and discourse patterns to forge a distinctive identity among disenfranchised adolescents in Germany. He showed how group identity is established by discourse strategies that assume the voice of an out-group to violate its norms. In doing so, the in-group gains a positive identity. In addition, he found that the in-group identity in this case was bolstered by participation in a discourse mode characterized by playful, fictional provocation and funny activities that included quick repartees. Deppermann (1999) pointed out that "it is the shared interactional style of the group that provides for the relevances from which identities are built" (p. 20).

In an effort to study the change in discourse practices with respect to femininity in the teenage years, Coates (1999) reported a study in which she asked four girls in North London, England, to audiotape periods when they spent time together on the weekends. They were recorded from the time they were 12 until they were 15 and were instructed to contact the researcher whenever they had filled up a 90-minute audiotape. Initially, it was found that their interactions were principally of a playful nature, but as they approached 14 and 15, their talk became more serious, included more elements of self-disclosure, and displayed increasing amounts of mutual support as they struggled to make sense of their changing worlds. A more important change was that from the girls' assumption of active social agency at 12 and 13 toward being subject to the dominant discourses of patriarchy as they grew older. The girls progressively portrayed themselves as losing the
power to resist the controlling ramifications of the dominant masculine discourses. In short, they were approximating the established discourse modes of adult female speakers. On following up two of the girls in their late teens, Coates' participants informed her that, although they had had a difficult time in their early teens, the struggle had made them more aware of their own emotions and had helped forge a firm friendship that became important to the further development of their femininities. Thus, Coates concluded that, although it is discouraging that privileged, middle-class White teenage girls still appropriated the dominant masculine discourses and then became subject to their constraining effects, there is nonetheless cause for encouragement that in the same process the friendships that they developed were resources that they could draw on to choose among different versions of femininity and reconcile contradictory ones.

Another phenomenon that has attracted attention recently is the appropriation of discourse styles of African-American English by Euro-American teenagers to situationally redefine their masculinity (e.g., Bucholtz, 1998; Cutler, 1998). Bucholtz theorized that there is an assumed gradient of masculinity tied to stereotypical activities and symbolisms among Euro-American high school males, with African American as most masculine, Euro-American as intermediate, and Asian American as least masculine. In both Bucholtz's and Cutler's work, it is demonstrated how Euro-American males use African-American English to index traits stereotypically associated with that culture. In their talk with their friends, they switch into this discourse mode when they want to index their own involvement in similar activities, express personal rebelliousness, and generally appropriate their conception of African-American masculinity through discursive means. Nevertheless, they engage in a particular form of distancing from the identity that they seem to be appropriating by using the African-American English forms in ways not fully competent. Here we observe the plasticity of identity performance that fashions a convenient, self-serving form of borrowed identity while not completely giving up one's own regular identity construction.

A similar process can be observed among Turkish youth in Germany (Hinnenkamp, 1998) who code switch between Turkish and German both to distinguish themselves from their immigrant parents, as a form of cultural emancipation, and also as a resource to entertainingly negotiate the crossing of ethnic boundaries. They do this intrasententially and even within particular words, thus indexing the co-occurrence of dialect forms while indicating that this process of identity construction is done on a fully elective basis. Thus, they are
consciously renegotiating the construction of their identities in the real time of interaction.

**SUMMARY AND CONCLUDING REMARKS**

We hope to have shown that discourse development is a complex process that involves an active novice who participates in language practices in situationally defined contexts. Of course, in a chapter as condensed as this one, it is impossible to cover all the fields and subfields that can be subsumed under the title of discourse development. The focus on the oral aspects of discourse development was a conscious one because we consider these the basis of literacy and written forms of discourse (Bamberg, 2002), bearing in mind the adage that “one writes as one speaks” (although also being aware that the development of written discourse follows other developmental strands and opens new dimensions of language awareness). Other notable omissions are the development of adult registers of discourse, the change and/or maintenance of discourse skills with advancing age and the pathologies that accompany it such as Alzheimer’s disease, and the development of discourse skills for populations that have special needs (e.g., individuals with dyslexia, specific language impairment, and learning disabilities). The latter two topics would comfortably fit together in a separate chapter on discourse development under special conditions. With respect to the development of the adult registers of discourse, we have hoped to demonstrate that discourse development clearly continues through the teenage years and, by extension, into adulthood.

Through the exposition of the varied research paradigms that address discourse development, we have tried to highlight the idea that development is principally in the eye of the researcher. However, we find the approach that has been contributed by Levy particularly promising for future directions of research. This is because it unequivocally combines the subdomains of conversation and extended discourse into a holistic entity while grounding the result in a practice-based approach to communicative development and human development in general. In doing so, it holds promise for furthering the understanding of how processes once considered separate feed into and reinforce one another. Combined with a focus on identity formation through discourse practices, such a conception of discourse development may form the core of a redefinition of the boundaries between language and communication.
At risk of being glib, we hazard an answer to the question posed at the beginning (i.e., What is language divorced from discourse?). Simply put, the answer is not much because communication would then be almost impossible. Hence, we suggest that in discourse language and communication come together, each providing potential resources for the other to draw on. This fusion should be expressed as a redefinition of discourse as communicative practices that bypass the trichotomy among the person, a language, and a world (that is external to language), where words or sentences refer to objects or events in the world and where language is the tool to (socially) connect people. Rather, we are suggesting that discourse as language/communicative practices is possibly the central mode of engagement not with but in the world. World and person making take place through discourse practices. To put it more radically, without language there would be no discourse, but, more important, without discourse language would remain an abstract system through which worlds and persons could not be coherently constructed.

In line with Kaplan’s (1983) insight that it is impossible to read off development directly from phenomena in the world, we would like to affirm that it is equally impossible to see or read off changes and permanence from phenomena in the world. What we as humans consider stable and what we consider mutable requires a position from which a figure and ground can be evoked so that time can be infused therein as a meaningful entity and serve as the tertium comparationis, the common third, from which a comparison of “constancies and changes over time” is possible. This position is neither a priori in the author nor in the world. It emerges in and through discourse in social practices.

In summary, we hope to have made clear the importance of the newer approaches to discourse development that emphasize the constitutive aspects of discourse in our meaning-making activities. The centrality of our fluid, although temporally coherent processes of identity formation inherent to these processes, should warrant much more concentrated research efforts in the future.

REFERENCES

11. DISCOURSE DEVELOPMENT


American and Latino children’s personal narrative. Paper presented at the 16th annual conference on Language development, Boston University, Boston, MA.


Discourse in Computational Linguistics and Artificial Intelligence

Johanna D. Moore
Peter Wiemer-Hastings
University of Edinburgh

Models of discourse structure and processing are crucial for constructing computational systems capable of interpreting and generating natural language. Research on discourse focuses on two fundamental questions within computational linguistics and artificial intelligence. First, what information is contained in extended sequences of utterances that goes beyond the meaning of the individual utterances? Second, how does the context in which an utterance is used affect the meaning of the individual utterances or parts of them?

Discourse research in computational linguistics and artificial intelligence (AI) encompasses work on spoken and written discourse, monologues as well as dialogues (both spoken and keyboarded). The questions that discourse research attempts to answer are relevant to all combinations of these features. The juxtaposition of individual clauses may imply more than the meaning of the clauses themselves, regardless of whether the clauses were contributed by the same speaker (writer).\(^1\) Likewise, the context created by prior utterances affects the current one regardless of which participant uttered it.

In this chapter, we first provide an overview of types of discourse structure and illustrate these with examples. We then describe the in-

\(^1\)Henceforth, we use speaker and hearer to indicate the producer and interpreter of discourse, respectively, whether it is spoken or written. Where the distinction between spoken and written discourse is important, we are more explicit.
fluential theories that account for one or more types of structure. We then show how these theories are used to address specific discourse processing phenomena in computational systems. Finally, we discuss the use of discourse processing techniques in a range of modern language technology applications.

OVERVIEW OF DISCOURSE STRUCTURE

Researchers in computational linguistics have long argued that coherent discourse has structure, and that recognizing the structure is a crucial component of comprehending the discourse (Grosz & Sidner, 1986; Hobbs, 1993; Moore & Pollack, 1992; Stone & Webber, 1998). Interpreting referring expressions (e.g., pronouns and definite descriptions), identifying the temporal order of events (e.g., the default relationship between the falling and pushing events in “Max fell. John pushed him.”), and recognizing the plans and goals of our interlocutors all require knowledge of discourse structure (Grosz & Sidner, 1986; Kehler, 1994b; Lascarides & Asher, 1993; Litman & Allen, 1987). Moreover, early research in language generation showed that producing natural-sounding multisentential texts required the ability to select and organize content according to rules governing discourse structure and coherence (Hovy, 1988b; McKeown, 1985; Moore & Paris, 1993).

Although there is still considerable debate about the exact nature of discourse structure and how it is recognized, there is a growing consensus among researchers in computational linguistics that at least three types of structure are needed in computational models of discourse processing (Grosz & Sidner, 1986; Hobbs, 1993; Moore & Pollack, 1992). These are described next.

Intentional structure describes the roles that utterances play in the speaker’s communicative plan to achieve desired effects on the hearer’s mental state or the conversational record (Lewis, 1979; Thomason, 1990). Intentions encode what the speaker was trying to accomplish with a given portion of discourse. Many have argued that the coherence of discourse derives from the intentions of speakers, and that understanding depends on recognition of those intentions (e.g., Grice, 1957; Grosz & Sidner, 1986). Research in response generation shows that, to participate in a dialogue, agents must have a representation of the intentional structure of the utterances they produce. Intentional structure is crucial for responding effectively to questions that address a previous utterance; without a record of what an utterance was intended to achieve, it is impossible to elaborate or clarify
that utterance (Moore, 1995; Young, Moore, & Pollack, 1994a). Moreover, speaker intentions are an important factor in generating nominal expressions (Appelt, 1985; Green, Carenini, & Moore, 1998) and selecting appropriate lexical items, including discourse cues (e.g., because, thus; Moser & Moore, 1995; Webber, Knott, Stone, & Joshi, 1999) and scalar terms (e.g., difficult, easy; Elhadad, 1995).

*Informational structure* consists of the semantic relationships between the information conveyed by successive utterances (Moore & Pollack, 1992). Causal relations are a typical example of informational structure, and psychologists working in reading comprehension have shown that these relations are inferred during reading (Gernsbacher, 1990; Graesser, Singer, & Trabasso, 1994; Singer, Revlin, & Halldorson, 1992). In addition, several researchers identified types of text whose organization follows the inherent structure of the subject matter being communicated (e.g., the structure of the domain plan being discussed; Grosz, 1974; Linde & Goguen, 1978) or the spatial (Sibun, 1992; Linde, 1974), familial (Sibun, 1992) or causal relationships (Paris, 1988; Suthers, 1991) between the objects or events being described, or the states and events being narrated (Lehnert, 1981). Several systems that generate coherent texts based on domain or informational structure have been constructed (Paris, 1988; Sibun, 1992; Suthers, 1991).

*Attentional structure* as defined by Grosz and Sidner (1986) contains information about the objects, properties, relations, and discourse intentions that are most salient at any given point in the discourse. In discourse, humans focus or center their attention on a small set of entities and attention shifts to new entities in predictable ways. Natural language understanding systems must track attentional shifts to resolve anaphoric expressions (Gordon, Grosz, & Gilliom, 1993; Grosz, 1977; Sidner, 1979) and understand ellipsis (Carberry, 1983; Kehler, 1994a). Natural language generation systems track focus of attention as the discourse as a whole progresses as well as during the construction of individual responses to influence choices on what to say next (Kibble, 1999; McCoy & Cheng, 1990; McKeown, 1985), determine when to pronominalize (Elhadad, 1992), make choices in syntactic form (e.g., active vs. passive; Elhadad, 1992; McKeown, 1985; Mittal, Moore, Carenini, & Roth, 1998), appropriately mark changes in topic (Cawsey, 1993), and generate elliptical utterances.

In addition to these three primary types of discourse structure, the literature on discourse in computational linguistics has discussed two additional types of structure. One of them, rhetorical structure, has had considerable impact on computational work in natural language generation.
Information structure consists of two dimensions: (a) the contrast a speaker makes between the part of an utterance that connects it to the rest of the discourse (the theme), and the part of an utterance that contributes new information on that theme (the rheme); and (b) what the speaker takes to be in contrast with things a hearer is or can be attending to. Information structure can be conveyed by syntactic, prosodic, or morphological means. Steedman argued that information structure is the component of linguistic structure (or grammar) that links intentional and attentional structure to syntax and prosody via compositional semantics for notions like theme (or topic) and rheme (or comment). Recently, a number of theories of information structure (Steedman, 1991; Vallduvi, 1990) have brought hitherto unformalized notions like theme, rheme, and focus within the compositional semantics that forms a part of formal grammar (Steedman, 2000, 2001).

Rhetorical structure is used by many researchers in computational linguistics to explain a wide range of discourse phenomena. There have been several proposals defining the set of rhetorical (or discourse or coherence) relations that can hold between adjacent discourse elements and researchers have attempted to explain the inferences that arise when a particular relation holds between two discourse entities, even if that relation is not explicitly signaled in the text. Researchers in interpretation have argued that recognizing these relationships is crucial for explaining discourse coherence, resolving anaphora, and computing conversational implicature (Hobbs, 1983; Lascarides & Asher, 1993; Mann & Thompson, 1988). Researchers in generation have shown that it is crucial for a system to recognize the additional inferences that are conveyed by the sequence of clauses they generate because these additional inferences may be the source of problems if the user does not understand or accept the system’s utterance. Moreover, to implement generation systems capable of synthesizing coherent multisentential texts, researchers identified patterns of such relations that characterize the structure of texts that achieve given discourse purposes, and many text generation systems have used these patterns to construct coherent monologic texts to achieve a variety of discourse purposes (Hovy, 1991; McKeown, 1985; Mellish, O'Donnell, Oberlander, & Knott, 1998; Mittal et al., 1998; Moore & Paris, 1993; Rösner & Stede, 1992; Scott & de Souza, 1990).

Much of the remaining debate concerning discourse structure within computational linguistics centers around which of these struc-

---

2Moore and Pollack (1992) argued that the rhetorical relations used in these systems typically conflate informational and intentional considerations, and thus do not represent a fourth type of structure.
tures are primary and which are parasitic, what role the structures play in different discourse interpretation and generation tasks, and whether their importance or function varies with the discourse genre under consideration. For example, a major result of early work in discourse was the determination that discourses divide into segments much like sentences divide into phrases. Each utterance of a discourse either contributes to the preceding utterances or initiates a new unit of meaning that subsequent utterances may augment. The usage of a wide range of lexicogrammatical devices correlates with discourse structure, and the meaning of a segment encompasses more than the meaning of the individual parts. In addition, recent studies show significant agreement among segmentations performed by naive subjects (Passonneau & Litman, 1997). Discourse theories differ about the factors they consider central to explaining this segmentation and the way in which utterances in a segment convey more than the sum of the parts. Grosz and Sidner (1986) argued that intentions are the primary determiners of discourse segmentation, and that linguistic structure (i.e., segment embedding) and attentional structure (i.e., global focus) are dictated by relations between intentions. Polanyi (1988) took an opposing view and claimed that hierarchical structure “emerges from the structural and semantic relationships obtaining among the linguistic units which speakers use to build up their discourses” (p. 602). In Hobbs (1985), segmental structure is an artifact of binary coherence relations (e.g., background, explanation, elaboration) between a current utterance and the preceding discourse. Despite these different views, there is general agreement concerning the implications of segmentation for language processing. For example, segment boundaries must be detected to resolve anaphoric expressions (Asher, 1993; Grosz & Sidner, 1986; Hobbs, 1979; Passonneau & Litman, 1997). Moreover, several studies have found prosodic as well as textual correlations with segment boundaries in spoken language (Grosz & Hirschberg, 1992; Hirschberg, Nakatini, & Grosz, 1995; Nakatani, 1997; Ostendorf & Swerts, 1995), and appropriate usage of these intonational indicators can be used to improve the quality of speech synthesis (Davis & Hirschberg, 1988).

In the sections that follow, we further examine the theories of discourse structure and processing that have had significant impact on computational models of discourse phenomena. A comprehensive survey of discourse structure for natural language understanding appears in Grosz, Pollack, and Sidner (1989), and therefore we focus on discourse generation and dialogue in this chapter. We then review the role that discourse structure and its processing play in a variety of current natural language applications. The survey in Grosz et al. (1989)
focuses largely on the knowledge-intensive techniques prevalent in the late 1980s. Here we emphasize statistical and shallow processing approaches that enable discourse information to be used in a wide range of today’s natural language technologies.

### COMPUTATIONAL THEORIES OF DISCOURSE STRUCTURE AND SEMANTICS

#### Discourse Representation Theory

Discourse Representation Theory (DRT) is a formal semantic model of the processing of text in context that has applications in discourse understanding. DRT was originally formulated in Kamp (1981) and further developed in Kamp and Reyle (1993), with a concise technical summary in van Eijck and Kamp (1997). DRT grew out of Montague’s model-theoretic semantics (Thomason, 1974), which represents the meanings of utterances as logical forms and supports the calculation of the truth conditions of an utterance. DRT addresses a number of difficulties in text understanding (e.g., anaphora resolution) that act at the level of the discourse.

This section gives a brief overview of the philosophy behind DRT, the types of structures and rules that DRT uses, and the particular problems that it addresses. We also describe some of the limitations of the standard DRT theory.

**Philosophical Foundations of DRT.** As mentioned earlier, DRT is concerned with ascertaining the semantic truth conditions of a discourse. The semantic aspects of a discourse are related to the meaning of the discourse, but not related to the particular situation (including time, location, common ground, etc.) in which the discourse is uttered. The advantage to this approach from the logical point of view is that the semantic representation for the discourse can be automatically (more or less) built up from the contents (words) and structure of the discourse alone without bringing in information about the external context of the utterance. Once constructed, it can be compared with a logical representation of some world (a model in DRT terms) to determine whether the discourse is true with respect to that model.

**DRT Structures.** The standard representation format in DRT, known as a discourse representation structure (DRS), consists of a box with two parts as shown in Fig. 12.1. The top part of the box lists the discourse referents, which act as variables that can be bound to different
entities in the world. The bottom section of the DRS lists the propositions claimed to be true of those referents in the described situation. Figure 12.1 gives the DRS of the sentence “John sleeps.” The representation can be read as, “There is an individual who is named John, and of whom the sleep predicate is true.” This is equivalent to the logical expression: \((x) \land \text{John}(x) \land \text{sleep}(x)\).

**DRT Rules and Processing.** To derive a structure like the one shown here, DRT uses a set of standard context-free grammar rules and a set of semantic interpretation rules based on the syntactic structure of the input sentence. Figure 12.2 shows a simple DRT rule for processing proper nouns. The left-hand side shows a segment of the syntactic tree that must be matched, and the right-hand side shows the result of applying the rule, including adding propositions to the DRS and changing the parse tree. This rule applies to the structure on the left in Fig. 12.3, which shows the parse tree of the example sentence, “John sleeps,” in a DRS. The rule produces the representation on the right in Fig. 12.3 by deleting part of the parse tree, inserting a variable in its place, and adding a proposition to the DRS.

Next, a similar rule is applied that reduces the verb phrase to a proposition, sleep, which is true of a new discourse referent, \(y\). Then the sentence rule deletes the remaining syntactic structure and equates the subject discourse referent with the object referent, \(x = y\). Finally, \(x\) is substituted for \(y\) in the sleep proposition, producing the structure shown in Fig. 12.1. The next sentence in a discourse is proc-
essed by adding its parse tree to the box and applying the semantic transformation rules to it.

The construction of a complete DRS enables the calculation of its truth value with respect to a model. A model is a formal representation of the state of the world. Figure 12.4 shows a model distinguished from a DRS by its double box. The elements in the top section of a model are interpreted differently than those in a DRS. In a DRS, they are variables that can bind to entities in the world. In a model, each referent indexes a particular entity in the world.

In this example, the DRS is true with respect to the model because there is a consistent mapping ($x = b$) between the discourse referents in the DRS and the individuals in the model, and the model contains all of the propositions that are in the DRS. Because the model is taken to be a snapshot of the world, it may contain many additional propositions not in the DRS without affecting the truth conditions of the DRS. Those that are not relevant to the DRS are simply ignored.

**Uses of DRT.** The major advantages of DRT are that it provides a simple, structure-based procedure for converting a syntactic representation of a sentence into a semantic one, and that semantic repre-
sentation can be mechanically compared to a representation of the world to compute the truth-conditional status of the text. DRT also addresses (at least partially) the difficult discourse problems of anaphora resolution, quantifier scoping, and presupposition.

When a pronoun is processed in DRT, the semantic interpretation rule adds an instruction of the form \( x = ? \), which is read as, “find some discourse referent \( x \) in the discourse context.” The referent must satisfy three constraints: consistency, structural, and knowledge. The consistency constraint specifies that the new mapping of a discourse referent must not introduce a contradiction into the DRS. In practice, this ensures that number and gender restrictions are applied. The structural constraint limits where in a complex DRS a coreferent can be found. In practice, this is similar to the constraints proposed in centering theory (Grosz, Joshi, & Weinstein, 1995). Finally, the knowledge constraint is intended to prohibit the inclusion of any coreference that would violate world knowledge or common sense. Unfortunately, the scope of this constraint makes a complete implementation of it impossible.

DRT is the foundation of recent research in psycholinguistics, which attempts to model human judgments of the acceptability of a range of anaphors. Gordon and Hendrick (1997) collected ratings from humans of the acceptability of various combinations of names, quantified, definite and indefinite noun phrases, and pronouns. Their results show that human judgments did not support some of the constraints on coreference acceptability that came from classical binding theory (Chomsky, 1981). Gordon and Hendrick (1998) claimed that a model of coreference based on DRT corresponds better with human acceptability judgments.

Of course pronominal coreference is not the only type of anaphora. Asher (1993) addressed other types of anaphora as described later. In sentences like (1), there is a structural ambiguity concerning the scope of the quantifier *every*. Specifically, there are two readings of the sentence: one in which each farmer owns a different donkey, and one in which all the farmers collectively own a particular donkey. When processing such a sentence with DRT, there is a choice of processing the quantifier before or after processing the indefinite noun phrase. The two orders of rule application produce the two different structures shown in Fig. 12.5.

(1) Every farmer owns a donkey.

Both of these DRSs include substructures that represent the quantifier as a conditional. They are read as, “if the conditions on the left
hold, the conditions on the right must also hold." In the DRS on the left, the donkey is within the scope of the conditional. Thus, for every farmer, there should be a (potentially different) donkey. In the DRS on the right, the referent for the donkey is global and outside the scope of the conditional. Thus, there is one donkey that every farmer owns. Although this example applies only within a sentence, it suggests how hierarchical discourse relations can be represented by variants of DRT as described next.

DRT’s treatment of presupposition in discourse is related to the way it handles quantifier scoping. In particular utterances such as (2a), certain propositions are said to project out of the sentence—that is, they are held to be true regardless of whether the premise of the conditional is. In (2a), there is a presupposition (at least for rhetorical purposes) that John has a dog, and that presupposition is true regardless of whether she has fleas. For constructs like “John’s dog,” DRT creates discourse referents at the global level of the representation that correspond to John and his dog.

(2) a. If John’s dog has fleas, she scratches them.
    b. She wears a flea collar though.
    b. They jump around a lot.

The DRS for (2a) is shown in Fig. 12.6 and is similar in structure to the one shown on the left in Fig. 12.5. Here the discourse referent for
the fleas is located within the internal left-hand box for the conditional
structure. Thus, a follow-up sentence like (2b) provides no difficulty
for anaphora resolution. But a sentence like (2b ) is not acceptable; the
referent for the fleas is inaccessible because it is embedded within a

**Limitations.** Although it addresses the issues described earlier, DRT
is somewhat limited as a theory of *discourse*. Most of its mechanisms
address within-sentence processing. The addition of new sentences to
a DRS is done in a simple way. The syntactic structure for the new sen-
tence is added to the current DRS, and the semantic construction rules
convert it to additional conditions. There is no accounting for the rela-
tionships that apply between utterances.

One direction of current research is on the tense and aspect of verbs
in sentences (Kamp & Reyle, 1993). This involves the addition of a dis-
course referent that refers to an event. A sentence like “John slept” is
represented with the predicate *sleep*(t1, j), with the additional infor-
mation *t1 n*, where *n* refers to the current time (now). Thus, this vari-
ant of DRT allows the explicit representation of the temporal relation-
ships between sentences in a discourse.

Another difficulty for standard DRT is anaphora with plural refer-
ents. There is no simple mapping between pronouns and plural refer-
ents. Instead, additional inference is required to produce appropriate
referents. In (3), the pronoun *they* refers not to any particular student,
but to the entire set of students. To address this problem, Kamp and
Reyle (1993) extended DRT to deal with conjunctions and quantifiers
like *most*.

(3) Every student passed the exam. They were very happy.

In discourses like (4), the pronoun *it* does not refer to an entity, but
to the fact that John failed the exam. Asher (1993) took this as a starting
point for his research, which deals with reference to abstract objects.
The current formalization of this research is Segmented DRT, which in-
cludes rhetorical relations between sentences; it is described later.

(4) John failed the exam, and it didn’t make him happy.

**Grosz and Sidner Theory**

In Grosz and Sidner's (1986) theory of discourse (henceforth G&S),
discourse structure is a composite of three interrelated structures
called *intentional, attentional, and linguistic structure*. In G&S, in-
**Intentional structure** consists of *discourse segment purposes* and the relationships between them. Discourse segment purposes reflect the communicative intentions of the discourse participants and lead to the utterances produced. Discourse segment purposes are thus an extension of the intentions in Grice’s (1957) theory of utterance meaning: They achieve their desired effect in part by being recognized.

In G&S, a discourse segment is a group of utterances that function together to realize a speaker's purpose. A segment DS\(_n\) originates with the speaker's intention: It is exactly those utterances that the speaker produces to satisfy a communicative intention I\(_n\) in the intentional structure. In other words, I\(_n\) is the discourse segment purpose (DSP) of DS\(_n\). The discourse structure is a hierarchy of segments originating with the structure of the speaker's intentions when producing the discourse. G&S defines two relations that may hold among DSPs. A purpose I\(_m\) *dominates* another purpose I\(_n\) when fulfilling I\(_n\) is part of fulfilling I\(_m\). A purpose I\(_n\) *satisfaction-precedes* another purpose I\(_m\) when I\(_n\) must be satisfied first. The dominance and satisfaction-precedence relations impose a hierarchical structure on the speaker's intentions, and this in turn determines the linguistic structure of discourse. DS\(_n\) is embedded in another segment DS\(_m\) just when the purposes of the two segments are in the dominance relation (i.e., I\(_m\) dominates I\(_n\)). The dominance relation among intentions fully determines the embeddedness relations of the discourse segments that realize them.

As an example, consider the discourse shown in Fig. 12.7 adapted from Mann and Thompson (1988). The whole discourse is a segment, DS\(_0\), that attempts to realize I\(_0\), the speaker’s intention for the hearer to adopt the intention of attending the party. As part of her plan to achieve I\(_0\), the speaker generates I\(_1\), the intention for the hearer to adopt the belief that there will be lots of good food. Then as part of her plan to achieve I\(_1\), the speaker generates I\(_2\), the intention that the hearer be-

<table>
<thead>
<tr>
<th>Intentional Structure</th>
<th>Linguistic Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>I(_0): Intend(_S)(Intend(_H) a)</td>
<td>DS(_0) (a) Come to the party for the new President.</td>
</tr>
</tbody>
</table>
| I\(_1\): Intend\(_S\)(Believe\(_H\) b) | [DS\(_1\) (b) There will be lots of good food.]
| I\(_2\): Intend\(_S\)(Believe\(_H\) c) | [DS\(_2\) (c) The Fluted Mushroom is doing the catering.] |

**FIG. 12.7.** For G&S, dominance in intentional structure determines embedding in linguistic structure.
lieve that a mutually known good caterer is providing the food. As shown on the left in Fig. 12.7, I_0 dominates I_1, which in turn dominates I_2. Due to these dominance relations, the discourse segment that realizes I_2 is embedded in the discourse segment for I_1, which is in turn embedded within the discourse segment for I_0 as shown on the right in the figure. The dominance of intentions directly determines embedding of segments.

One of the most important aspects of the work of G&S is the investigation of the relation between intentional and attentional structure in discourse. They model attentional state using a stack of focus spaces and a set of transition rules for pushing and popping spaces from the stack. Each focus space is associated with a discourse segment and contains the entities salient either because they have been mentioned explicitly in the segment or because they become salient in the process of comprehending the utterances in the segment. The DSP is also included in the focus space associated with a segment. G&S argue that attentional structure is parasitic on intentional structure; more precisely, the relationships among DSPs determine pushes and pops of focus spaces.

G&S provides a model of the use of referring expressions and aids in (a) determining the range of possible referents that an initial description makes available, and (b) choosing among the possibilities to determine the entity(ies) to which a pronoun or definite description refers. Determining how referring expressions interact with attentional state is crucial for solving these problems. A detailed account of the G&S approach to anaphora resolution is beyond the scope of this chapter. Instead we provide an example. Further detail can be found in Grosz and Sidner (1986) and Grosz et al. (1989).

Consider the example discourse in Fig. 12.8. G&S would break this discourse into two segments, DS_0 and DS_1, with the embedding shown in the figure. This structure accounts for why the utterances in 1–2 and 10–14 are more closely related to one another than to those in the intervening segment DS_1. In addition, G&S dictates that the focus space relating to the embedded segment DS_1 would be popped off the stack by the time the definite noun phrase the tent in 14 is interpreted, and thus explains how participant A can determine that B is referring back to the tent introduced in utterance 2 and not the tent introduced in utterance 7. Lexicogrammatical clues to this segmentation are given by the last trip and this trip.

As we have seen, DRT also deals with anaphora resolution, but DRT does not properly constrain the range of possible referents, and therefore both tents would be accessible. DRT overgenerates antecedents because it does not consider intention or the relation of intention to
discourse segmentation and attentional structure, and therefore cannot rule out entities introduced in utterances 3–9. Also note that a full account of pronoun resolution requires a more fine-grained model of attentional state than that provided by attentional structure. An account of that phenomenon that is compatible with G&S is provided by centering theory, which is described in detail in Grosz et al. (1995). Finally, it has been recognized that there are significant difficulties in recognizing the speaker’s discourse plan. These are thoroughly described in Grosz et al. (1989).

Rhetorical Structure Theory

As we have seen, in Grosz and Sidner’s theory, speaker intentions and the relations of dominance and satisfaction precedence that may obtain among them are the primary determiners of discourse structure. Other researchers such as Hobbs (1983, 1985) downplay the importance of intention, arguing that the role of the speaker’s intention is indirect and there are many discourse problems for which the speaker’s intentions are uninformative.

The relation-based theories attempt to explain how coherent discourse conveys more than the sum of its parts. They aim to provide a more detailed account of the inferences that hearers can be expected to make when interpreting a series of utterances based on the assumption that they form a coherent discourse.

Hobbs (1979, 1983, 1985) characterized coherence in terms of a set of binary coherence relations between a current utterance and the pre-
ceding discourse. He identified four reasons why a speaker breaks a discourse into more than one clause and classified the relations accordingly. For example, if a speaker needs to connect new information with what is already known by the hearer, the speaker chooses one of the linkage relations, such as BACKGROUND or EXPLANATION. If a speaker wishes to move between specific and general statements, he or she must employ one of the expansion relations, such as ELABORATION or GENERALIZATION. According to Hobbs, how the speaker chooses to continue a discourse is equivalent to deciding which relation to employ. From the hearer’s perspective, understanding why the speaker continued as he or she did is equivalent to determining what relation was used.

Hobbs (1979) originally proposed coherence relations as a way of solving some of the problems in interpreting discourse (e.g., anaphora resolution). He defined coherence relations in terms of inferences that can be drawn from the propositions asserted in the items being related. For example, Hobbs (1985) defined ELABORATION as follows:

ELABORATION: $S_1$ is an ELABORATION of $S_0$ if the hearer can infer the same proposition $P$ from the assertions of $S_0$ and $S_1$. (p. 25)

Here $S_1$ represents the current clause or larger segment of discourse and $S_0$ an immediately preceding segment. $S_1$ usually adds crucial information, but this is not part of the definition because Hobbs wishes to include pure repetitions under ELABORATION.

Hobbs’ theory of coherence is attractive because it relates coherence relations to the functions that speakers wish to accomplish in a discourse. Thus, Hobbs’ theory could potentially be used in text generation to indicate what coherence relation should be used to achieve a particular goal of the speaker. For example, Hobbs (1979) noted two functions of ELABORATION. One is to overcome misunderstanding or lack of understanding, and another is to “enrich the understanding of the listener by expressing the same thought from a different perspective.” However, note that such specifications of the speaker’s intentions are not an explicit part of the formal definition of the relation. For this reason, many researchers in text generation have turned to an alternative theory of text structure—Rhetorical Structure Theory (RST; Mann & Thompson, 1988).

As in Hobbs’s account, RST characterizes coherence in terms of a set of relations between contiguous spans of text in a discourse. One of the main aims of RST was to account for conventional inferences that arise when interpreting discourse. For example, Mann and Thompson (1986) argued that in addition to the propositions represented explic-
itly by independent clauses in a text, there are many implicit propositions, called relational propositions, which arise out of the combination of clauses. They argued that the predicates of these propositions come from a small set of general, frequently occurring relational predicates (e.g., cause, solutionhood, concession). These propositions arise from the hearer’s search for coherence among utterances that occur together—that is, from the hearer’s assumption that the parts of a discourse form an intentionally constructed sequence of linguistic expressions. They emphasize that relational propositions are inferences that arise from the combination of the two parts of a text and cannot be derived from either of the parts independently.

Consider the following example from Mann and Thompson (1986):

(5) a. I’m hungry.
   b. Let’s go to the Fuji Gardens.

Mann and Thompson (1986) argued that the assumption that this is a coherent discourse gives rise to the relational proposition that (5b) provides a solution to the problem posed by (5a). The relational predicate associated with this proposition is called solutionhood. Note that although the solutionhood relation is not explicitly expressed anywhere in the text, it arises out of the juxtaposition of the two discourse elements. As we discuss later, it is crucial that systems generating such texts recognize these implicit relations that are conveyed because they may be the source of problems if the user does not understand or accept the system’s utterance.

Mann and Thompson made an argument that relational propositions are more basic than other sorts of inferences that arise from texts and cited as evidence the fact that virtually every language has conjunction morphemes to signal them (e.g., in English, because, therefore, so, however). Rhetorical Structure Theory attempts to define a set of rhetorical relations that accounts for these relational propositions.

The definition of each rhetorical relation in RST indicates constraints on the two entities being related, constraints on their combination, as well as a specification of the effect that the speaker is attempting to achieve on the hearer’s beliefs or inclinations. Thus, RST provides an explicit connection between the speaker’s intention and the rhetorical means used to achieve it.

As an example, consider the RST definition of the MOTIVATION relation shown in Table 12.1. As shown, an RST relation has two parts: a nucleus (N) and a satellite (S). The MOTIVATION relation associates text expressing the speaker’s desire that the hearer perform an action (the nucleus) with material intended to increase the hearer’s desire to
perform the action (the satellite). For example, in the following text, (6a) and (6b) are related by MOTIVATION:

(6) a. Come to the party for the new president.
    b. There will be lots of good food.

The nucleus of the relation is that item in the pair that is most essential to the writer’s purpose. In the prior example, assuming that the writer’s intent is to make the hearer go to the party, clause (6a) is nuclear. In general, the nucleus could stand on its own, but the satellite would be considered a nonsequitur without its corresponding nucleus. In this example, without the recommendation to “come to the party,” the satellite in (6b) is out of place. Moreover, RST states that the satellite portion of a text may be replaced without significantly altering the intended function of the text. The same is not true for the nucleus. For example, replacing (6b) with:

(7) b. All the important people will be there.

does not greatly change the function of the text as a whole. However, replacing the recommendation in the nucleus—for example,

(8) a. Don’t go to the party.

significantly alters the purpose of the text.

Mann and Thompson also argued that if all the satellite units from a given text are deleted but all the nuclei remain, we should be left with a coherent text with a message resembling that of the original; it should be something like a synopsis or summary of the original text. As we see later, this observation has been useful in recent work on summarization.

In RST, schemas define the structural constituency arrangements of text. They are abstract patterns consisting of a small number of con-
stituent text spans, a specification of the relations between them, and a specification of how certain spans (nuclei) are related to the whole collection. Schemas are thus loosely analogous to grammar rules, and they constrain the possible RST structures. A graphical depiction of one schema defined by Mann and Thompson (1988) appears in Fig. 12.9. This schema consists of a nucleus and two satellites: one providing MOTIVATION for the material in the nucleus, and the other providing ENABLEMENT for the material in the nucleus.

RST schemas are recursive: text serving as a nucleus or satellite in one schema may itself be described by a schema that can be further decomposed into spans related in one of the ways dictated by the schemas. As defined by Mann and Thompson (1988), the schemas do not constrain the ordering of the nucleus and satellites, and each constituent relation may occur any number of times within a schema.

For example, the following text is described by the schema depicted in Fig. 12.9:

(9) a. Come to the party for the new president.
   b. There will be lots of good food.
   c. The Fluted Mushroom is doing the catering.
   d. The party is in the ballroom at eight o’clock on Friday.

In this example, (9a) is the nucleus of the entire text, and it presents an action that the speaker wishes the hearer to perform. (9b–c) presents information intended to increase the hearer’s desire to perform the action, and is therefore a satellite related to (9a) by the MOTIVATION relation. (9b–c) is further decomposed into a nucleus, (9b), and a satellite, (9c), which in this case are related by EVIDENCE because (9c) is intended to increase the hearer’s belief in (9b). In (9d), the speaker provides information intended to increase the hearer’s ability to perform the action in the nucleus, and thus (9d) is a satellite span related to (9a) by the ENABLEMENT relation. The RST analysis of 9 is shown in Fig. 12.10.
To be acceptable, an RST analysis of a text must meet several additional criteria. It must be complete and connected (i.e., there must be one schema application under which the entire text is subsumed and all minimal units of the text must be accounted for in the analysis). In addition, each minimal unit can appear in exactly one schema application, and the spans constituting each schema application must be adjacent in the text. These constraints guarantee that a correct RST analysis forms a tree structure.

As we describe in more detail later, RST has been used extensively by researchers in text generation. More recently, RST has been used as the basis for rhetorical parsing, which has been applied to the problem of text summarization (Marcu, 1999). Although it did not correctly identify as many of the rhetorical relations as humans do (47% compared with 83%), the relations that it did identify were mostly correct (78% compared with 83% for humans; Marcu, 1999).

### Segmented Discourse Representation Theory

The three theoretical approaches presented so far each focus on different aspects of what makes discourse coherent. Recently Asher and Lascarides developed a theory that combines the logic-based structures of DRT with the focus on rhetorical relations from RST to address a wide range of discourse phenomena. This theory, called Segmented Discourse Representation Theory (SDRT), started in Asher (1993) and has been further developed in Asher and Lascarides (1995) and Asher and Lascarides (1998, 2003).

In DRT, the discourse update procedure that joins the DRSs of two utterances together consists simply of appending the two structures.
Any unresolved references to, for example, pronouns are resolved structurally (i.e., all available [not embedded] discourse referents are potential antecedents). Thus, DRT overgenerates antecedents; it allows coreferences that humans would never consider for a variety of discourse-related reasons. The connections are made on the basis of structure alone, not on the content.

SDRT greatly expands the power of the discourse update procedure by including rhetorical relations. Every time a DRS for a new utterance is added, some relation must be computed between it and one of the preceding utterances. The set of relations is open ended, but includes Narration, Elaboration, Continuation, Background, Explanation, Result, Evidence, Parallel, and Contrast. The relations are derived from theories in the field of pragmatics (e.g., Grice, 1957).

For each new utterance in a dialogue, a DRS is created in the same way as described previously. When it is added to the structure for the dialogue (the discourse update procedure), there must be some link established via a rhetorical relation with a preceding utterance. The inclusion of the relation constrains how the preceding discourse utterances can be accessed. Thus, the set of possible antecedents is not just based on structure, it is based on the pragmatically preferred reading.

For example, Fig. 12.11 shows an SDRS created from the discourse in (10), if $K_1$ to $K_6$ are DRSs that represent respectively the content of the utterances (10a) to (10e):

(10) a. Andrew’s family had a great time at the beach last week.  \hspace{1cm} 1
   b. They went snorkeling. \hspace{1cm} 2
   c. They saw a starfish. \hspace{1cm} 3
   d. They saw a grouper fish too. \hspace{1cm} 4
   e. Then they had dinner at a cafe on the beach. \hspace{1cm} 5
The symbols label subcomponents (representations of utterances) of the discourse. Relationships between them are given in the same way as normal DRT conditions—for example, Narration(2, 5) in this case. The key benefit of SDRT is that the specification of the relationships between the utterances constrains further processing—for example, anaphora resolution. If discourse (10) were continued with the sentence “It was delicious,” the pronoun it could not refer to the grouper fish because its representation is embedded within a substructure.

We know of no current implementation of SDRT. It was previously used in a project for analyzing air traffic control conversations in French (Asher, Aurnague, Bras, & Vieu, 1995). It is the center of some recent research in natural language generation (Danlos, Gaiffe, & Roussarie, 2001) and in dealing with sentence fragments in discourse (Schlangen, Lascarides, & Copestake, 2001).

GENERATING COHERENT DISCOURSE

As noted earlier, a comprehensive survey of discourse structure for natural language understanding appears in Grosz et al. (1989). Thus, here we focus on the role of discourse in natural language generation.

Content Selection and Organization

Early work in natural language generation (Appelt, 1985; Cohen, 1978; Cohen & Perrault, 1979) focused on generating utterances that would allow a hearer to recognize the speaker’s intention to perform a particular speech act. These systems formalize the preconditions and effects of illocutionary acts, and reason about the beliefs of the hearer and speaker and the effects of speech acts on these beliefs. This approach explicitly represents the relation between the speaker’s intentions and the speech acts that achieve them. This is a necessary component of any system that must participate in a dialogue with its users. However, these systems could generate only short (one- or two-clause) texts; they do not represent or use knowledge about how speech acts may be combined into larger bodies of coherent text to achieve a speaker’s goals.

To build systems capable of producing larger bodies of text, researchers sought approaches that would enable systems to produce texts that adhere to standard patterns of discourse and flow smoothly from topic to topic. Until recently, two main approaches to the generation of connected discourse were prevalent: graph traversal and schema based. Graph traversal produces texts whose structure mir-
rors the structure of the domain knowledge representations being traversed, whereas schema-based approaches use domain-independent rhetorical strategies to select information and impose a structure on the text. Both approaches incorporate focus considerations to determine what to say next when several options are available.

Graph Traversal: Paraphrasing the Knowledge Base

By studying a range of naturally occurring texts, computational linguists identified types of text for which discourse structure closely follows the conceptual relations (e.g., causal, temporal, spatial) connecting the domain entities being described. For example, Paris (1988) observed that one strategy for describing a complex physical object is to trace through the process that allows the object to perform its function. Sibun (1992) found that texts describing the layout of houses or the relations among members of a family also follow domain relations. When text structure follows domain structure, the most appropriate generation mechanism selectively traverses existing links in a knowledge base (Suthers, 1991). To generate text, graph traversal is typically combined with a focusing mechanism, which appeals to some model of attentional structure to select the most appropriate thing to say next when multiple domain links are available.

Uses of this technique are limited to cases where the system's representation of the domain mirrors or can be engineered to mirror the structure of natural utterances in the domain. This requirement may place an undue burden on knowledge engineers, who are already trying to mediate between the sometimes conflicting demands of writing programs that are correct, efficient, and maintainable (Swartout, 1983). Moreover, the approach does not admit communicative strategies that depart from the organization of the knowledge base, thereby restricting the types of texts that a system can produce. Finally, because the approach does not model the effects of its utterances, it does not provide a representation of the text from which a system can determine how to interpret and respond to feedback from the user.

Schemas: Representing Patterns of Rhetorical Structure

Some of the limitations of traversal techniques can be overcome by using domain-independent rhetorical strategies that control both what is said and how it is structured. Many researchers have attempted to understand the nature of the relationships that hold between the utterances of a coherent discourse (Lehnert, 1981; Polanyi,
1988; Reichman-Adar, 1984; Rumelhart, 1975). At the heart of many of these efforts is a set of rhetorical predicates or relations that characterize the organization of coherent texts of the type studied. We have seen that many linguists and computational linguists have proposed classifications of these relations and attempted to identify their functions (Hobbs, 1983; Grimes, 1975; Lascarides & Asher, 1993; Longacre, 1983; Mann & Thompson, 1988).

McKeown (1985) found a correlation between the discourse purpose of a text and patterns of rhetorical predicates in the text, with a family of similar patterns representing a strategy for achieving a given discourse purpose. McKeown encoded these patterns of rhetorical predicates into script-like (Schank & Abelson, 1977) structures called schemas. By associating each rhetorical predicate with an access function for an underlying knowledge base, these schemas can be used to guide both the selection of content and its organization into a coherent text to achieve a given communicative goal. The schema-based and other approaches based on rhetorical structuring (Hovy, 1988b) have proved successful for many text generation applications.

The graph traversal and rhetorical structuring approaches to content determination are not mutually exclusive. For example, the Intelligent Labelling EXplorer (ILEX; Mellish et al., 1998) uses a combination of these techniques. In ILEX, facts about the domain are organized into a text potential—a graph of facts interconnected in terms of thematic and rhetorical relations representing the information ILEX can express and the ways it can be related. Given an entity to describe, ILEX finds all of the facts associated with that entity and selects among these opportunities for what to say next using a set of heuristics. Once a fact is chosen, all of its connections are examined, the next fact is chosen, and so on.

The schema-based approach has several advantages. First, because it decouples discourse strategies from low-level details of knowledge representation, knowledge engineers have more flexibility to design knowledge bases to satisfy other desiderata, such as maintainability and run-time efficiency. Second, discourse strategies based on rhetorical knowledge enable systems to generate a range of different texts from the same knowledge representation.

However, Elhadad (1996, 1997) has argued the need to go beyond schemas to produce effective argumentation. Moore (1995) has shown that schemas and rhetorical structure trees are insufficient as a discourse model for dialogue systems because they do not include a representation of the intended effects of the components of the text produced, nor how these intentions are related to one another or to the informational structure of the utterances produced. A schema can be
viewed as the result of a compilation process that dispenses with the rationale for all of its component steps. What remains is the top-level communicative goal that invoked the schema and the sequence of actions (i.e., instantiated rhetorical predicates that cause sentences to be generated) that are used to achieve that goal. Because of this compilation, schemata provide a computationally efficient way to produce multisentential texts for achieving discourse purposes. They are rhetorical recipes that encode frequently occurring patterns of discourse structure. Using schemata, the system need not reason directly about how speech acts affect the beliefs of the hearer and speaker, nor about the effects of juxtaposing speech acts. The system is guaranteed that each schema will lead to a coherent text that achieves the specified discourse purpose.

However, this compilation renders the system incapable of responding appropriately if the hearer does not understand or accept the utterance. Because intentional structure has been compiled out of the schema representation, the system cannot determine whether any of the discourse actions in the recipe have failed to achieve their intended effects or what other strategies could be used to achieve those effects. Intentional structure is crucial for interpreting and responding to questions or challenges that address a previous utterance: Without a record of what an utterance was intended to achieve, it is impossible to elaborate, clarify, or defend that utterance. This is because there is not a one-to-one mapping between intentional and informational structure. That is, there is not a one-to-one mapping between the ways in which content can be related in a discourse and the ways in which intentions combine into a coherent discourse plan to affect a hearer’s mental state (Moore & Paris, 1993; Moore & Pollack, 1992). Thus, it is impossible to recover intentional structure from informational structure or vice versa. Therefore, it is not possible to reconstruct the intended effects of individual actions in an instantiated schema, which contains only propositions and rhetorical relations between those propositions.

**Plan-Based Approaches**

To overcome the limitations inherent in schema-based approaches, researchers have applied techniques from AI planning research to the problem of constructing discourse plans that explicitly link communicative intentions with communicative actions and the information that can be used in their achievement (Moore, 1995; Young et al., 1994a). Text planning generally makes use of plan operators—discourse ac-
tion descriptions that encode knowledge about the ways in which information can be combined to achieve communicative intentions:

- **effect(s):** communicative goal(s) the operator is intended to achieve.
- **preconditions:** conditions that must hold for an act to successfully execute. For example, it may be the case that the hearer must hold certain beliefs or have certain goals for a particular discourse strategy to be effective.
- **constraints:** specifications of the knowledge resources needed by the discourse strategy.
- **subplan:** optionally, a sequence of steps that implement the discourse strategy.

Simplified examples of typical discourse planning operators, taken from Young and Moore (1994), are shown in Table 12.2. In this framework, the representation of communicative action is separated into two types of operators: action and decomposition. Action operators capture the conditions (preconditions and constraints) under which an action can be executed and the effects the action achieves if executed under the appropriate conditions. Preconditions specify conditions that the agent should plan to achieve (e.g., the hearer knows a certain term), whereas constraints specify conditions that the agent should not attempt to plan to change (e.g., facts and rules about the domain). Effects describe the changes that a discourse action is intended to have on the hearer’s mental state. If an action is composite, there must be at least one decomposition operator indicating how to break the action down into more primitive steps. Each decomposition operator protoc183

<table>
<thead>
<tr>
<th>TABLE 12.2</th>
<th>Discourse Plan Operators</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Operator 1:</strong> Action operator for <code>Cause-to-Bel</code></td>
<td></td>
</tr>
<tr>
<td>HEADER: <code>Cause-to-Bel(?p)</code></td>
<td></td>
</tr>
<tr>
<td>CONSTRAINTS: <code>not(Bel(?p))</code></td>
<td></td>
</tr>
<tr>
<td>PRECONDITIONS: <code>nil</code></td>
<td></td>
</tr>
<tr>
<td>EFFECTS: <code>Bel(?p)</code></td>
<td></td>
</tr>
<tr>
<td><strong>Operator 2:</strong> Decomposition operator for <code>Cause-to-Bel</code></td>
<td></td>
</tr>
<tr>
<td>HEADER: <code>Cause-to-Bel(?p)</code></td>
<td></td>
</tr>
<tr>
<td>CONSTRAINTS: <code>nil</code></td>
<td></td>
</tr>
<tr>
<td>STEPS: Begin, Inform(?p), Support(?p), End</td>
<td></td>
</tr>
<tr>
<td><strong>Operator 3:</strong> Decomposition operator for <code>Support</code></td>
<td></td>
</tr>
<tr>
<td>HEADER: <code>Support(?p)</code></td>
<td></td>
</tr>
<tr>
<td>CONSTRAINTS: <code>causes(?q, ?p)</code></td>
<td></td>
</tr>
<tr>
<td>STEPS: Begin, <code>Cause-to-Bel(?q)</code>, <code>Cause-to-Bel(causes(?q, ?p))</code>, End</td>
<td></td>
</tr>
</tbody>
</table>
vides a partial specification for a subplan that can achieve the action's effects, provided the preconditions are true at the time the steps in the decomposition are executed.

As an example of how action and decomposition operators are used to encode discourse actions, consider operators 1 and 2 in Table 12.2. These two operators describe the discourse action Cause-to-Bel, which is capable of making the hearer believe a proposition. Operator 1 is an action operator; it indicates that Cause-to-Bel can be used to achieve the state where the hearer believes a proposition \( ?p \), if the hearer does not already believe \( ?p \). Operator 2 in Table 12.2 is one decomposition operator for the Cause-to-Bel action. It says that one plan for making a hearer believe a proposition is to inform the hearer of that proposition and provide support for that proposition.

As illustrated in operator 3 of Table 12.2, decomposition operators may also have constraints, which indicate the conditions under which the decomposition may be applied. Such constraints often specify the type of information needed for particular communicative strategies, and satisfying them causes the planner to find content to be included in its utterances. For example, operator 3 encodes the discourse strategy that one way for a speaker to support a proposition is to describe a plausible cause of that proposition. More specifically, the constraint on operator 3 says that to support proposition \( ?p \), there must be another proposition, \( ?q \), such that \( \text{causes}(?q, ?p) \) is true in the domain. When the planner attempts to use a decomposition operator to support a proposition, it must try to satisfy all of its constraints. If a constraint contains no unbound variables, it is simply checked against the knowledge source to which it refers. However, if the constraint contains free variables (e.g., \( ?q \) in operator 3), the system must search its knowledge bases for acceptable bindings for these variables. In this way, satisfying constraints directs the planner to select appropriate content to include in explanations. In operator 3, if an appropriate \( ?q \) can be found, then the speaker can support \( ?p \) by making the hearer believe \( ?q \) and making the hearer believe that \( \text{causes}(?q, ?p) \). Thus, we see that action and decomposition operators specify how information can be combined in a discourse to achieve effects on the hearer's mental state. That is, action operators and their decompositions encode the link between intentional and informational structure.

A detailed description of the algorithm for synthesizing plans from such operators is beyond the scope of this chapter and may be found in (Young & Moore, 1994; Young, Pollack, & Moore, 1994b; Moore, 1996).

3Note that \( \text{not}(\text{Bel}(?p)) \) is a constraint because we do not wish our discourse planner to attempt to plan to make the hearer not believe \( ?p \) to use this operator.
However, it is important to note that the plans produced from such operators can serve as the speaker’s model of the effects that individual parts of the text were intended to have on the hearer and how they fit together to achieve her top-level intention. From a text plan, we can determine which intentions dominate other intentions (i.e., we can determine when an intended action is in the plan to serve a higher intention). This allows the speaker to interpret and respond to feedback indicating that the hearer does not fully understand or accept what the speaker says and localize the failure in some portion of the utterance that failed to achieve its intended purpose.

To illustrate this, we consider an utterance from a naturally occurring tutoring dialogue and see how reasoning about the text plan that produced the utterance could enable a system to respond appropriately to a range of student responses. Consider the dialogue fragment shown in Table 12.3, which was taken from a corpus of student–tutor interactions in which students are using a computer simulation environment that trains them to troubleshoot the complex electronic circuitry found in aircraft. The tutoring system presents the student with a troubleshooting problem to solve, allows the student to solve the problem with minimal tutor interaction, and then provides a critique of the student’s solution in a postproblem review session. During the review session, the system replays the student’s solution step by step, pointing out good aspects of the student’s solution as well as ways in which the student’s solution could be improved.

In turn 1 of the dialogue in Table 12.3, the tutor indicates that the problem-solving step of replacing a particular component was sub-optimal (as indicated by < — > preceding the step description). The student does not immediately accept this assessment and probes the tutor to find out why this action was assessed negatively, turn 2.

In turn 3, the tutor explains why the student’s action was judged suboptimal, with the following utterance, repeated here for convenience:
(11) P1. It would have been better to troubleshoot this card by taking measurements instead of swapping.
   P2. You can’t count on having spare cards available for swapping.

Our analysis of this example is that the tutor’s primary intention is to convince the student of P1—that it is better to troubleshoot a component by taking measurements than to swap the component. To achieve this goal, the tutor asserts this proposition and then supports it by claiming P2—that the student cannot always count on having spare parts available for swapping. The plan for generating this text is shown in Fig. 12.12.

To handle feedback indicating that the hearer does not fully understand or agree with what the speaker says, the speaker must be able to determine what portion of the utterance failed to achieve its intended purpose. Therefore, the speaker must have a model of the effects that individual parts of the text were intended to have on the hearer and how they fit together to achieve the speaker’s top-level intention. From the plan in Fig. 12.12, we can determine that the speaker’s intention to make the hearer believe that one cannot count on having spare cards available for swapping serves the higher level intention of making the hearer believe that it would have been preferable to troubleshoot the card by taking measurements.

Now let us consider several possible student responses to the tutor’s utterance in turn 3 of the sample dialogue, given in examples (12)–(14), and see how a computer system, acting as tutor, could use this discourse plan to help determine appropriate responses in each case.

(12) We always have spare cards in our shop.
(13) Yeah, it would have been better to troubleshoot the card, but we always have spare cards in our shop.
(14) Yeah, you can’t count on having spares, but it’s still better to swap.

In example (12), the student rejects P2—the proposition that spare cards may be unavailable. This blocks the support that P2 would have provided to convince the hearer that troubleshooting by taking measurements is a better strategy than swapping. To see how the system can determine this, consider again the discourse plan in Fig. 12.12. Response (12) indicates that the effect Bel(P2) asserted by the Inform(P2) was not achieved. From the plan representation, it is possible to trace a path of failed effects from Bel(P2) across causal links and up decompositional links through the actions Inform(P2), Cause-to-
Bel(P2), Support(P2), Cause-to-Bel(P1), and eventually to the top-level intended effect Bel(P1). Using this information, the system can determine that appropriate responses to (12) can be generated by trying to convince the student that spare cards are not, in fact, always available (i.e., replanning the subtree rooted at the node Cause-to-Bel(P2), most likely by providing support for P2) or by finding some other support for the claim that troubleshooting by taking measurements is a better strategy than swapping (i.e., replanning the subtree rooted at the node Support(P1)).

An appropriate response to (13) would be different. In (13), the hearer again expresses disbelief in the supporting proposition P2 (i.e.,

\[ P1 = \text{It would have been better to troubleshoot by taking measurements instead of swapping.} \]
\[ P2 = \text{You can’t count on having spare cards available for swapping.} \]
\[ G = \text{to troubleshoot the test station} \]
\[ A = \text{taking measurements} \]
\[ B = \text{swapping components} \]
the effect Bel(P2) asserted by the Inform(P2) was not achieved). However, here the student gives more information about the success of the speaker’s original plan by indicating that he believes it would have been better to take measurements than to swap—that is, here the effect Bel(P1) has been achieved. As in (12), the speaker’s intention to get the hearer to believe P2 has failed, and thus the support that P2 would have provided for P1 is again blocked. However, in (13), the tutor need not do anything about this failure. This is because the intention to get the hearer to believe P2 was held in service of the intention to get the hearer to believe P1. Since (13) explicitly indicates that Cause-to-Bel(P1) has achieved its intended effect—namely, Bel(P1)—the outcome of any intended act that served as a precondition to this intention or as a step in a subplan for this intention can be ignored—unless of course the tutor has some other reason for wanting the student to believe P2.

Now consider how to handle (14), where the hearer agrees with P2, but is not convinced of P1. Implicit in the speaker’s original argument was his belief that, as a rule, not being able to count on having spare cards makes troubleshooting by taking measurements a preferable strategy to swapping. That is, the discourse plan in Fig. 12.12 is predicated on the truth of Bel(causes(P2,P1)). Note that the node Cause-to-Bel(causes(P2,P1)) was not expanded in this discourse plan because this proposition was an effect of the initial step. Together with the statement in P2, causes(P2,P1) would have provided support to convince the hearer that it is better to troubleshoot before swapping. (14) indicates that the support for P1 has failed. At this point, the tutor must either convince the student that the causal relationship between P2 and P1 does indeed hold or must find another way to support P1. As in the previous case, an appropriate response results from the replanning of subtrees whose execution is affected by this failure. Specifically, those subtrees rooted at Cause-to-Bel(causes(P2,P1)) across causal links and up decompositional links to Support(P1) and eventually to Cause-to-Bel(P1). Note that this does not include the subtree rooted at Cause-to-Bel(P2) and thus, unlike in (12), the system will behave correctly and will not attempt to reestablish P2—the proposition that spare cards may not always be available.

In these examples, each of the hearer’s replies provides feedback about a subset of the speaker’s intentions. To respond appropriately, the speaker must reason about the relationships between the intentions in his communicative plan to determine what implication the hearer’s feedback has on the ultimate success of his other intentions. These examples show that the information in discourse plans provides guidance for the planning of subsequent responses.
Note that discourse plans in this framework also include a representation of informational structure. When a discourse strategy requires that a particular informational relation be recognized by the hearer for a discourse to have its intended purpose, a condition expressing this requirement appears in the (sub)plan that requires this. For example, in Fig. 12.12, the strategy for supporting P1 is successful only if the two effects Bel(P2) and Bel(causes(P2,P1)) are achieved. Conditions such as this one allow the planner to recognize how the content expressed in portions of the text plan is related. Among other tasks, the informational structure is used by the realization component when transforming the discourse plan into a series of natural language utterances. The representation of informational structure, together with information about intentional structure, allows the text generator to order clauses and select appropriate content-bearing discourse cues (e.g., because, in addition).

Discourse plans such as these, which have been used in many systems, are hierarchical structures whose leaves specify a partially ordered sequence of speech acts to be performed. The internal structure of the plan represents dominance and satisfaction precedence relations between discourse intentions, and steps in subplans post goals to make the hearer recognize informational relations between plan components. Although these plan structures contain much information that is crucial for generating coherent multisential natural language texts, they cannot be transformed directly into natural language. They do not include all of the information required by existing syntactic realizers, which transform abstract syntactic specifications of natural language sentences (or phrases) into their corresponding surface forms. Examples of such systems are FUF (Elhadad, 1992) and REALPRO (Lavoie & Rambow, 1997).

To construct specifications from which English sentences can be generated, many decisions about organization and lexicalization remain. A system must choose a total ordering for the steps in the discourse plan and decide how to apportion propositional content to sentences and sentence constituents (Hovy, 1988a; McDonald, 1983; Meteer, 1992). It also must choose referring expressions (Appelt, 1985; Dale, 1992; Reiter, 1990) and lexical items to express the content in the speech acts (Bateman & Paris, 1989; Danlos, 1984; Granville, 1984; Matthiessen, 1991; McDonald, 1991; Pustejovsky & Nirenburg, 1987; Reiter, 1991). As the references indicate, there has been considerable research on many of these issues.

A remaining issue, and one that has received relatively little attention in the computational generation literature, concerns the use of discourse cues. Discourse cues are words or phrases, such as be-
cause, first, although, and also, that mark structural and semantic relationships between discourse entities. They play a crucial role in many discourse processing tasks, including plan recognition (Litman & Allen, 1987), text comprehension (Cohen, 1984; Hobbs, 1985; Mann & Thompson, 1986; Reichman-Adar, 1984), and anaphora resolution (Grosz & Sidner, 1986). Moreover, research in reading comprehension indicates that felicitous use of cues improves comprehension and recall (Goldman, 1988), but that indiscriminate use of semantic cues may have detrimental effects on recall (Millis, Graesser, & Haberlandt, 1993). In addition, there is evidence that the benefit of discourse cues may depend on the subjects’ reading skill or level of prior knowledge of the domain (Goldman & Murray, 1992; Meyer, Brandt, & Bluth, 1980; McNamara, Kintsch, Songer, & Kintsch, 1996).

The problems of determining when discourse cues should be used in the final text, where the cues should be placed, and which cues would be most effective in increasing the hearer’s comprehension of a text are a current area of research. McKeown and Elhadad studied several connectives (e.g., but, since, because) with the aim of identifying features of the propositions connected by the cues that can be used to select appropriate cues during text generation (Elhadad & McKeown, 1990; McKeown & Elhadad, 1991). Researchers concerned with generating text from RST trees (where leaf nodes contain content and internal nodes indicate the RST relation that holds between subtrees) have proposed algorithms for determining sentence boundaries and choosing cues based on the rhetorical relation between spans of text, the order of the relata, and the complexity of the related text spans (Rösner & Stede, 1992; Scott & de Souza, 1990).

As noted earlier, RST analyses presume that there is a primary rhetorical relation between any two consecutive elements of a coherent discourse, and RST analyses do not permit multiple relations between text spans. This means that consecutive elements in RST trees are related either by an informational or an intentional relation. However, Moore and Pollack (1992) showed that discourse interpretation and generation cannot depend on analyses in which informational and intentional structure are in competition; intentional and informational analyses must coexist. Therefore, we must devise algorithms for generating appropriate texts from a discourse model that represents these two types of structure, such as the discourse plans described earlier.

### Intentional and Informational Structure and Cue Choice

It should be clear that the informational (or semantic) relation between discourse items affects the choice of cue. For example, to mark an exemplification relation, a speaker can felicitously use a cue such as for
example or for instance, but could not use a cue such as thus or first. It is less clear to many how intentional structure affects cue usage. Note here that we are concerned with those cues that convey semantic information between discourse elements, such as because, thus, or for example. It is clear that intentional structure affects cues that indicate purely structural aspects of discourse, such as topic shifts (e.g., now, anyway) and digressions (by the way; Grosz & Sidner, 1986; Hirschberg & Litman, 1993).

To illustrate the effect of intentional structure on cue choice, let us consider the two example discourses in (15) and (16) in which the informational relation between discourse entities and the placement of the cue is held constant, but the intentional structure varies. Figure 12.13 shows the intentional and informational relations between two discourse actions and the text that these actions produce. In this example, the tutor is trying to convince the student of (15b)—that there is a break in a certain signal path. To achieve this goal, the tutor informs the student of (15b) and supports it with (15a). In the domain, there is a causal connection between (15a) and (15b), the bad signal at pin 26 causes there to be a break in the particular path discussed. Thus, the tutor can use the discourse strategy (encoded in operator 3) of supporting a proposition by describing a plausible cause of that proposition. Figure 12.13 represents a typical deductive argument; to convince the hearer of an effect (15b), cite its cause (15a) as support.

In example 15, the intention to make the student believe (15b) dominates the intention to make the hearer believe (15a). At the informational level, there is a causal relation between (15a) and (15b). In the text, the proposition expressing the cause (and the dominated discourse purpose) precedes the one expressing the effect (and the dominating discourse purpose). The discourse cue, thus, is placed with the

![Diagram](attachment://diagram.png)

(15) a. You know that the signal on pin 26 is bad.
   b. Thus, there's a break in the path created by TPA63.

FIG. 12.13. Arguing from cause to effect.
proposition expressing the effect. This cue indicates both the causal relation at the informational level as well as the dominance relation between the speaker’s intentions.

In contrast, consider the relation between the discourse actions and the accompanying text in Fig. 12.14. In this example, the tutor is trying to convince the student of (16a)—that the signal on pin 26 is bad and is using (16b) as support. That is, the speaker is trying to convince the hearer that a state exists by citing an effect of that state. This is a typical abductive argument. In this example, the informational relation between (16a) and (16b) is the same as in example (15)—that is, (16a) causes (16b). However, the two texts differ at the intentional level. In (16), the intention to make the hearer believe (16a) dominates the intention to make the hearer believe (16b). This difference in intentional structure is reflected in the discourse cue chosen. As in example (15), the cause precedes the effect, and the discourse cue is placed with the text expressing the effect. However, a different discourse cue (because) must be used to indicate the difference in intentional structure. In (16), the intentional roles of cause and effect are reversed. The proposition expressing the cause is now expressing the dominating discourse purpose, and the one expressing the effect is now expressing the dominated discourse purpose. The cue is now placed with the proposition expressing the dominated discourse purpose. Since the causal relation at the informational level has remained unchanged from example (15), the difference in cue must be due to the difference in intentional structure.

These examples show that algorithms based on a discourse model that forces a choice between intentional and informational structure, such as RST, cannot be complete. Algorithms for cue usage must take

\[
\text{Cause-to-Bel(a)} \\
\text{cause(a,b)} \\
\text{Cause-to-Bel(b)}
\]

(16) a. You know that the signal on pin 26 is bad  
b. because there's a break in the path created by TPA63.

both informational and intentional structure into account. Moser and Moore (1995, forthcoming) have done an extensive empirical study to identify the factors that predict appropriate use of cues. Di Eugenio, Moore, and Paolucci (1997) are using machine learning to induce decision trees that can be used to determine cue choice in automatic text generation.

**CURRENT DISCOURSE APPLICATIONS**

In this section, we describe some of the new directions in computational linguistics and AI in processing discourse. Many of the current techniques break with the theoretical traditions described in previous sections. Instead, they rely on shallow text processing techniques and statistical methods that support the inference of discourse information in a task-specific or domain-specific way.

**Summarization**

As its name suggests, the goal of a text summarization system is to produce a summary of a text that can be quickly read by a user. Because of the huge amounts of textual data available on the Web and elsewhere, text summarization can provide a great benefit to those who need to scan or stay current in a topic, but care only about the main points and not the details.

Most current systems perform summarization in three steps:

1. Identify the important text units of a document (or set of related documents).
2. Extract from each unit the most central sentence or sentences.
3. Join them to form the summary.

This section describes how discourse processing techniques are used to perform these steps in a variety of summarization systems.

A critical first step in determining which parts of a document are most important is locating the boundaries between topic segments. This allows a summarization system to know which topics should be represented in the summary, and prevents the system from misreading important text on a new topic as less important text continuing the current topic.

The standard technique for performing automatic text segmentation is to use some measure of sentence similarity to find consecutive clusters of sentence that have something in common. The usual simi-
Similarity metrics are based on word overlap, typically by calculating word stems, and then comparing sentences with either a keyword overlap mechanism or vector-based formula. Then some type of clustering algorithm is used to detect boundaries between groups of consecutive sentences that are highly similar (Choi et al., 2001; Hearst, 1997).

Another approach is to use the pattern of word occurrences across the text as an indication of topic segments. Lexical chaining was introduced by Morris and Hirst (1991). For each reference to an entity in the text, a graph is created that follows references to the same or closely related terms. The related terms are inferred from synonym sets provided by a lexical database like WordNet (Fellbaum, 1998). The chains show which terms are essential to the text as a whole (longer chains represent more significant terms) and locate where those terms are mentioned (Barzilay & Elhadad, 1997; Hirst & St-Onge, 1998; Stairmand, 1996).

A genre-specific segmentation method was developed by Teufel (1999). For scientific papers, she used a machine learning technique to associate a variety of discourse cues with argumentative zones—that is, sections of the papers with different functional purposes. Two examples of argumentative zones are general scientific background and descriptions of other people’s work. These zones provided the basis for performing topic segmentation of the texts.

Marcu (2000) used a similar technique that works on unrestricted texts. His approach, rhetorical parsing, also used machine learning to determine the rhetorical relations based on a variety of shallow cues such as discourse cues and punctuation. For example, the use of the cue term although indicates either a Concession or an Elaboration relation with a neighboring textual unit. Marcu demonstrated that the hierarchical rhetorical trees that this technique produces are useful for text summarization because they highlight topic switches in the text. The rhetorical relations also indicate the central text units of each section. Systems that do not perform rhetorical parsing tend to rely on word overlap measures to determine which text units are most central within a section.

The key sentences of a segment cannot normally be used directly in a summary because of anaphoric references. Coreference resolution must be performed to replace abbreviated references with their fully specified form. For example, a document might refer to the same company as “Apple Computer Inc.,” “Apple,” “the computer company,” “the company,” and “it.” Because the first reference may not be in the sentence that contains the most central information, coreference resolution is essential. As mentioned earlier, the various computational theories of discourse structure each have something to say about the con-
straints on coreference. Because most summarization systems do not perform in-depth processing of the texts, they tend to use domain-specific shallow methods to track coreference.

Coreference resolution has been the focus of many current corpus-based applications. Lappin and Leass (1994) described a model for resolving pronominal anaphora using a surface parse of the sentences and a heuristic measure of salience of potential referents based on features such as their proximity and syntactic position. This approach was extended by Kennedy and Boguraev (1996) to use an even shallower syntactic representation in which words were marked with their syntactic function, but no parse tree was required. Other researchers have used machine learning approaches like Bayesian induction and decision trees to learn methods for coreference resolution from annotated corpora (Aone & Bennett, 1995; Connelly, Burger, & Day, 1994; Kehler, 2000; McCarthy & Lehnert, 1995). A recent DARPA-sponsored information extraction initiative (Sundheim, 1995) had a subtask that required participants to resolve coreference among proper names, aliases, definite noun phrases, and more.

As in other natural language processing tasks, performance is normally calculated by comparing with human judgments and reported in terms of recall, precision, and Fscore. Recall is the number of true positives divided by the sum of the true positives and the false negatives. Precision is the number of true positives divided by the sum of the true positives and the false positives. Because there is normally a trade-off between precision and recall, the Fscore combines them and is defined as: \(2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}\). On the coreference task, participants have achieved success rates of over 60% recall, 70% precision, and 65% Fscore (Sundheim, 1995).

Once coreference resolution is performed by a summarization system, the fully specified sentences can then be processed into a coherent and readable summary. Using techniques from natural language generation, references to the same entities can be replaced with anaphoric expressions to increase the cohesiveness of the text. Sentences may also need to be restructured to account for discourse focus (Kan & McKeown, 1999).

Another DARPA initiative has recently compared the performance of several text summarization systems (Mani et al., 1998). Because it is so difficult to determine what an ideal summary would be, the evaluation focused on extrinsic metrics—that is, ones that measure how well the end product of the system enables a human to perform some task. In this case, one task (the ad hoc task) was to determine whether a document was relevant to a particular topic. A good summary allows the human analyst to quickly determine whether the source document is
relevant. In this evaluation, the best systems achieved performance of over 60% recall, 80% precision, and Fscores of around 70%. There was also a categorization task, in which the summaries were used by a human analyst to classify the source document as either fitting into one of five topics or none of the above. Here the scores were lower, with top precision scores of around 70%, recall scores around 45%, and Fscores around 50%.

**Question Answering**

Question answering (QA) is an offshoot of the information retrieval task. In information retrieval, the task is to select from a large database of texts a small number that matches some query, similar to exploring the Internet with a search engine. In QA, the task is more specific: Instead of returning an entire document, the system should return just the specific sentence that answers the question.

The standard approach to this task involves first performing information retrieval to find relevant documents. Then each document is searched for the sentence(s) that are most relevant to the question. The sentences are ranked for relevance and informativeness, and the highest ranking sentence is returned as the answer.

Here also it is clear that discourse information plays a crucial role. For example, in searching for the answer to the question, “Who wrote “To Kill a Mockingbird”?, a system might find the text:

Now Harper Lee is 70, a white-haired woman who protects her privacy as insistently as J.D. Salinger or Thomas Pynchon. She is the author of a single book, “To Kill a Mockingbird,” but that book is one of the most widely read American novels of the century.

A QA system must resolve the coreference to succinctly provide the correct answer. Coreference resolution has been used to increase the performance of a number of recent QA systems (Breck, Burger, Ferro, House, Light, & Mani, 1999; Humphreys, Gaizauskas, Hepple, & Sanderson, 1999; Litkowski, 1999). The systems used a variety of more local techniques, for example, producing variants of the questions. Because these are not discourse related, they are not described here.

In the most recent QA section of the DARPA-sponsored TREC (Text REtrieval Conference) program, the task was to answer a set of approximately 700 fact-based short answer questions by extracting for each a small amount of text (250 bytes) from a 3-gigabyte collection of newswire text. An example question is, “How much folic acid should an expectant mother get daily?” The systems were allowed to provide a
ranked set of answers to each question. The scores were based on how far down the stack of answers the correct answer was or 0 for no correct answer. The best system from among the 28 participants achieved a score of 75% (the correct answer was on average one quarter down the ranked list of answers) and did not answer 15% of the questions.

**SUMMARY**

In this chapter, we have discussed the types of discourse structure that researchers in computational linguistics and AI have developed to address a range of problems in discourse interpretation and generation. In conclusion, we would like to point to several fruitful areas for future research.

First, discourse actions like all other actions have context-dependent effects. Indeed, one of the main aims of discourse research is to understand how the context in which an utterance occurs affects the meaning of that utterance. If we are taking a plan-based view of discourse interpretation and generation, many inferences beyond what are listed as the direct effects of discourse operators may be licensed. In AI, this is the well-known ramification problem. In discourse, some of these inferences, the Gricean conversational implicatures, are an important part of normal cooperative conversation. As we have argued, discourse interpreters must make these inferences if they are to properly understand the discourse, and discourse generators must be aware of the implicatures licensed by their utterances to produce natural sounding utterances and avoid leading the hearer to make false implicatures. Although early research (e.g., Hirschberg, 1985; Joshi, Webber, & Weischedel, 1984) identified the problem and attempted to specify the conversational implicatures licensed by certain types of utterances, there has been no general solution. Recently, there has been renewed interest in this problem, and considerable progress has been made (see Green & Carberry, 1999; Stone, 2001). However, much work remains to be done to provide an adequate model of discourse actions and their effects, which can be used in realistic computational systems.

Second, computational accounts of discourse have largely ignored the impact that human processing limitations must have on language. However, some research has shown that taking resource limitations seriously can provide an explanation of phenomena such as how much risk speakers take of being misunderstood in a given conversational setting (Carletta, 1992) and why speakers sometimes produce redundant utterances (Walker, 1996). Moreover, McKeown’s (1985) schema-based generation system showed how constraints on focus of attention
could affect the content and organization of a discourse. To adequately model human discourse interpretation and generation in computational systems, we must further investigate the interpretation and generation problems as resource-bounded tasks.

Third, we must take seriously a model of belief and context update. Without it, our theories cannot adequately account for why speakers ever do more than simply assert the facts they want their hearers to believe. Most models simply assume that the effect of asserting a proposition P is that the hearer believes P. In fact, a speaker may go to great lengths to convince the hearer of the truth of a proposition. She may first assert it, then support it, and even provide support for that support. In such cases, the speaker presumably believes that the combination of utterances is what leads the hearer to accept the main proposition, and we need to model this behavior. Recent work by Galliers (1990), Carberry and Lambert (1999), and Lascarides and Asher (1999) began to address this issue.

Fourth, more integrated accounts of the relationship between the various types of discourse structure are needed. For example, a more detailed understanding of how speakers’ intentions are realized via informational structure support more principled and effective text and speech generation. Moreover, a more detailed understanding of the relationship between discourse structure at the segment level and the information structure of Vallduvi (1990) and Steedman (1991) is crucial.

Fifth, over the last decade, statistical techniques have greatly improved performance on tasks such as parsing and word sense disambiguation, and probabilistic models are now widely used for a range of language processing tasks (Jurafsky & Martin, 2000; Manning & Schütze, 1999). We are just beginning to see progress in statistical methods for discourse and dialogue, and there is much work to be done to incorporate more sophisticated models of the discourse phenomena that we wish to approximate with statistical methods.

Finally, with the advent of virtual and lifelike animated agents, many new discourse processing tasks are emerging. Our theories need to be broadened to include spoken language, where much discourse information is conveyed by intonation, and to account for the information contained in a speaker’s gestures, which can be used to disambiguate or enrich the meaning of the speaker’s utterances (Cassell, Sullivan, Prevost, & Churchill, 2000).

REFERENCES


12. DISCOURSE IN CL AND AI

ternational Workshop on Natural Language Generation. Niagara on the Lake, Canada.


This page intentionally left blank
For the field of discourse processing to advance, researchers must understand the underlying cognitive processes involved. The creation of models of discourse processing permit researchers to develop predictions based on their understanding of these cognitive processes. For these models to be testable, they must be implemented in a way that they can generate specific quantitative predictions that can be compared against research data. Current quantitative models have proved to be successful in accomplishing these goals. Researchers can run simulations of human discourse processing and test the effectiveness of the models. This permits an improved understanding of how humans process discourse, as well as the development of practical applications that simulate human processing.

There is a wide range of quantitative models that have been applied to discourse, including models based on statistical analyses of the properties of language, models of discourse structure, models of language generation, and models of conversational interaction. This chapter focuses on a subset of these models called cognitive quantitative models. Cognitive quantitative models are defined as models that provide quantitative predictions or explanations of discourse phenomena while using cognitively plausible mechanisms. These mechanisms are general cognitive processes and structures such as working memory, long-term memory, knowledge representations, learning, and problem solving. Thus, the models described must address how...
language is processed within the context of human information processing.

**AN IDEAL COGNITIVE MODEL OF TEXT AND DISCOURSE PROCESSING**

What would an ideal quantitative cognitive model of discourse encompass? It would have to account both for features found in text and discourse and for properties of human information processing. For example, text and discourse features that would be factored into the model include the syntax, semantics, coherence, pragmatics, and rhetorical structures found in a text being read or in a conversation being heard. Properties of human information processing that would be factored into the model would include a person's world knowledge, working memory capacity, ability to make inferences, ability to process metaphors, and knowledge of social rules of conversation. Thus, it would need to effectively model the meaning expressed in a text—the meaning derived by humans and their world knowledge. The interactions among these factors are quite complex, involving the integration of the physical characteristics of the text with the many cognitive and social processing structures. An ideal model would need to account for this integration and complexity. There are numerous questions that such a model would have to address. How do individual cognitive processes affect the processing of discourse? How do these cognitive processes combine to produce comprehension of the discourse? How do the cognitive processes operate across a wide range of types of discourse (e.g., ranging from conversations to narrative texts)? How do social factors within conversational discourse affect the processing?

The study of discourse processing is an interdisciplinary field, so it is important to incorporate theories and techniques from diverse fields. What provides strength to developing quantitative models of discourse is the understanding and contributions from these different, but overlapping fields. Although some fields provide quantitative modeling techniques, other fields provide constraints to these techniques by identifying how language is used and how humans process language. The following are some of the primary fields that provide these contributions:

- Computer science. Artificial intelligence, information-processing algorithms, use of computer simulations.
- Computational linguistics. Parsing, computer-generated semantic/syntactic representations.
• Cognitive psychology. Theories of representation, models of storage, retrieval and processing of information, empirical methods for testing models.

• Linguistics. Descriptive models of language and the use of language.

• Sociolinguistics/sociology. Conversational structure and the effects of social context.

• Rhetoric. Text and discourse structures.

• Education. Approaches to learning from text, instructional design.

Our ideal model would incorporate research from these fields to constrain the design of the model. For example, linguistics, sociolinguistics, and rhetoric provide examples of the types of language that must be processed and constraints on how they are processed. Although not always quantitative in nature, these examples and constraints can be quantified to be applied within quantitative models. Cognitive psychology provides constraints on the cognitive structures used, the representation of information in memory, and how much mental effort can be used for the processing of information. Additionally, quantitative modeling techniques from some of these fields can be incorporated. For example, computational linguistic and artificial intelligence (AI) techniques can be used to develop models of syntactical and semantic processing.

Finally, for our ideal model to be testable, it would have to generate quantitative predictions. Thus, it would have to account for different quantitative measures of discourse and text processing that can be extracted from human studies. These measures would include comprehension (including verbatim recall, textbase, summarization, problem solving, metaphor processing), reading time, ability of parsing texts, speed of retrieval of information from long- and short-term memory, eye movements, and choice of different speech acts under different social conditions. By plugging in a set of parameters representing both the conditions of the text and the abilities of the reader, the ideal model could then generate a complete set of predictions for any of these quantitative measures.

At present, there is no current model that can account for all factors in discourse and cognitive processes. No model is currently able to do a seamless integration of the different fields, nor account for all the diverse quantitative measures of performance. Finally, it is not easy to develop a quantifiable set of parameters that would adequately describe the complete state of a discourse and of a person’s cognitive and social states. However, a variety of the models have been developed re-
recently that still provide effective explanations, either over a wide range
of research findings in human discourse processing or within some
particular components discourse.

MODERN DISCOURSE MODELING

Over the past 25 years, researchers have increasingly focused on the
development of computational models of discourse processing. Sev-
eral factors have contributed to this increase. One factor is the greater
availability of computing power to develop, run, and analyze the mod-
els. This reduces the effort in development and testing, permitting
faster, more accurate modeling. It also allows more complete modeling
in terms of the amount and range of data that can be explained. A sec-
ond factor is the development of techniques for recording discourse
and storing it electronically. For example, the availability of word proc-
ecessors has permitted large corpora of text and discourse to be accessi-
ble for analysis. Finally, discourse processing has benefited from im-
proved models in areas that are relevant to discourse processing, such
as models of memory developed within cognitive psychology.

In earlier cognitive models, researchers were often limited to study-
ing surface-level features of text and discourse, such as syntactic pars-
ing, lexical processing, and verbatim recall. However, with more pow-
erful computing available, there has become a much greater focus on
modeling deeper levels of comprehension and discourse processing
(see Graesser, Swamer, Baggett, & Sell, 1996). For example, newer
models focus on representing a reader’s world knowledge, how a
reader processes inferences or metaphors, and the pragmatic effects
involved in dialogues. Modeling these aspects can account for a much
greater amount of the variance of the data from research on compre-
hension than models that merely focus on surface features. This ap-
proach opens the field to a much wider range of discourse that can be
analyzed. Indeed many of the models reviewed later have been applied
to limited types of text and discourse, but have the potential to be ap-
plied to other types of discourse. Thus, the field has much room for
growth.

There are two perspectives that can be taken in a review of quantita-
tive cognitive models. One perspective is to examine individual models
to see the components and how the models can be applied. The second
perspective is to consider theories of discourse and cognition and then
examine what can be modeled. Although the first perspective yields
greater understanding of individual modeling approaches, the second
perspective yields an understanding of how multiple modeling approaches can explain research results. This chapter takes both perspectives. It first explores some general issues about text and discourse. Then it provides an overview of some of the quantitative models and reviews aspects of discourse and cognitive processes that have been modeled.

**APPROACHES TO MODELING DISCOURSE AND TEXT**

Before examining any models in depth, it is important to introduce some of the major themes and issues found across all models of discourse and text. These common themes address how information is represented in the models, what architectures of cognition are adopted, and what assumptions about human information processing are made.

**Representation in Models**

Comprehension obviously requires access to knowledge of the world. There is a wide range of studies on the integration of discourse and world knowledge (for early studies, see Bransford & Johnson, 1972; Spilich, Vesonder, Chiesi, & Voss, 1979). To model knowledge acquired from discourse, a representational system must be developed to account for a comprehender's semantic or conceptual knowledge structures as well as the verbatim level of the discourse. The critical issue is how to represent meaning. Within cognitive modeling, a number of approaches to representing meaning have been used, including feature systems (see Smith & Medin, 1981), semantic networks (see Collins & Quillian, 1969), and scripts or schemas (see Schank & Abelson, 1977). Although each captures some aspect of meaning in memory, propositional representations have been most widely applied to discourse modeling.

**Propositional Representations.** Propositions can be considered the smallest unit of meaning of discourse. Thus, they can be incorporated into semantic networks, scripts, or schemas. In a propositional representation, discourse meaning is represented as a network of propositions in which each proposition consists of a predicate (typically a verb or adjective) and one or more arguments (nouns or additional embedded proposition; see Kintsch, 1998; van Dijk & Kintsch, 1983). For ex-
ample, the sentence “Yesterday, John threw Sara the red ball” would be represented by the three propositions:

P1  THROW[JOHN, P2, SARA]
P2  RED[BALL]
P3  YESTERDAY[P1]

A propositional representation is designed to capture the meaning rather than the surface features provided by words or sentences. For example, the proposition TALK[MARY, JOHN] could represent any of these sentences: “Mary talked to John,” “It was Mary who talked to John,” and “John was talked to by Mary.”

Research on the psychological reality of propositions has shown that propositions are valid units of comprehension. When given a cue that is from one proposition, subjects are more likely to recall information from within that proposition than from propositions in the same sentences (Wanner, 1975). Further, words from the same proposition prime each other more than words across propositions even if the prime and target words are greatly separated in the surface structure of the sentence (Ratcliff & McKoon, 1978). Manuals that describe techniques for propositionalizing texts can be found in Turner and Greene (1978) and Bovair and Kieras (1985).

Levels of Representation. Many approaches to studying text processing start with identifying the propositions in a text and in a reader’s knowledge. Then these propositions are integrated to form larger, more coherent knowledge structures, such as causal chains or situation models (see reviews in Fletcher, 1994; Kintsch, 1994; Singer, 1994). Therefore, it is important to distinguish the possible levels of representation of the structures of information. van Dijk and Kintsch (1983) proposed three levels of representation. At the base level, the surface form of the discourse represents the actual order of the words of the discourse. At a second level, the textbase is a representation of the network of propositions extracted from the text. Propositions are connected to each other as a network through common arguments in the propositions and other discourse connectives, typically forming hierarchical structures. As a text is being read, new propositions that are present together in working memory are attached to the structure, building a coherent knowledge structure (Fletcher & Bloom, 1988). Studies of reading have shown better comprehension and increased reading time for propositions at higher levels in the hierarchy of the textbase (Britton, Meyer, Hodge, & Glynn, 1980, Cirilo & Foss, 1980; Kintsch & Keenan, 1973; Meyer, Brandt, & Bluth, 1980).
A textbase is limited in that it is only the reader’s representation of what was seen in the text without considering the reader’s knowledge. The situation model is a more complete structure combining the textbase with propositions representing the reader’s knowledge from long-term memory. It serves as a general mental representation or mental model of a person’s understanding of a situation described by the discourse. This representation can include information about characters’ goals, actions, events, and space. Extensive evidence exists that readers construct situation models from text (see Anderson, Garrod, & Sanford, 1983; Fletcher & Chrysler, 1990; Glenberg, Meyer, & Lindem, 1987; Millis, King, & Kim, 2000; Morrow, Bower, & Greenspan, 1989; Zwaan, Magliano, & Graesser, 1995). In addition, the three levels (surface form, textbase, and situation model) can be empirically distinguished from each other (Fletcher & Chrysler, 1990; Kintsch, Schmalhofer, Welsch, & Zimny, 1990).

**Architectures**

Along with using appropriate representations, a model of discourse must be appropriately structured to match how humans process the discourse. Cognitive architectures can be described as “relatively complete proposals about the structure of human cognition” (Anderson, 1993, p. 3). Unlike individual theories of cognition, the development of more generalized architectures permits models that can account for a wide range of psychological phenomena. Generalized architectures have been a major focus of development in cognitive science over the past 20 years (see Anderson, 1983, 1993; Newell, 1991; Rumelhart & McClelland, 1986a). These architectures have similarly influenced models of discourse processing because they make the assumption that processing language involves the use of general mechanisms and cognitive processes. Some of the primary architectures are described next.

**Production Systems.** The concept behind production system architectures is that skills can be encapsulated as a set of production rules. Production rules are represented as condition-action pairs or if-then rules (Anderson, 1983, 1993). For example, a production system for parsing a text might include the production rule:

IF the current word is a determiner (e.g., the),
THEN expect to be processing a noun phrase

In production system architectures, each production rule can be characterized as an individual piece of knowledge or step in a cognitive
skill. Production rules are abstract so that they can be applied in different possible situations. By chaining together multiple production rules, complex cognitive skills can be performed. In modeling these complex skills, a person’s goals and the results of intermediate steps are maintained in working memory. Nevertheless, in any situation, multiple production rules may be applicable, and a production system must account for how decisions are made about what rules will fire. Although production systems have been primarily applied to modeling skill and problem solving, models of comprehension have also incorporated production systems to predict actions performed during processes of interpretation and comprehension (see Meyer & Kieras, 1997a, 1997b; Just & Carpenter, 1996).

**Connectionism.** Connectionist architectures typically represent components of language as a network of interacting processing units. Each unit is connected to all other units with some associated strength or level of activation. Using a distributed representation, information (such as the knowledge of a word) is represented as a pattern of activation across a set of units rather than as one single node in the network. For example, a representation of the concept *dog* might be represented as a pattern of activation across a number of features relevant to the concept. Thus, *dog* would have activation of feature units such as *fur*, *tail*, *four legs*, and *licks*, but no activation for feature units such as *wings* or *meows*.

Distributed connectionist architectures have been used for modeling many low-level aspects of language processing—for example, visual and auditory word recognition (McClelland & Rumelhart, 1981; Seidenberg & McClelland, 1989), speech production (Dell, 1986), morphological processing for learning the past tense of verbs (Rumelhart & McClelland, 1986b) and word roles in sentence processing (McClelland & Kawamoto, 1986). In addition, there has been some development of models of higher level discourse processing. For example, simple recurrent networks and other connectionist techniques have been used to predict the next word in a sentence (Elman, 1991), model components of discourse comprehension (Golden & Rumelhart, 1993; Sharkey, 1990), account for the activation of inferences (St. John, 1992), and predict dialogue patterns (Graesser, Swamer, & Hu, 1997). See Chater and Christiansen (1999) for a review of connectionist models in natural language processing.

One factor that has limited the use of some connectionist models in higher level discourse processing has been the use of distributed representations in which a concept is represented as a pattern of activation across a set of nodes. It is more difficult to represent objects like
propositions or concepts in a distributed manner. For example, to represent a concept like *dog*, a researcher would have to decide on a technique to generate a pattern of activation across the processing units. In some models, this has been done by indexing concepts by their features, however this requires making assumptions about the dimensions of features involved in representing concepts. Alternatively, it can be done in a principled manner by basing the features on a lexicon such as Wordnet (Fellbaum, 1998) or analyses of statistical properties of the words, such as the frequencies of occurrence of words with other words.

Another solution is to use symbolic representations within connectionist models. In a symbolic connectionist model, concepts and propositions are represented as nodes, with connections among the nodes representing strength of associations. For example, propositions can be connected to each other based on their degree of argument overlap or other discourse connectives such as causal relations or syntactic cues. Additional propositions activated from long-term memory can similarly be connected to the propositional network. For example, in Fig. 13.1, the propositions from the sentence used in Till, Mross, and Kintsch (1988), “The townspeople were amazed that all the buildings had collapsed except the mint” are shown connected by argument overlap (represented as circles) along with associations from long-

![Fig. 13.1. Symbolic connectionist model of the sentence “The townspeople were amazed that all the buildings had collapsed except the mint.”](image-url)
term memory (represented as squares). The dashed lines represent a negative or inhibitory connection between alternative interpretations of the concept mint. These hybrid connectionist models avoid some of the complexities inherent in representing real language input while still using the architecture for modeling aspects of constraint satisfaction, memory retrieval, and learning (Kintsch, 1988, 1998). Additional details about this approach are discussed later in the context of Kintsch’s Construction-Integration model.

Latent Semantic Analysis

Models that use computational linguistic techniques are a more recent technique for deriving representations and modeling discourse processing. Within computational linguistics, techniques have been developed for modeling syntax, semantics, as well as some pragmatic levels of discourse. These models use approaches such as n-gram statistical analyses, Bayesian analyses, Markov models, collocation analyses, symbolic parsers, and the development and analyses of ontologies for representing meaning (see chapter by Moore and Wiemer-Hastings; Boguraev & Pustejovský, 1996; Jurafsky & Martin, 2000, for reviews of computational linguistic approaches). The critical assumption of computational linguistic models is that language has inherent structure and computational techniques can be applied to analyzing language to automatically extract that structure (see Biber, Conrad, & Reppen, 1998; Grodzins & Sidner, 1986).

Despite modeling language, many of the computational linguistic models serve as processing models rather than models of cognition. These models are not compared against human cognitive performance, are not derived from cognitive theories, and are not used to provide explanations for mechanisms found in human information processing. Nevertheless, some computational linguistic models do use cognitive approaches. In this chapter, we focus on a cognitively based computational linguistic architecture called latent semantic analysis (LSA). In LSA (Foltz, 1996a, 1996b; Landauer & Dumais, 1997), a large corpus of running text can be analyzed to derive semantic representations of words and larger units of text (sentences, paragraphs). These texts units are represented as vectors on 100 to 400 indexing dimensions. The vectors can be used directly in models that require distributed representations of concepts or they can be used in symbolic models by computing measures of the degree of association among the vectors for the concepts. LSA uses the assumption found in computational linguistics that, as humans encounter discourse, they must apply techniques to extract information from the structure in language.
By applying statistical techniques to large amounts of discourse, LSA can approximate the way humans process language to acquire semantic knowledge. Additional details on LSA are provided in a later section of this chapter. Other related models have also provided cognitive accounts for the acquisition of semantic or syntactic information through text-based input (see Biber, Conrad, & Reppen, 1998; Burgess & Lund, 2000).

Unlike production system and some hybrid connectionist architectures, LSA does not posit a complete set of cognitive components, such as working memory, mechanisms for retrieval from long-term memory, or rules for the applications of procedural knowledge. Nevertheless, it does posit mechanisms for the acquisition and representation of discourse information in memory and can serve as a general architecture of human knowledge representation. It can be applied to deriving meaning of words, sentences, and paragraphs, and for measuring similarity of meaning among units of texts. Because these components cannot be easily modeled by production systems or most connectionist approaches, this architecture can be integrated with other architectures for developing computational models that need to interpret or generate discourse.

Common Themes in Cognitive Models of Discourse

A number of common themes found across different models have already been touched on here. For example, models use a multileveled representation, models are based on general cognitive processes and mechanisms rather than specialized text-processing mechanisms, and recent models tend to focus on modeling knowledge and comprehension rather than surface-level features of text and discourse. The following are two additional themes contained in many recent models.

Hybrid Architectures. The architectures previously discussed may be blended together. Thus, a fixed separation among these architectures does not really exist in models of discourse processing. Indeed, many of the models are hybrids, incorporating parts of multiple architectures. The Capacity Constrained Construction-Integration model (Goldman, Varma, & Coté, 1996) incorporates the production system architecture found in the Just and Carpenter (1992) CAPS/READER model with the connectionist architecture in Kintsch’s (1988, 1998) Construction-Integration model. As a second example, Kintsch (2000) incorporated the Construction-Integration model with the computational linguistic technique of LSA (Foltz, 1996a, 1996b; Landauer & Dumais, 1997) to model metaphor processing. Although the Construc-
tion-Integration model excels at modeling the assembly and activation of propositions in working memory, adding LSA permits the inclusion of modeling the automatic generation of associations to other related concepts in long-term memory.

The fact that models adopt different components of different architectures illustrates that no architecture accounts for all aspects of discourse processing. By combining the best features of multiple architectures, models can account for a wider range of discourse phenomena. Nevertheless, combining architectures results in greater complexity. Many of the current models treat discourse processing as a complex dynamical system (see Graesser & Britton, 1996). In these models, there are large or infinite numbers of possible states and transitions. Because of this, it becomes more difficult to trace the behavior of a model. Consequently, one must often rely on analyses of the output of the model to determine how well it explains aspects of human discourse processing.

**Modeling Discourse as Cohesive Mental Representations.** Cognitive models of discourse processing all share the theme that, to understand discourse, the discourse must be assembled in memory in a coherent representation. Readers apply knowledge and strategies in processing discourse to create their representation. These strategies include making causal connections, drawing inferences to maintain coherence, and using background knowledge. In most models, this process of generating coherence is a function of the ability to assemble and link propositions to each other and to prior knowledge (see Lorch & O’Brien, 1995, for research on models of coherence).

The theme of coherent representations is well illustrated in the Structure Building Framework (Gernsbacher, 1990, 1996). In the Structure Building Framework, there are three processes. The first process is to lay a foundation through setting up the mental structures needed to establish the coherence. Readers spend more time processing the initial words of a sentence (Aaronson & Ferres, 1983) and the initial sentences in a paragraph (Haberlandt, 1984). This increased time can be explained by readers taking the time to create an initial framework for processing the new discourse. The second process is to map coherent information onto the structures that have been developed. Sentences that are highly coherent with the established structure are easier to map (see Anderson, Garrod, & Sanford, 1983; Bower, Black, & Turner, 1979, and the next section on coherence for additional research on the effects of text coherence on comprehension). Finally, if information being processed is not coherent with the current structure, a third process—shifting—must be performed to
build a new substructure. This overall process of laying foundations, building, and shifting results in a branching structure that is an overall cohesive mental representation of the discourse.

Thus, the process of interpreting discourse involves developing coherence at a local level of connecting discourse together and at a global level of connecting it to prior knowledge. Much of the research in modeling is dedicated to investigating the factors involved in these processes (e.g., inferencing, effects of cohesive devices in discourse, causal analyses, and multidimensional aspects of situation models). See the chapter by Zwaan & Singer for more detail on these approaches.

**DISCOURSE MODELS**

In this section, three models are described that illustrate some of the themes described previously. For example, the models use production system connectionist and computational linguistic techniques as well as hybrids of the techniques. All three of the models are also implemented computationally and have software available. This makes it possible for researchers to generate and test the models on their own data or expand the models to new aspects of discourse processing. Footnotes associated with each model indicate means to acquire more information about implementing the models.

**The Construction-Integration Model**

The Construction-Integration model\(^1\) is based on theories of discourse comprehension developed by Kintsch over the last 25 years (Kintsch, 1988, 1998; Kintsch & van Dijk, 1974; van Dijk & Kintsch, 1983). Comprehension of discourse is viewed as a two-stage cyclic process of constructing representations and integrating it with prior knowledge. A text is parsed into propositions, and a few propositions at a time (about the size of a phrase or short sentence) go through the construction and then the integration phase in each cycle.

In the initial construction stage, propositions from one part of a discourse (e.g., a single sentence) are connected to each other in a network based on overlap of arguments between propositions, the embedding of one proposition in another, or other coherence devices (such as causal relations). In addition, concepts from general knowledge associated with the propositions are also connected to their prop-

---

\(^1\)A version of the Construction-Integration program is available at: [http://psych-www.colorado.edu/users/emross/Work.html](http://psych-www.colorado.edu/users/emross/Work.html). The program runs on Macintosh computers.
ositions. Finally, some additional easily inferred propositions may be added into the network. In this network, connection strength among propositions and concepts is based on amount of overlap or degree of association, contradictory propositions or concepts have negative strengths for their connections. This bottom–up, data-driven process based on weak rules of constructing associations results in an elaborated network of propositions that may contain incoherent and inconsistent interpretations of the discourse (see Fig. 13.1 for an example of the propositional structure created during the construction phase).

The integration phase resolves these inconsistencies. In the integration phase, a connectionist algorithm is used to allow activation to spread among the propositions and concepts. This activation spreads until the network settles. The result is that more strongly connected propositions and concepts receive greater activation. Additionally, inconsistencies in interpretation among competing propositions are resolved with the propositions most strongly and consistently associated with prior knowledge receiving the greatest activation. Finally, some of the mostly highly activated propositions are held over (within the constraints of working memory) for the next cycle of construction and integration.

The resulting patterns of activation of concepts and propositions have been used to model a wide range of discourse phenomena. The change in activation of propositions over cycles in the integration phase predicts priming of the activation of concepts based on reading words in the contexts of texts (Kintsch, 1988; Till, Mross, & Kintsch, 1988). The overall activation of propositions predicts recall of information from texts, including recall of surface feature, textbase, and situation model-level information (Kintsch, 1988; Kintsch, Welsch, Schmalhofer, & Zimny, 1990). The model has further been applied to areas of comprehension including predicting the effects of prior knowledge on comprehension (Mannes & St. George, 1996; McNamara, Kintsch, Songer, & Kintsch, 1996), comprehension and recall of narrative texts (Fletcher, van den Broek, & Arthur, 1996), inferential processes (Kintsch, 1988), metaphor processing (Kintsch, 2000), and processes used in solving arithmetic word problems (Kintsch, 1988). The models and variants have also been used in a number of text-based applications in education (see section on applications).

**Capacity-Constrained Comprehension in Working Memory**

Although propositions are assembled in working memory in the Construction-Integration model, the model does not make strong assumptions about the size of available working memory capacity. In the
model, it typically carries over a fixed number propositions in each cycle (usually two). Much research has shown that readers’ abilities can depend on the amount of resources available for processing. One such resource that has been studied is working memory (Daneman & Carpenter, 1980; Goldman, Hogaboam, Bell, & Perfetti, 1980; Perfetti, 1994). The amount of resources available for processing and storage available in working memory during discourse tasks can be measured and is typically referred to as their working memory span (Daneman & Carpenter, 1980).

The 3CAPS model\(^2\) (Concurrent, Capacity-Constrained Activation-Based Production System; Carpenter, Miyake, & Just, 1994; Just & Carpenter, 1992; Just, Carpenter, & Keller, 1996) proposes that there is a limited amount of activation available in working memory for processing and storage. In the model, all representational elements in discourse (e.g., words, propositions, grammatical structures, concepts activated from long-term memory) have an associated activation level. However, if the overall processing system does not have sufficient total activation, then some of the representational elements do not receive sufficient activation, fall below a threshold, and cannot be used in comprehending the discourse.

The 3CAPS model uses a hybrid of a production system architecture and connectionist network. A set of production rules are used for changing the activation of each element through propagating activation from one element to another. For example, a production rule could be, “If the current word is a determiner (e.g., the),” then “expect to be processing a noun phrase.” Individual production rules can fire repetitively to increase the activation of elements, and all production rules can fire in simultaneously. The result is a changing pattern of activation across the elements during the processing of discourse. This changing activation is connectionist in its approach, creating a spreading activation based on procedural rules of what discourse elements to process in working memory.

By limiting the total amount of activation that can flow within the system, the model can explain many of the effects found related to using working memory for processing in discourse. For example, maintaining information in working memory as one moves from one sentence to the next is critical for maintaining a coherent representation. Readers with larger working memory spans are better at integrating information over greater intersentential distances than readers with smaller working memory spans (see Daneman & Carpenter, 1980).

\(^2\)http://coglab.psy.cmu.edu/3caps.htm is a Web page on computational modeling with 3CAPS as well as links to other projects using 3CAPS.
The model can account for these individual differences in reading ability as well through modeling greater or smaller total amounts of activation that can be applied to processing and storage in working memory. The model has been used to account for a range of additional results in text and discourse, such as individual differences in syntactic processing and linguistic ambiguity, and effects of extrinsic memory loads and aging on working memory capacity (see Carpenter, Miyake, & Just, 1994; Miyake & Shah, 1999, for reviews).

In keeping with the theme of the hybridization of models, the 3CAPS model has also been incorporated with the Construction-Integration model. Goldman and Varma (1995) extended the 3CAPS model by embedding the processes performed by the CI model within a capacity-constrained working memory. This extension further permitted the 3CAPS architecture to be used for processing much longer discourse instead of individual sentences or short sentence sequences. Through modeling recall of passages, they found the model could sometimes perform better than the CI model at maintaining information from prior sentences, and that overall it could simulate patterns of recall of information in children and adults.

Latent Semantic Analysis

In discourse models developed in the past, meaning had to be hand coded, typically by creating propositions from text and having researchers make decisions about how those propositions are associated to concepts in memory. This limited the size of texts that could be analyzed. It also required some ad hoc assumptions about semantic associations among concepts and propositions. In the last 5 years, computational linguistic techniques have become available that permit automatically deriving representations of meaning and knowledge.

LSA3 is a computational linguistic technique that serves as a theory of knowledge induction and representation. LSA infers relations among the meanings of words and passages of discourse. It makes these inferences using no humanly constructed dictionaries, knowledge bases, semantic networks, grammars, syntactic parsers, morphologies, or the like. It takes as its input only raw text parsed into words defined as unique character strings and separated into meaningful passages or samples such as sentences or paragraphs.

3http://lsa.colorado.edu provides demonstrations of applications of LSA, an interface for testing one’s own data using LSA using a variety of corpora, reprints of papers related to LSA, and additional information on how to use LSA.
The primary assumption of LSA is that there is some underlying or latent structure in the pattern of word usage across contexts (e.g., paragraphs or sentences within texts), and that statistical techniques can be used to estimate this latent structure. Through an analysis of the associations among words and meaningful contexts, the method produces a high-dimensional representation in which words that are used in similar contexts are represented as semantically similar. Using this representation, words, sentences, or larger units of text may be compared against each other to determine their semantic relatedness. Mathematical details of LSA may be found in Deerwester, Dumais, Furnas, Landauer, and Harshman (1990); Landauer and Dumais (1997); and Landauer, Foltz, and Laham (1998).

To analyze a text or texts, LSA first generates a matrix of occurrences of each word in each context (e.g., sentences or paragraphs) over a large corpus (typically tens of thousands to millions of words) of running text. In this preprocessing stage, each cell of the matrix contains a transformation of the frequency of the occurrences of each word. The transformation typically used is the log of the frequency of the word divided by the entropy of its occurrence frequencies across all contexts. This entropy transformation provides a measure of the amount of information provided by that word within that context. Transforms of this or similar kinds have long been known to provide marked improvement in information retrieval (Harman, 1986) and have been found important in several applications of LSA. They are probably most important for correctly representing a passage as a combination of the words it contains because they emphasize specific meaning-bearing words.

LSA then applies singular-value decomposition (SVD), a form of factor analysis, or more properly the mathematical generalization of which factor analysis is a special case. The SVD scaling decomposes the word-by-context matrix into a set of $k$, typically 100 to 500, orthogonal factors (or dimensions). This generates an approximation of the original matrix. Instead of representing contexts and terms directly as vectors of independent words, LSA represents them as continuous values on each of the $k$ orthogonal indexing dimensions derived from the SVD analysis. Because the number of factors or dimensions is much smaller than the number of unique terms, words are not independent. For example, if two terms are used in similar contexts, they have similar vectors in the reduced-dimensional LSA representation. One advantage of this approach is that matching can be done between two pieces of textual information even if they have no words in common. To illustrate this, suppose LSA were trained on a
large number of documents on the Panama Canal; consider the following two sentences:

1. The U.S.S. Nashville arrived in Colon harbor with 42 marines.
2. With the warship in Colon harbor, the Colombian troops withdrew.

The vector for the word warship would be somewhat similar to that of the word Nashville because both words occur in the same context of other words such as Colon and harbor. Now if one thinks about computing all possible associations of words based on whether they occur together or whether they occur within the same (or 2nd, 3rd, . . . nth order) context, one is just solving a large constraint satisfaction or series of simultaneous equations. However, this description considerably simplifies and underrepresents the power of the analysis; it is not just a matter of finding words that co-occur with other words that co-occur, but the construction of a joint semantic space in which the relations of all words and passages to all words are accounted for simultaneously. The LSA technique captures a much deeper latent structure than simple term-term correlations and clusters.

One can interpret the analysis performed by SVD geometrically. The result of the SVD is a $k$-dimensional vector space containing a vector for each term and each document. In this space, the cosine of the angle between two vectors is usually taken as the measure of their similarity. Thus, by determining the vectors of two pieces of textual information, we can determine the semantic similarity between them. By choosing appropriate comparisons of various kinds, a wide range of questions about semantic content of discourse can be modeled.

LSA’s performance has been assessed across a number of areas of text and discourse processing. In studies examining the semantic representation of words, a large corpus of representative text is first analyzed to derive the semantic representations of the words. Then the cosine among terms is computed to determine their semantic relatedness. These cosines are then compared against the human relatedness judgments. LSA has been used to model subjects’ ratings of word-word relations, to mimic synonym, antonym, singular-plural, and compound-component word relations, aspects of some classical word sorting and priming studies and of human vocabulary test synonym judgments (Landauer & Dumais, 1997). The results of these studies illustrate that the representation derived by LSA is close to that of a human representation of association among words.

LSA has been applied to several areas of modeling higher levels of discourse processing including text coherence (Foltz, 1996a, 1996b;
Foltz, Kintsch, & Landauer, 1998). By computing the semantic similarity of each sentence to the following sentence across texts, an overall measure of coherence can be derived. For example, if LSA were trained on a corpus of information about the functioning of the human heart, the cosine between the following two sentences could be computed:

There are many kinds of heart disease, some of which are present at birth and some of which are acquired later.

A congenital heart disease is a defect that a baby is born with.

Cosines run between –1 and 1, with greater values indicating greater semantic relationship. The cosine between these two sentences was 0.69, indicating that the sentences are highly semantically related. Although the two sentences share the terms *heart* and *disease* along with some function words, one can intuitively see that the sentences are related because they share additional context due to words such as *birth*, *baby*, and *congenital*. Comparing individual terms in the LSA semantic space, *birth* has a cosine of 0.56 with *baby* and a cosine of 0.33 with *congenital*. In addition, the terms *birth* and *baby* receive much greater weights in the LSA analyses than terms like *heart* or *disease* because they do not occur as frequently within the context of the heart articles. Thus, they contribute more to the calculation of semantic relatedness. A measure of term overlap between sentences would not be able to account for the nuances in semantics in the sentences.

Using this approach, one can compute the coherence of texts at the word, sentence, paragraph, or text section level. LSA-based measures of coherence of texts have been found to predict the comprehensibility of texts better than term overlap measures (Foltz, Kintsch, & Landauer, 1998). Therefore, it can serve as a model of the activation of semantic concepts for maintaining coherence while reading texts. In addition, LSA-based coherence measures can identify places in the text where coherence is low, indicating where a reader is likely needs to make bridging inferences.

LSA has further been used as a predictor of subjective ratings of text properties. For example, it can compare the semantic content contained in a text against the semantic content in other texts. This approach can be used as the basis for scoring essays (Foltz, 1996a, 1996b; Landauer, Laham, & Foltz, 2001). For example, the semantic content of an essay can be compared against the semantic content of other essays for which the grades are known. Generally, essays that have similar content tend to receive similar scores. Thus, measures can be derived for automatically scoring and commenting essays. Studies of these measures show that the derived scores agree with hu-
man scores as often as humans agree with each other. The technique can be further used for determining the appropriate level of text a student should read. By analyzing the semantic content in a student’s essay, it can determine the appropriate level of information that a student would need to learn from a text (Wolfe et al., 1998).

Because LSA serves as a method to automatically generate associations among concepts and propositions, it can work with other architectures in which this was formerly done by hand coding. Kintsch (1998, 2000, 2001) has used LSA-derived meaning representations to demonstrate their possible role in construction-integration-theoretic accounts of sentence comprehension, metaphor, predication, and context effects in decision making. Thus, LSA is not a replacement for other models within discourse, but it can be incorporated into other architectures as an adjunct for modeling semantic aspects in discourse processing.

MODELING ASPECTS OF DISCOURSE, TEXT, AND COGNITIVE PROCESSES

Cognitive models must account for both aspects of the text or discourse and the cognitive processes involved. Some of these aspects have been described previously, such as the effects of working memory capacity on comprehension. To illustrate how models can be used in research and applications, a short overview of some additional aspects are described next. See the chapter by Zwaan and Singer for additional details on aspects related to text comprehension.

Coherence

The role of coherence in comprehension has already been mentioned. However, it is necessary to describe how coherence is modeled. There are many dimensions of a discourse that contribute to coherence, including coreference, causal relationships, connectives, and syntactic signals (Fletcher, Chrysler, van den Broek, Deaton, & Bloom, 1995). In addition, coherence can be both at local and global levels. To model coherence, propositions from the text must be connected to each other based on judgments of whether the propositions cohere. This linking of propositions may be due to one or more of these dimensions of coherence and may occur both within and across sentences. For example, models by Kintsch (1988; van Dijk & Kintsch, 1983) have emphasized that the effect of coreference and propositions are connected on the basis of propositional argument overlap or embedding of a proposition within another.
Additional models have taken into account other aspects of coherence. In models that examine causal relations within narrative texts (Fletcher, van den Broek, & Arthur, 1996; Langston & Trabasso, 1999; Trabasso, Secco, & van den Broek, 1984), propositions or clauses are linked on the basis of a causal structure of the text. Langston and Trabasso (1999) described a hybrid connectionist model that accounts for the integration of new sentences and how those sentences affect the overall memory representation of the text over time. The model first constructs a semantic representation through connecting clauses based on causal inferences. As each clause is added, activation is spread across the network, and the connection strengths and activation values of all clauses in the network are modified. Over each cycle, the resulting activation of the network provides the reader’s memory representation of the text at that particular point of reading. Results show that the connection strengths among clauses within the network can account for a wide range of phenomena found in discourse comprehension, including coherence of the overall text, ratings of causal relatedness among sentences, time to integrate sentences, and immediate and long-term recall of sentences. The model therefore illustrates the dynamic process of connecting and integrating causal information and the effect of prior contexts on new information.

As described in a previous section, LSA has also been used to model the semantic similarity of each sentence to the next to predict the comprehensibility of passages of text (Foltz, Kintsch, & Landauer, 1998). Thus, there is a variety of approaches to modeling coherence. Overall, these models are able to demonstrate which aspects within the discourse contribute to coherence and how the coherence of the discourse affects overall comprehension. Ideally, coherence should be modeled as a mixture of all of these dimensions. The effects of each dimension on how a reader generates a coherent representation and how these dimensions interact need to be modeled.

**Inferential Processing**

To build a coherent representation, it is not just a matter of passively building a representation using only the discourse. Language processing also involves active processing of the information. Inferencing, the creation of additional propositions that are based on information in a text and furnished from world knowledge, is a primary active process. Bridging inferences are made to maintain local coherence, whereas elaborative inferences are additional propositions based on world knowledge that elaborate the text being read. These inferences can be made at both the textbase and situation model level (see Graesser,
Numerous factors can affect inference processing, both due to information contained in the text and due to cognitive processing abilities. For information in the text, factors include the hierarchical level in the text (Walker & Meyer, 1980), distance between textual items requiring bridging inferences (Hayes-Roth & Thorndyke, 1979), and the role of style of text on the construction of a situation model (Perrig & Kintsch, 1985). Cognitive processing factors include reading span (Daneman & Carpenter, 1980) and having appropriate domain knowledge (Spilich, Vesonder, Chiesi, & Voss, 1979).

The process of inferencing involves extracting meaning from the text and activating appropriate knowledge from memory. Thus, to model inferences, it requires a representation of the meaning of text and information in long-term memory. Like modeling coherence, inferences can be modeled using propositional analyses. By measuring how well propositions are connected in the discourse structure, gaps in the discourse can be identified. These gaps tend to correspond to places in the text that would require a reader to make inferences.

As an example, Graesser and colleagues developed the QUEST model of question answering (Graesser & Franklin, 1990; Graesser, Lang, & Roberts, 1991; Graesser, McMahen, & Johnson, 1994). Based on the idea that inferencing is like answering questions like “why” and “how” within a text, the model can predict the availability of inferences. In the QUEST model, information from the text and world knowledge is represented as nodes within a conceptual graph structure. In this structure, nodes are connected based on conceptual relations, such as reason, consequence, and outcome. Thus, the conceptual structure captures effects of causal relations such as goal hierarchies and event chains. To answer questions (make inferences), the model searches in a principled manner along the connections between the nodes. This permits the model to determine whether there is an appropriate connection for answering the question. For different types of questions (“why” or “how”), the model uses different searches on the conceptual structure. For example, “how” questions search backward along the consequence connections. If there is no appropriate connection between nodes required to answer a question or if the connection between the nodes is quite distant, the inference is unlikely to be generated. Therefore, the QUEST model permits the prediction of whether causal inferences are available during the processing of text.

In a related approach that addresses how concepts are activated during inferencing, van den Broek and colleagues (1996) described a landscape view of reading. In this modeling approach, as each sen-
entence of a narrative text is processed, it activates different concepts associated with the sentence represented as propositions. Across sentences there is a fluctuating pattern of activation of concepts. This pattern reflects the associations that the reader would need to make the appropriate inferences to understand each sentence.

Although many models use similar approaches to modeling the connection of propositions and concepts for modeling inferences, it should be noted that there are differences in theoretic approaches to inferences. Some take a constructionist approach, in which elaborative inferences occur in linking the text with world knowledge through general cognitive processes (see Garrod & Sanford, 1994; Graesser, Singer, & Trabasso, 1994; Long & Golding, 1993). Others take a minimalist approach, in which local coherence is based on inferences drawn only for concepts that can be easily and automatically activated (McKoon & Ratcliff, 1992). These differences can affect the extent to which a model might activate concepts in a person’s world knowledge.

Situation Models

As described previously, comprehension goes beyond just building a representation based on a discourse. Comprehension also requires integrating the discourse representation with world knowledge into a situation model. Research has moved toward greater emphasis on modeling the situation model as opposed to the textbase (Anderson, Garrod, & Sanford, 1983; Fletcher & Chrysler, 1990; Glenberg, Meyer, & Lindem, 1987; Millis, King, & Kim, 2000; Morrow, Bower, & Greenspan, 1989).

However, a model of a situation model is more difficult to develop than that of a textbase. Much of the structure of a textbase can be derived directly from the text, whereas the structure of a situation model has to be derived from eliciting a representation of a person’s knowledge. Situation models are not fixed because the situation model emerges as a function of the task and of the person’s current activated knowledge. In addition, situation models can link information based on a variety of dimensions, such as time, causality, or imagery.

Like many of the modeling approaches described earlier, the primary approach to modeling situation models has been through the use of propositional representations. To develop the structure of the situation model, a researcher must infer what possible propositions would be available or activated from memory and how those propositions would be connected. This linking can be done in a principled manner in a number of ways. One approach is to use knowledge elicitation techniques to determine relationships among concepts. This can be
done through having subjects sort concepts, make relatedness ratings among concepts, or test priming of concepts (Schvaneveldt, 1990). A second approach is to use techniques such as LSA to determine semantic association among concepts. Because these associations are derived from natural text, they should reflect the structure of associations found in a reader’s situation model (see Foltz & Wells, 1999; Kintsch, 1998, for tests and examples of this approach).

In establishing coherence at a situation model level, readers must consider how one part of a text relates to the next on a number of dimensions. Zwaan and others (Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995) proposed an event-indexing model, in which readers construct a multithreaded situation model when reading narrative texts. Through modeling five dimensions of continuity within situation models (spatial settings, time, causation, intentions, and characters), they were able to show that reading times increase based on how many of the dimensions lacked coherence. Additionally these identified dimensions were separable, with each having its individual contribution to predictions of reading time of narrative discourse. The modeling approach opens the field to investigating additional factors involved in constructing situation models as well as investigations on the techniques for representing the information contained in situation models.

Metaphors

Much of the modeling effort in text and discourse focuses on the interpretation of the literal meaning of texts. However, there are discourse devices that require readers to use processes other than making literal interpretations. Metaphors are widely used as a research testbed for studying nonliteral comprehension because they require the integration of the textual information with world knowledge to be correctly interpreted. For example, to understand the metaphor my lawyer is a shark requires knowledge about both lawyers and sharks, as well as using some analogical processes that can relate the concepts.

Although many theories of metaphor comprehension have been developed (see Cacciari & Glucksberg, 1994; Glucksberg, 1998), computational models of metaphor have only emerged recently. Kintsch (1988, 2000, 2001) provides a model of metaphor comprehension through combining LSA and the Construction-Integration model. The model is based on the concept of predication—that the meaning and senses of words can only be distinguished by the context in which they appear. In his model, the meaning of concepts are represented in an LSA-derived, high-dimensional semantic space. The meaning of a sen-
tence or metaphor can be represented as the combination of vectors of the words for that sentence or metaphor. For example, the topic of a metaphor (e.g., lawyer) is predicated by the context of the vehicle (shark). Although the vector representation for lawyer is similar to concepts such as justice and crime, adding the vector representation of shark to the vector representation of lawyer makes the metaphor's overall representation more similar to concepts such as vicious in the LSA-derived semantic space.

By comparing the vector representations of metaphor with concepts related to the metaphors, the model is able to predict experimental results such as the nonreversability of metaphors and the use of literal statements to prime metaphors. The model is consistent with prior research showing that predication of metaphors is performed in the same way as the predication of literal statements. The predication model has been extended to modeling other factors of discourse such as causal inferences and homonym disambiguation (see Kintsch, 2001). Overall, the approach illustrates that novel models can be derived through hybrids of existing models and architectures and extend quantitative modeling to new areas of discourse.

Models of Dialogue Patterns

Dialogue patterns provide a final example of a level of discourse that can be handled by quantitative models. Although extensive research has been conducted in conversational analysis, the majority has relied on hand-coded analyses of transcripts. This has greatly limited the size and complexity of discourse analyzed and often led to more qualitative analyses of discourse. Applying statistical and computational linguistic techniques, however, now make automatic analyses and modeling possible.

Patterns of dialogue vary across different contexts of oral discourse. Analyses of these patterns can be instrumental for testing theories of discourse as well as producing applications to process discourse. In one set of studies, Graesser and colleagues (Graesser, Swamer, Baggett, & Sell, 1996; Graesser, Swamer, & Hu, 1997) modeled dialogue patterns of speech acts from tutorial dialogues, phone conversations, and children’s conversations, using a simple recurrent connectionist network to predict transitions among categorized speech acts. In a recurrent network, connections from the hidden units feed to a set of context units. These context units maintain the activations of the hidden units from previous steps within the model and feed the information back to the hidden units. Thus, the model can maintain context of previous information to which it was exposed. Using this ap-
proach, the model was able to predict categorization of speech acts as accurately as humans.

In another study, Kiekel, Cooke, Foltz, and Shope (2001) used LSA as well as statistics of transition probabilities for the analysis of discourse of military personnel performing team flying/navigation tasks. They found that, through automatic analysis of the team members’ transcripts, they could predict overall team performance based on the patterns of discourse.

Computational analysis of dialogues is still in its infancy. Applying both computational linguistic and connectionist modeling techniques to conversational dialogues has great potential. It permits fast analyses on collections of large transcripts that could not be analyzed previously. The results can provide additional insights into the patterns and use of language in conversation.

APPLICATIONS OF MODELS

Quantitative cognitive models of discourse and text also provide useful applications. Because they closely model human processes in understanding text, they can be used in a number of applications that require computers to process the meaning of text in a manner similar to that of humans. These applications include the development of measures of text readability and comprehensibility, methods for message understanding, educational training systems, and dialogue processing.

Propositional modeling has been applied in a number of ways to measuring the comprehensibility of texts. Miller and Kintsch (1980) and Kintsch and Vipond (1979) found that they could predict the comprehensibility of texts through propositionalizing texts and measuring the amount of argument overlap. These predictions were more accurate at predicting readers’ comprehension scores than standard measures of readability (e.g., Flesch, 1948). Britton and Gulgoz (1991) and McNamara, Kintsch, Songer, and Kintsch (1996) have applied Kintsch’s model to instructional texts by detecting places in the text that required readers to make inferences. By repairing those places in the text, they created better texts that produced improved learning in students. The Construction-Integration model has also been adapted to modeling user interactions with computer systems, expanding its generality beyond just texts. For example, Doane et al. (1992) modeled expert–novice differences in comprehending UNIX commands, and Kitajima and Polson (1995) modeled errors in performing complex human–computer interaction tasks.

However, because propositional models require hand coding of the propositions, applications are often limited in the size of text that can
be analyzed. Automatically computing representations of meaning permits the development of applications that understand the meaning of the discourse. In this vein, LSA has been applied to a variety of areas that require analyzing the semantic content of information being expressed. These areas include: predicting the comprehensibility of texts through coherence analysis (Foltz, Kintsch, & Landauer, 1998), predicting which texts influenced students’ knowledge (Foltz, 1996a, 1996b), predicting the appropriate level of text to provide to a student to optimize learning (Wolfe et al., 1998), automatically scoring essay exams (Landauer, Laham, & Foltz, 2001), and integrating LSA into a conversational tutoring system that understands student text input to provide appropriate automatic feedback to students (Graesser et al., 2000).

CONCLUDING REMARKS

Discourse is a complex cognitive process, so modeling discourse requires taking into account many different aspects of discourse and human information processing. There have been efforts within cognitive science to develop universal models of cognition—for example, ACT (Anderson, 1983, 1993) and SOAR (Newell, 1991). These efforts to develop such models are noble because they treat all cognition as a set of processes that can be modeled by a single architecture; in doing so, they are able to account for a wide range of cognitive phenomena. Discourse processing has taken a similar approach, and models have incorporated similar architectures. Models such as the Construction-Integration model, 3CAPS, and LSA have generalizable architectures that permit them to account for a number of aspects in discourse processing.

Marr (1982) made a distinction between three levels at which any machine carrying out an information-processing task could operate. The computational level specifies the goal of the computation and the logic of the strategy by which it can be carried out. The representation and algorithm level specifies how the computational theory could be implemented, what the representation for the input and output would be, and what algorithms would be used. The hardware level specifies the actual physical structures that could implement the representation and algorithm. Models of discourse processing have moved beyond the level of computational theory to the level of representation and algorithm. The models are now quite specific as to how their computational theories and representations can be implemented. Marr’s hardware implementation level, which would be the actual linking of these models to the biological structures, is still in its infancy, however. Al-
though there has been research on biological functions of lower level processing of language (word recognition, word associations, aphasias), this has not yet translated to levels that they can significantly aid in the development of models of higher level discourse processing (see Caplan, 1994).

This brings us back to the ideal model of discourse and text processing. What architectural components make up our ideal model? At the current state of research, some of the components that would go into such a model include:

- Using computational linguistic techniques for deriving semantic and syntactic representations of words, propositions, and sentences.
- Using connectionist techniques for modeling activation of concepts and propositions and performing constraint satisfaction among propositions.
- Using production systems for representing rules that can be applied to decisions that have to be made for processing the discourse.
- Incorporating constraints from general cognitive processes. For example, storage and retrieval constraints are imposed by current models of working memory and long-term memory.
- Incorporating constraints and information from text and discourse, such as the role of syntax, semantics, rhetorical structure, and social context on processing.

To know whether we have achieved the ideal model, we need to evaluate it. Fletcher (1999) identified criteria for computational models of reading. These criteria apply equally well across all theories of discourse processing, and three of his criteria are paraphrased next with some additional commentary. In addition, two other criteria are included.

1. Do the models make it easier to compare the underlying theories with human performance? Computer-based models have made it much easier to derive quantitative predictions. Recent models are developed on computer platforms that can be used by almost anyone with access to computers or the Web. The 3CAPS, Construction-Integration, connectionist, and LSA models all have software available so they can be tested for specific predictions. This has permitted greater dissemination and made it easier for researchers to test their theories against research data. Models are still not at the point that we can simply enter a series of conditions and parameters for a new ex-
periment and have it instantly produce predicted results. Most models still need a fair bit of adjustment and require researchers to fully understand the theories and mechanisms behind the models before they can use them appropriately. This requires understanding the role of different parameters in the models and being able to set them in a way that is both cognitively plausible and within the bounds of the theories on which the model is based. Nevertheless, some of the models described in this chapter have been used inappropriately, resulting in erroneous research conclusions.

2. Do the models allow us to explore and understand theories in ways that would not be possible? A fair bit of modeling in general is still at a point in which a model is developed, tested against one aspect of a theory, published, and then discarded. This is because some models are still not easily implemented by others and are not robust enough to be applied to new domains. This limits their general usefulness. Nevertheless, the more generalized architectures that can be implemented and tested by multiple research groups permit researchers to explore theories in a variety of novel ways. This makes it easier for others to develop new testable research hypotheses based on the models. Thus, findings in one research group can be expanded and extended by other groups because the models can be adapted to new research data.

3. Are the models useful for solving real-world problems? Moving theoretical models to real-world applications is a critical aspect for discourse processing. One sign of a successful theory is its applicability to the real world. There is a wide range of needs for techniques that can process, understand, and generate text and discourse. For example, information retrieval, natural language understanding, machine translation, and intelligent tutoring systems can benefit from such models. Accurate theories of discourse processing should be able to help in these application areas. As described earlier, models are starting to solve real-world problems such as measuring readability, discourse segmentation, and natural language tutoring.

4. How wide a range of discourse phenomena can the models explain? This question is related to Question 2, but focuses on the issue as to whether a model is capturing aspects of more general cognitive processing architectures. If discourse processing is partly explained by these general cognitive processes, then a successful model should be generalizable to a wide range of discourse phenomena. A model that is generalizable can satisfy this criteria at two levels.

4a. How well can a model be transferred to other sets of data within the same area of discourse processing? A model must be robust enough to account for variations in the text or individual differences in humans when it is applied to highly related tasks.
4b. How well can the model be transferred to other areas within discourse processing? A model that captures aspects of general cognitive processing should be applicable to a wide range of areas because many of the same cognitive functions can be used within the different areas. For example, a model of memory for reading text could also be used for accounting for memory for conversations if they use the same memory functions.

5. How well can we differentiate one model from another? If we have multiple successful architectures, how do we distinguish which architecture is the true model of discourse processing? Can we generate testable hypotheses from a model that can help differentiate it from another model? This may help identify which models are best accounting for different cognitive processes or best model aspects of text. New hybrid models can then incorporate the best of the prior models.

In developing models of text and discourse processing, these criteria need to be continuously considered. Although current models do satisfy some of these criteria, they still show limitations that need to be addressed. For example, implementing and adapting models to novel types of discourse can still be quite difficult. This brittleness may indicate areas in which a model needs to adapt its architecture.

More research needs to be performed with quantitative models of text and discourse. At this point in time, a majority of quantitative models have been applied to narrow areas of discourse, primarily investigating comprehension in short expository or narrative texts being read under highly controlled conditions. For models to truly capture the full range of discourse processing, they have to expand to explain additional factors in texts and human processing. The following are some factors that need to be considered. Human affective responses in narrative texts are seldom modeled, but there is evidence that it affects comprehension and interpretation (see Gerrig, 1993). Individual differences affect many discourse processing skills, such as syntactic processing, working memory span, processing meaning, and discourse interpretation. New models need to incorporate these individual differences (Perfetti, 1994). Models of text comprehension have concentrated on learning from individual texts. However, learning is the integration of information across multiple texts, lectures, and conversations. New models need to account for the integration of information from multiple sources (Perfetti, Britt, & Georgi, 1995). Clark (1997) made a strong point that social features influence understanding. Although almost all discourse occurs within social contexts, most current models do not take into account any social features. Finally, models need to expand to new forms of text and discourse. For exam-
ple, the models need to consider comprehension and processing of hypertexts and computer-based training systems (Foltz, 1996b; Graesser, Swamer, & Hu, 1997).

Our ideal model of text and discourse processing has not yet been implemented. However, new models, architectures, and representational techniques continue to be developed, which can help us converge toward our ideal model. Nevertheless, even before we have this model, there is still a wide range of novel discourse and text phenomena to which our current models and techniques can be applied.

ACKNOWLEDGMENTS

The author is grateful to Adrienne Lee, Sara Gilliam, and Jana Olivas for comments on drafts of this chapter.

REFERENCES


Author Index

A

Abel, M., 257, 278
Abelson, R. P., 4, 18, 23, 461, 484, 491, 522
Abrahams, R. D., 45, 49, 60, 74
Adam, B., 297, 314
Adger, C. T., 193, 197
Afflerbach, P., 6, 22, 203, 232,
Agle, G., 268, 280
Aksu-Koç, A., 409, 428
Albin, M. L., 209, 210, 232
Albrecht, J. E., 102, 114, 118, 119
Albro, E. R., 408, 436
Alexander, P. A., 201, 202, 203, 204,
205, 206, 207, 209, 210, 211, 213, 214, 215, 218, 219, 220,
222, 224, 225, 227, 228, 230,
231, 232, 233, 236, 237, 238, 239, 241
Allbritton, D., 378, 389
Allen, J., 9, 19, 309, 318, 440, 470, 482
Allington, R. L., 187, 193, 212, 233
Almasi, J. F., 209, 223, 224, 227, 229, 233
Almor, A., 99, 100, 114
Alonso, M. A., 94, 115
Alton-Lee, A., 192, 193
Altwerger, B., 212, 236
Alvermann, D. E., 222, 224, 228, 229, 230, 233
Ames, C., 208, 213, 233
Ancunzo, M., 399, 428
Anders, P. L., 216, 233
Anderson, A. H., 94, 98, 114, 129, 130,
136, 157, 158, 159, 247, 248,
254, 264, 273, 279, 281, 397, 401, 428
Anderson, B. R. O. G., 31, 74
Anderson, J. R., 91, 109, 111, 114, 493,
498, 509, 513, 517
Anderson, R. C., 105, 114, 203, 210, 234
Anderson, T. H., 206, 234
Anderson-Inman, L., 215, 216, 219,
220, 234
Andringa, E., 328, 330, 332, 333, 344,
351, 355
Andrusiak, P., 103, 120
Ansary, C., 110, 115
Anselmi, D., 399, 428
Archer, J., 208, 233, 234
Argyle, M., 244, 246, 253, 257, 264,
274, 279
Ariel, M., 147, 157
Armbruster, B. B., 206, 234
Arthur, E. J., 500, 507, 518
Arya, P., 223, 229, 233
Atkinson, J. M., 31, 74,
Atkinson, M., 398, 429
Atkinson, P., 291, 299, 314,
Atwell, N., 182, 193
Au, T. K., 97, 114
Austin, J. L., 29, 74, 127, 157
Axelrad, E., 338, 351
Aylett, M. P., 138, 139, 142, 144, 147,
150, 157
Ayling, K., 267, 272, 284

B

Bach, K., 29, 74
Baddeley, A. D., 85, 114, 142, 162
Baecker, R., 256, 282
Baggett, W. B., 490, 511, 519
Bailey, F. G., 30, 74, 312, 313, 315
Bailey, R., 327, 348, 351
Baillet, S. D., 101, 117
Baird, W., 206, 214, 237, 330, 353
Bakhtin, M. M., 201, 234, 297, 314
Ball, J. A., 138, 151, 152, 161
Ball, D. L., 175, 193, 225, 234
Ball., L., 271, 283
<table>
<thead>
<tr>
<th>Author</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ballenger, C.</td>
<td>190, 191, 193</td>
</tr>
<tr>
<td>Bamberg, M.</td>
<td>395, 405, 406, 409, 411, 419, 427, 429</td>
</tr>
<tr>
<td>Barclay, J. R.</td>
<td>91, 93, 102, 115</td>
</tr>
<tr>
<td>Bard, E. A.</td>
<td>132, 138, 139, 143, 144, 147, 150, 157, 264, 279</td>
</tr>
<tr>
<td>Barnes, D.</td>
<td>167, 177, 193, 196</td>
</tr>
<tr>
<td>Baron, C.</td>
<td>289, 316</td>
</tr>
<tr>
<td>Bard, E. A.</td>
<td>132, 138, 139, 143, 144, 147, 150, 157, 264, 279</td>
</tr>
<tr>
<td>Barth, J.</td>
<td>27, 34, 74</td>
</tr>
<tr>
<td>Bartlett, F. C.</td>
<td>90, 102, 114</td>
</tr>
<tr>
<td>Barton, S. B.</td>
<td>103, 114</td>
</tr>
<tr>
<td>Bates, E.</td>
<td>91, 92, 117, 397, 429</td>
</tr>
<tr>
<td>Bauman, R.</td>
<td>52, 74, 299, 315</td>
</tr>
<tr>
<td>Baumann, J. F.</td>
<td>212, 234</td>
</tr>
<tr>
<td>Bavelas, J. B.</td>
<td>146, 157</td>
</tr>
<tr>
<td>Bawarshi, A.</td>
<td>325, 351</td>
</tr>
<tr>
<td>Beach, R.</td>
<td>305, 315, 328, 329, 351, 354</td>
</tr>
<tr>
<td>Beardsley, M.</td>
<td>323, 355</td>
</tr>
<tr>
<td>Beattie, G.</td>
<td>246, 253, 274, 279</td>
</tr>
<tr>
<td>Beck, I. L.</td>
<td>209, 224, 227, 229, 233</td>
</tr>
<tr>
<td>Becker, A.</td>
<td>312, 315</td>
</tr>
<tr>
<td>Becker, J. B.</td>
<td>424, 430</td>
</tr>
<tr>
<td>Beckwith, R.</td>
<td>9, 22</td>
</tr>
<tr>
<td>Beeman, W. O.</td>
<td>32, 63, 74</td>
</tr>
<tr>
<td>Bell, A.</td>
<td>126, 157</td>
</tr>
<tr>
<td>Bell, L. C.</td>
<td>501, 519</td>
</tr>
<tr>
<td>Belson, W. A.</td>
<td>154, 157</td>
</tr>
<tr>
<td>Benjamin, W.</td>
<td>299, 315</td>
</tr>
<tr>
<td>Bennett, L.</td>
<td>209, 237, 475, 478</td>
</tr>
<tr>
<td>Benton, S. L.</td>
<td>209, 210, 232</td>
</tr>
<tr>
<td>Bereiter, C.</td>
<td>215, 240</td>
</tr>
<tr>
<td>Bergin, D. A.</td>
<td>207, 234</td>
</tr>
<tr>
<td>Berknerkotter, C.</td>
<td>201, 234</td>
</tr>
<tr>
<td>Berko-Gleason, J.</td>
<td>402, 413, 414, 415, 422, 429, 431</td>
</tr>
<tr>
<td>Berlyne, D. E.</td>
<td>231, 234</td>
</tr>
<tr>
<td>Berman, R. A.</td>
<td>405, 409, 410, 429</td>
</tr>
<tr>
<td>Bern, H.</td>
<td>101, 115</td>
</tr>
<tr>
<td>Berninger, G.</td>
<td>398, 431</td>
</tr>
<tr>
<td>Bernstein, B.</td>
<td>29, 30, 31, 32, 71, 74, 75</td>
</tr>
<tr>
<td>Beseler, S.</td>
<td>170, 194</td>
</tr>
<tr>
<td>Besnier, N.</td>
<td>61, 75</td>
</tr>
<tr>
<td>Beasgten, Y.</td>
<td>94, 114</td>
</tr>
<tr>
<td>Bever, T. G.</td>
<td>84, 120</td>
</tr>
<tr>
<td>Biber, D.</td>
<td>7, 10, 20, 43, 52, 53, 56, 66, 75, 496, 497, 517</td>
</tr>
<tr>
<td>Biggs, M.</td>
<td>413, 435</td>
</tr>
<tr>
<td>Bigler, K.</td>
<td>110, 118</td>
</tr>
<tr>
<td>Bistoiz, G. L.</td>
<td>14, 21, 54, 77</td>
</tr>
<tr>
<td>Black, J. B.</td>
<td>3, 20, 101, 115, 408, 429, 498, 517</td>
</tr>
<tr>
<td>Black, S. D.</td>
<td>186, 193</td>
</tr>
<tr>
<td>Blakar, R.</td>
<td>129, 157</td>
</tr>
<tr>
<td>Blom, J. P.</td>
<td>63, 75</td>
</tr>
<tr>
<td>Blommaert, J.</td>
<td>287, 315</td>
</tr>
<tr>
<td>Bly, B.</td>
<td>143, 157, 379, 392</td>
</tr>
<tr>
<td>Bly, S.</td>
<td>248, 256, 257, 258, 272, 279</td>
</tr>
<tr>
<td>Bobrick, B.</td>
<td>69, 70, 75</td>
</tr>
<tr>
<td>Bobrow, D.</td>
<td>245, 284</td>
</tr>
<tr>
<td>Bock, J. K.</td>
<td>130, 140, 147, 157</td>
</tr>
<tr>
<td>Boden, D.</td>
<td>31, 75</td>
</tr>
<tr>
<td>Bogdonovich, J.</td>
<td>380, 391</td>
</tr>
<tr>
<td>Boguraev, B.</td>
<td>475, 482, 496, 517</td>
</tr>
<tr>
<td>Bolte, J. D.</td>
<td>215, 217, 219, 234</td>
</tr>
<tr>
<td>Bortfeld, H.</td>
<td>126, 128, 129, 130, 131, 132, 157</td>
</tr>
<tr>
<td>Bortolussi, M.</td>
<td>18, 20, 338, 342, 351, 352</td>
</tr>
<tr>
<td>Botkin, P. T.</td>
<td>142, 159</td>
</tr>
<tr>
<td>Bourdieu, P.</td>
<td>29, 75, 301, 315</td>
</tr>
<tr>
<td>Bourg, T.</td>
<td>343, 351</td>
</tr>
<tr>
<td>Bovair, S.</td>
<td>492, 517</td>
</tr>
<tr>
<td>Bower, G. H.</td>
<td>3, 20, 296, 315, 408, 429, 493, 498, 509, 521</td>
</tr>
<tr>
<td>Boyle, E.</td>
<td>130, 134, 158</td>
</tr>
<tr>
<td>Boynton, M.</td>
<td>403, 431</td>
</tr>
<tr>
<td>Bradner, E.</td>
<td>249, 283</td>
</tr>
<tr>
<td>Brandt, D. M.</td>
<td>470, 483, 492, 521</td>
</tr>
<tr>
<td>Bransford, J. D.</td>
<td>3, 20, 91, 93, 102, 115, 491, 517</td>
</tr>
<tr>
<td>Brauner, J. S.</td>
<td>151, 152, 161</td>
</tr>
<tr>
<td>Brennan, S. E.</td>
<td>125, 126, 128, 129, 130, 131, 132, 133, 134, 136, 146, 147, 153, 157, 158, 161, 163, 164, 245, 247, 252, 275, 279, 285</td>
</tr>
<tr>
<td>Brewer, W. F.</td>
<td>325, 343, 351</td>
</tr>
<tr>
<td>Bricker, P. D.</td>
<td>261, 262, 281</td>
</tr>
<tr>
<td>Briggs, C.</td>
<td>299, 311, 315</td>
</tr>
<tr>
<td>Bright, W.</td>
<td>66, 77</td>
</tr>
<tr>
<td>Brit, M. A.</td>
<td>516, 522</td>
</tr>
<tr>
<td>Britton, B. K.</td>
<td>492, 498, 512, 517</td>
</tr>
<tr>
<td>Britton, J.</td>
<td>177, 193</td>
</tr>
<tr>
<td>Brockmeier, J.</td>
<td>420, 429</td>
</tr>
<tr>
<td>Bromme, R.</td>
<td>142, 158</td>
</tr>
<tr>
<td>Bronfenbrenner, U.</td>
<td>170, 193</td>
</tr>
<tr>
<td>Brooks, C.</td>
<td>308, 315</td>
</tr>
<tr>
<td>Brown, A. L.</td>
<td>212, 224, 239</td>
</tr>
<tr>
<td>Brown, C. M.</td>
<td>89, 120</td>
</tr>
<tr>
<td>Brown, G.</td>
<td>27, 75, 397, 401, 429</td>
</tr>
<tr>
<td>Brown, J. S.</td>
<td>110, 115, 231, 234</td>
</tr>
<tr>
<td>Brown, L.</td>
<td>49, 75</td>
</tr>
<tr>
<td>Brown, M.</td>
<td>378, 379, 383, 391, 392</td>
</tr>
<tr>
<td>Brown, P.M.</td>
<td>31, 75, 101, 117, 137, 143, 158, 159, 359, 389</td>
</tr>
</tbody>
</table>
AUTHOR INDEX

Brown, R., 31, 75, 214, 236
Bruce, B. C., 220, 221, 222, 227, 234
Bruce, S. P., 221, 227, 234
Bruce, V., 247, 258, 279, 280
Bruner, J., 310, 315, 316, 398, 404, 430
Buchalter, D., 358, 391
Buchanan, C. M., 424, 430
Bucholtz, M., 426, 430
Budwig, N., 417, 418, 421, 430, 434
Buehl, M. M., 201, 233
Bullitt-Jones, M., 380, 389
Burger, L. K., 423, 437, 475, 476, 479
Burgess, C., 497, 517
Burke, K., 32, 75
Burke, P. J., 39, 75
Burnett, K., 216, 234
Buttrick, S., 138, 159
Buxton, W. M., 204, 213, 241, 256, 282

C

Cacciari, C., 366, 389, 510, 518
Cain, K., 99, 119
Cairns, H. S., 399, 430
Camaioni, L., 397, 429
Camp, C. J., 108, 115
Campana, M. E., 16, 22
Campana, M. J., 90, 119
Caplan, D., 514, 518
Capps, L., 32, 80, 419, 421, 434
Caramazza, A., 96, 97, 115
Carlson, G. N., 105, 106, 118
Carlson, T. B., 126, 138, 158, 359, 390
Carlsson, M. A., 339, 352
Carpenter, P. A., 90, 115, 494, 497, 501, 502, 508, 518, 520
Carreiras, M., 94, 98, 115, 116
Carriedo, N., 94, 115
Carroll, J. B., 2, 20
Carroll, J. M., 84, 115, 131, 158
Carstensen, L. L., 130, 132, 163
Carston, H., 370, 389, 390
Carter, A., 398, 430
Caverly, D. C., 200, 235
Caazal, C. B., 168, 169, 182, 189, 193, 195, 224, 235
Chalfonte, B., 247, 255, 257, 261, 262, 276, 277, 280, 282
Chambliss, M., 206, 235
Champion, T., 300, 307, 315
Chang, F. R., 97, 115
Charniak, E., 105, 115
Chater, N., 494, 518
Cherny, L., 247, 285
Chesebro, J., 268, 279
Chi, M. T. H., 13, 20
Chiesi, H. L., 3, 20, 43, 491, 508, 522
Chilton, P., 31, 75
Chiu, M., 13, 20
Chodorow, M. S., 97, 118
Chomsky, N., 171, 194, 447, 479
Christian, D., 193, 197
Christian, S., 195, 194
Christiansen, M. H., 494, 518
Christie, B., 244, 251, 265, 267, 269, 271, 272, 274, 276, 284
Chrysler, S. T., 493, 506, 509, 518
Church, M., 384, 392
Chuska, R. K., 176, 194
Cicourel, A. V., 29, 32, 38, 39, 49, 61, 75, 76, 78, 79
Cirilo, R. K., 492, 518
Clark, A., 397, 401, 428
Clark, B., 370, 392
Clark, E. V., 84, 95, 108, 115
Clark, L. F., 88, 116
Clawson, D., 512, 518
Coates, J., 146, 157
Coates, L., 425, 430
Cobb, P., 183, 194, 225, 235
Cohen, E. G., 188, 194
Cohen, K., 254, 279
Cohen, P., 248, 257, 261, 262, 263, 275, 283, 459, 479
Cohn, C., 70, 76
Colby, K. M., 4, 20
Cole, M., 221, 223, 235
Cole, P. 29, 76
Collins, A., 110, 115, 231, 234, 491, 518
Collins, R., 29, 31, 76
Collis, G. M., 397, 436
Colston, H., 359, 384, 390
Conforti, J., 399, 435
Conley, J., 32, 50, 76
Conrad, F. G., 154, 159, 163
Conrad, S., 7, 20, 43, 75, 496, 497, 517
Cook, G., 337, 351
Cook, M., 246, 253, 279
Cooke, N. J., 512, 520
Cook-Gumperz, J., 35, 39, 76, 79
Cooper, R., 258, 280
Coppock, J., 305, 319
Corbett, A. T., 90, 96, 97, 104, 115
Corno, L., 208, 213, 235
Corsaro, W. A., 31, 45, 76
Coser, L. A., 30, 76
Costermans, J., 110, 115
Coté, N., 13, 20, 54, 75, 216, 235, 497, 519
Coulthard, R. M., 3, 23, 172, 196, 224, 241
Coupland, J., 58, 76, 143, 160
Coupland, N., 58, 76, 143, 160, 404, 430
Cox, M., 182, 193
Craddock-Willis, K., 404, 435
Crapanzano, V., 52, 76
Culnan, M. J., 265, 267, 280
Cummins, J., 185, 194
Cupchik, G. C., 338, 345, 351
Curiel, J. M., 93, 121
Currie, M., 287, 308, 309, 315
Curtis, M. E., 97, 101, 117
Cutler, C., 426, 430
Czarniawska, B., 31, 76
Daft, R., 276, 280
Dahrendorf, R., 30, 76
Dail, R.D., 300, 307, 315
Dailey, A., 323, 354
Daly, B., 271, 280
Daly-Jones, O., 252, 255, 257, 258, 277, 280
Damarin, S., 219, 241
Damon, W., 424, 430
Damrad-Frye, R., 405, 411, 429
Daneman, M., 501, 508, 518
Danet, B., 143, 159
Daniels, P. T., 66, 77
Darling-Hammond, L., 187, 194
Davidson, R. G., 403, 430
Davis, J.R., 443, 480
Davis, K., 310, 316
Davis, M. H., 143, 159
Davis, R., 277, 285
de Castell, S., 217, 238
de Leeuw, N., 13, 20
Deaton, J. A., 506, 518
Deci, E. L., 208, 235
Deerwester, S., 503, 518
Deese, J., 413, 430
Delamont, S., 313, 315
Dell, G. S., 137, 144, 148, 158, 159, 494, 518
Delpit, L., 167, 193, 194, 196
Demetrac, M. J., 401, 430
Denis, M., 105, 115, 345, 351
Dennison, R. S., 209, 240
Depperman, A., 425, 430
Derrida, J., 325, 351
Deutsch, W., 133, 159
Dewey, J., 209, 235
Dews, S., 363, 384, 390
Dickinson, D., 182, 193
Dijkstra, K., 343, 351
Dillon, D. R., 222, 224, 228, 230, 233
Dillon, J. T., 222, 235
Dix, A. J., 245, 247, 252, 275, 277, 282
Dixon, P., 18, 20, 338, 342, 351, 352
Doane, S. M., 512, 518
Doctorow, E. L., 302, 316
Doherty, G., 133, 159
Doherty-Sueldon, G., 132, 139, 143, 147, 150, 157, 247, 252, 264, 279, 280
Dolan, R. J., 90, 116
Dole, J. A., 211, 212, 235
Donders, F. C., 85, 115
Dorval, B., 400, 430
Dos Passos, J., 308, 316
Dosher, B. A., 90, 104, 115
Drew, P., 31, 74
Dubois, D., 105, 115
DuBois, J. W., 8, 20
Dubrovsky, V., 267, 270, 280, 284
Duckworth, E., 177, 194
Duffy, G. G., 90, 97, 118, 211, 235
Duguid, P., 231, 234
Duhon, K. E., 209, 233
Duke, N. K., 187, 194
Dumais, S. T., 496, 497, 503, 504, 518, 521
Duncan, S., 246, 253, 261, 274, 280
Duranti, A., 40, 74, 77, 78
Dweck, C. S., 208, 213, 235
Dyson, A., 300, 316

E

Eagleton, T., 326, 352
Eberhard, K. M., 139, 163
Eccles, J. S., 213, 235, 424, 430
Eckerman, C., 400, 430
Eckert, P., 424, 425, 431
Eddy, W. F., 90, 115
Edelsky, C., 212, 236
G

Gagnon, N., 109, 120, 289, 319
Galda, L., 328, 352
Galegher, J., 247, 261, 262, 276, 282
Gallagher, M. C., 212, 239
Gallagher, T., 399, 431
Gallas, K., 170, 194
Gallego, M. A., 221, 236
Gallimore, R., 14, 23
Garlock, C. M., 261, 262, 281
Garner, R., 207, 209, 213, 215, 217, 236
Garnham, A., 98, 99, 116, 119
Garvey, C., 96, 97, 115, 398, 399, 431
Gazaniga, M. S., 421, 431
Gee, J., 14, 20, 166, 192, 194, 223, 236, 413, 431
Geelhoed, E., 245, 248, 252, 258, 272, 275, 277, 285
Geertz, C., 63, 77, 312, 313, 316
Geisler, C., 290, 316
Geller, V., 265, 267, 268, 270, 272, 281
Gelman, R., 134, 163, 204, 227, 236
Genereux, R., 205, 238
Genette, G., 299, 316
Genishi, C., 300, 316
Georgi, M. C., 516, 522
Geren, K. J., 310, 316, 410, 431
Gernsbacher, M. A., 6, 16, 20, 21, 89, 91, 99, 116, 119, 164, 376, 377, 390, 416, 441, 480, 498, 519
Gernsbacher, M. C., 7, 16, 20, 376, 377, 390
Gerrig, R. J., 102, 103, 104, 116, 118, 127, 138, 139, 141, 142, 149, 154, 159, 161, 162, 327, 352, 378, 383, 389, 390, 516, 519
Ghadessy, M., 66, 77
Gibson, J. J., 244, 280
Giddens, A., 305, 316
Gilbert, P., 309, 316
Giles, H., 58, 76, 143, 160
Gillett, G., 404, 407, 410, 432
Gilliam, S., 10, 20
Gilligan, C., 422, 431
Gillingham, M. G., 209, 215, 216, 217, 221, 236
Gilman, A., 31, 75
Gimbert, C., 404, 435
Giora, R., 363, 366, 377, 391
Gipps, C., 178, 191, 196
Girod, M., 221, 241
Givón, T., 416, 431
Glavanov, D., 112, 120
Glenberg, A. M., 114, 116, 493, 509, 519
Glenn, C., 413, 436
Gluckman, M., 61, 77
Glynn, S. M., 492, 517
Godfrey, J., 8, 20
Goetz, E. T., 105, 114, 204, 209, 210, 234, 236, 240, 346, 352
Goffman, E., 126, 160, 176, 194, 292, 316, 361, 391
Golden, R. M., 494, 519
Goldenberg, C., 14, 20, 22
Goldfield, B. A., 174, 196
Gonzales, P., 32, 80
Goodchild, F., 211, 240
Goodman, K. S., 212, 237, 293, 316
Goodwin, C., 32, 40, 61, 74, 77, 78, 123, 127, 160, 300, 316, 423, 431
Goodwin, M. H., 31, 61, 74, 77, 78, 423, 431
Goody, E. N., 31, 78
Goody, J., 66, 67, 68, 78
Gordon, P. C., 88, 99, 116, 441, 447, 480
Gospodinoff, K., 3, 22
Gough, P. B., 3, 21
Grace, P., 310, 316
Hockett, C. F., 66, 79
Hodge, B., 29, 31, 77
Hodge, M. H., 492, 517
Hodge, R., 29, 79
Hodges, R., 403, 431
Hoffstaedter, P., 335, 353
Hogaboam, T., 501, 519
Hogan, M. P., 220, 222, 234
Hollan, J., 278, 281
Holland, N. N., 323, 353
Holliman, E., 8, 20
Holmes, J., 189, 195
Holt, K., 222, 238
Honeck, R., 364, 393
Hood, L., 399, 429
Hoorn, J., 347, 350, 353
Hopper, A., 245, 281
Horgan, D., 399, 432
Jackson, M., 246, 248, 254, 281
 Jacob, E., 189, 195
Jacobs, F., 188, 196
Jacoby, S., 32, 80
Jaffe, J., 262, 281
Jain, R., 408, 437
Jakobson, L. S., 110, 120
Jakobson, R., 32, 79, 299, 317
James, W., 209, 237
Jarvella, R. J., 91, 117
Jarvis, P. E., 142, 159
Jauss, H. R., 323, 353
Jesso, B., 404, 413, 435
Jetton, T. L., 202, 204, 205, 206, 209, 210, 211, 213, 214, 215, 218, 222, 224, 225, 226, 228, 230, 231, 232, 233, 237, 238, 240
John, V. P., 182, 189, 193
Johnson, B. K., 508, 519
Johnson, C. E., 399, 432
Johnson, C. J., 212, 240, 398, 432
Johnson, K., 267, 268, 269, 270, 271, 280, 281
Johnson, M. K., 3, 20, 369, 378, 380, 381, 391, 392
Johnson, N. S., 3, 22, 306, 317
Johnson, T., 146, 157
Johnson, 419, 517
Johnson-Laird, P. N., 91, 117
Jones, G., 277, 285
Jordan, C., 189, 195
Jordan, P., 19, 21
Jorgensen, J., 383, 384, 388, 391, 392
Judy, J. E., 204, 211, 233
Juel, C., 188, 195, 230, 237
Jurafsky, D., 9, 10, 21, 478, 482, 496, 520
Just, M. A., 90, 115, 494, 497, 501, 502, 518, 520
Kahn, D., 61, 79
Kahn, K., 245, 284
Kahneman, D., 258, 281
Kalim, J. D., 338, 351
Kamil, M. L., 216, 238
Kamler, B., 182, 195, 294, 317
Kaplan, B., 395, 417, 428, 432, 437
Kaplan, J., 384, 390
Karolides, N. J., 212, 238
Karsenty, L., 130, 160
Kasckhak, M., 7, 20
Kassler, M. A., 55, 78, 336, 337, 352
Katz, A., 364, 384, 392
Katz, L., 297, 298, 300, 303, 307, 315
Kawamoto, A. H., 494, 521
Keefe, D. E., 107, 117
Keenan, E. O., 399, 432
<table>
<thead>
<tr>
<th>Author</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lehnert, W. G.</td>
<td>4, 21, 108, 117, 408, 432, 441, 460, 475, 482, 483</td>
</tr>
<tr>
<td>Leichner, R.</td>
<td>260, 282</td>
</tr>
<tr>
<td>Lemke, J. L.</td>
<td>172, 195, 216, 224, 225, 238, 288</td>
</tr>
<tr>
<td>Lempers, J. D.</td>
<td>398, 437</td>
</tr>
<tr>
<td>Lengel, R.</td>
<td>276, 280</td>
</tr>
<tr>
<td>Lensmire, T.</td>
<td>309, 317</td>
</tr>
<tr>
<td>Leon, S. D.</td>
<td>126, 129, 157</td>
</tr>
<tr>
<td>Leonard, G.</td>
<td>338, 351</td>
</tr>
<tr>
<td>Lerer, S.</td>
<td>70, 80</td>
</tr>
<tr>
<td>Leung, A.</td>
<td>18, 20, 338, 352</td>
</tr>
<tr>
<td>Levelt, W. J. M.</td>
<td>126, 130, 139, 140, 147, 161</td>
</tr>
<tr>
<td>Levi, P.</td>
<td>302, 317</td>
</tr>
<tr>
<td>Levin, J. A.</td>
<td>186, 193</td>
</tr>
<tr>
<td>Levine, J. M.</td>
<td>181, 196</td>
</tr>
<tr>
<td>Levinson, S. C.</td>
<td>31, 75, 143, 158, 359, 389</td>
</tr>
<tr>
<td>Levy, E.</td>
<td>165, 195</td>
</tr>
<tr>
<td>Levy, F.</td>
<td>396, 404, 415, 416, 417, 418, 427, 433, 434</td>
</tr>
<tr>
<td>Lewis, C.</td>
<td>183, 195</td>
</tr>
<tr>
<td>Lewis, D.</td>
<td>440, 482</td>
</tr>
<tr>
<td>Lewis, S.</td>
<td>134, 161, 248, 282</td>
</tr>
<tr>
<td>Lichtenstein, E. H.</td>
<td>325, 343, 351</td>
</tr>
<tr>
<td>Lichtenstein, S.</td>
<td>142, 161</td>
</tr>
<tr>
<td>Lieberman, P.</td>
<td>132, 161</td>
</tr>
<tr>
<td>Lieblich, A.</td>
<td>287, 317</td>
</tr>
<tr>
<td>Lightman, A.</td>
<td>336, 353</td>
</tr>
<tr>
<td>Lima, P.</td>
<td>381, 391</td>
</tr>
<tr>
<td>Lindem, K.</td>
<td>493, 509, 519</td>
</tr>
<tr>
<td>Lipson, M. Y.</td>
<td>205, 239</td>
</tr>
<tr>
<td>Lloyd, P.</td>
<td>129, 161</td>
</tr>
<tr>
<td>Locke, J.</td>
<td>203, 238</td>
</tr>
<tr>
<td>Lockridge, C. B.</td>
<td>134, 146, 147, 153, 161</td>
</tr>
<tr>
<td>Long, E.</td>
<td>305, 316</td>
</tr>
<tr>
<td>Long, J. L.</td>
<td>217, 239</td>
</tr>
<tr>
<td>Long, J. M.</td>
<td>205, 237</td>
</tr>
<tr>
<td>Lorch, Jr. R. F.</td>
<td>104, 121</td>
</tr>
<tr>
<td>Lorch, R. F.</td>
<td>102, 119, 498, 521</td>
</tr>
<tr>
<td>Lories, G.</td>
<td>110, 115</td>
</tr>
<tr>
<td>Lotan, R. A.</td>
<td>188, 194</td>
</tr>
<tr>
<td>Love, G.</td>
<td>268, 283</td>
</tr>
<tr>
<td>Lucariello, J.</td>
<td>363, 386, 392</td>
</tr>
<tr>
<td>Lucas, M. M.</td>
<td>105, 118</td>
</tr>
<tr>
<td>Lucas, S.</td>
<td>289, 316</td>
</tr>
<tr>
<td>Luff, P.</td>
<td>255, 256, 263, 264, 280</td>
</tr>
<tr>
<td>Luke, A.</td>
<td>217, 238</td>
</tr>
<tr>
<td>Luke, C.</td>
<td>217, 238</td>
</tr>
<tr>
<td>Lund, K.</td>
<td>497, 517</td>
</tr>
<tr>
<td>Lutz, M. F.</td>
<td>104, 118</td>
</tr>
<tr>
<td>MacDonaught, R.</td>
<td>264, 273, 279</td>
</tr>
<tr>
<td>MacWhinney, B.</td>
<td>8, 21</td>
</tr>
<tr>
<td>Madden, C. J.</td>
<td>94, 121</td>
</tr>
<tr>
<td>Madison, C. L.</td>
<td>403, 431</td>
</tr>
<tr>
<td>Magliano, J. P.</td>
<td>6, 13, 23, 88, 93, 118, 120, 121, 331, 343, 349, 350, 351, 354, 355, 408, 437, 493, 510, 523</td>
</tr>
<tr>
<td>Mage, R. A.</td>
<td>188, 194</td>
</tr>
<tr>
<td>Malt, B.</td>
<td>98, 118</td>
</tr>
<tr>
<td>Mandel, G.</td>
<td>91, 116</td>
</tr>
<tr>
<td>Mangiola, L.</td>
<td>181, 195</td>
</tr>
<tr>
<td>Mangold, R.</td>
<td>133, 161</td>
</tr>
<tr>
<td>Mannes, S.</td>
<td>500, 521</td>
</tr>
<tr>
<td>Manning, A. N.</td>
<td>302, 317, 478, 483</td>
</tr>
<tr>
<td>Mannele, S.</td>
<td>403, 433</td>
</tr>
<tr>
<td>Mantei, M.</td>
<td>256, 282</td>
</tr>
<tr>
<td>Marchant, G.</td>
<td>132, 150, 161</td>
</tr>
<tr>
<td>Marchman, V. A.</td>
<td>409, 433</td>
</tr>
<tr>
<td>Marcinkiewicz, M.</td>
<td>7, 22</td>
</tr>
<tr>
<td>Marcus, M.</td>
<td>7, 22</td>
</tr>
<tr>
<td>Marini, A.</td>
<td>205, 238</td>
</tr>
<tr>
<td>Markus, L.</td>
<td>249, 282</td>
</tr>
<tr>
<td>Markus, M. L.</td>
<td>265, 267, 280</td>
</tr>
<tr>
<td>Marques, K.</td>
<td>10, 11, 21</td>
</tr>
<tr>
<td>Marr, D.</td>
<td>513, 521</td>
</tr>
<tr>
<td>Marshall, C. R.</td>
<td>134, 135, 139, 158, 252, 258, 279</td>
</tr>
<tr>
<td>Martin, J. H.</td>
<td>9, 10, 21, 478, 482, 496, 520</td>
</tr>
<tr>
<td>Martin, J. R.</td>
<td>223, 238</td>
</tr>
<tr>
<td>Martindale, C.</td>
<td>323, 354</td>
</tr>
<tr>
<td>Martinez-A.</td>
<td>16, 23</td>
</tr>
<tr>
<td>Martinez-M. A.</td>
<td>398, 433</td>
</tr>
<tr>
<td>Marton, F.</td>
<td>339, 352</td>
</tr>
<tr>
<td>Mason, K.</td>
<td>132, 150, 161</td>
</tr>
<tr>
<td>Matthews, A.</td>
<td>97, 118</td>
</tr>
<tr>
<td>Matute-Bianchi, M. E.</td>
<td>191, 196</td>
</tr>
<tr>
<td>Mauner, G.</td>
<td>106, 118</td>
</tr>
<tr>
<td>Mayer, M.</td>
<td>406, 410, 433</td>
</tr>
<tr>
<td>Mayfield-Stewart, C.</td>
<td>54, 75</td>
</tr>
<tr>
<td>Mazoyer, B. M.</td>
<td>16, 22, 23</td>
</tr>
<tr>
<td>McAllister, J.</td>
<td>132, 150, 161</td>
</tr>
<tr>
<td>McCabe, A.</td>
<td>289, 301, 303, 307, 317, 318, 404, 405, 413, 414, 415, 422, 431, 433, 435</td>
</tr>
<tr>
<td>McCann, A. D.</td>
<td>204, 209, 237</td>
</tr>
<tr>
<td>McCarran, N.</td>
<td>359, 390</td>
</tr>
<tr>
<td>McCarthy, S. J.</td>
<td>212, 236</td>
</tr>
<tr>
<td>McCarthy, J. C.</td>
<td>245, 247, 252, 275, 277, 282, 475, 483</td>
</tr>
<tr>
<td>McClain, K.</td>
<td>225, 235</td>
</tr>
<tr>
<td>Mullen, J., 200, 235</td>
<td></td>
</tr>
<tr>
<td>Muller, U., 277, 285</td>
<td></td>
</tr>
<tr>
<td>Mullin, J., 246, 248, 254, 264, 273, 279, 281, 397, 401, 402, 428</td>
<td></td>
</tr>
<tr>
<td>Münte, T. F., 94, 95, 118</td>
<td></td>
</tr>
<tr>
<td>Murachver, T., 108, 110, 116</td>
<td></td>
</tr>
<tr>
<td>Murdoch, K., 146, 162</td>
<td></td>
</tr>
<tr>
<td>Murnane, R. J., 165, 195</td>
<td></td>
</tr>
<tr>
<td>Murphy, G. L., 92, 118</td>
<td></td>
</tr>
<tr>
<td>Murphy, O., 283, 316</td>
<td></td>
</tr>
<tr>
<td>Murphy, P. K., 201, 203, 204, 206, 207, 209, 211, 217, 225, 233, 239, 241</td>
<td></td>
</tr>
<tr>
<td>Murphy, S., 293, 316</td>
<td></td>
</tr>
<tr>
<td>Myers, J. L., 90, 97, 101, 102, 103, 105, 114, 118, 119, 363, 393</td>
<td></td>
</tr>
<tr>
<td>Myers-Scotton, C., 64, 80</td>
<td></td>
</tr>
<tr>
<td>Nadel, J., 403, 437</td>
<td></td>
</tr>
<tr>
<td>Nadig, A., 150, 152, 162</td>
<td></td>
</tr>
<tr>
<td>Nakatani, C. H., 277, 278, 285, 443, 484</td>
<td></td>
</tr>
<tr>
<td>Nardi, B., 249, 260, 264, 273, 278, 282, 283</td>
<td></td>
</tr>
<tr>
<td>Narens, L., 110, 118</td>
<td></td>
</tr>
<tr>
<td>Narváez, 331, 355</td>
<td></td>
</tr>
<tr>
<td>Nascimento, S., 378, 391</td>
<td></td>
</tr>
<tr>
<td>Nathanson, C. E., 31, 80</td>
<td></td>
</tr>
<tr>
<td>Nayak, N., 378, 379, 391, 392</td>
<td></td>
</tr>
<tr>
<td>Neisser, U., 420, 434</td>
<td></td>
</tr>
<tr>
<td>Nell, V., 329, 354</td>
<td></td>
</tr>
<tr>
<td>Nelson, K., 404, 415, 417, 420, 434, 437</td>
<td></td>
</tr>
<tr>
<td>Nelson, T. H., 215, 239</td>
<td></td>
</tr>
<tr>
<td>Nelson, T. O., 110, 118</td>
<td></td>
</tr>
<tr>
<td>Newell, A., 205, 239, 493, 513, 521</td>
<td></td>
</tr>
<tr>
<td>Newlands, A., 130, 132, 134, 139, 142, 147, 150, 157, 158, 264, 279</td>
<td></td>
</tr>
<tr>
<td>Nicholls, J. G., 208, 239</td>
<td></td>
</tr>
<tr>
<td>Nickerson, R. S., 142, 162</td>
<td></td>
</tr>
<tr>
<td>Nicolle, S., 370, 392</td>
<td></td>
</tr>
<tr>
<td>Nicolopoulos, A., 414, 434</td>
<td></td>
</tr>
<tr>
<td>Nicolaus, K., 411, 434</td>
<td></td>
</tr>
<tr>
<td>Ninio, A., 396, 399, 404, 434</td>
<td></td>
</tr>
<tr>
<td>Nix, D., 215, 239</td>
<td></td>
</tr>
<tr>
<td>Nkanga-Ngila, B., 16, 23</td>
<td></td>
</tr>
<tr>
<td>Noice, H., 18, 22</td>
<td></td>
</tr>
<tr>
<td>Noice, T., 18, 22</td>
<td></td>
</tr>
<tr>
<td>Nolen, S. B., 208, 239</td>
<td></td>
</tr>
<tr>
<td>Nooteboom, S. G., 150, 162</td>
<td></td>
</tr>
<tr>
<td>Nordlie, J., 367, 392</td>
<td></td>
</tr>
<tr>
<td>Norton, J., 308, 318</td>
<td></td>
</tr>
<tr>
<td>Norhwine, A., 262, 283</td>
<td></td>
</tr>
<tr>
<td>Novick, J. M., 138, 144, 152, 160</td>
<td></td>
</tr>
<tr>
<td>Nückles, M., 142, 158</td>
<td></td>
</tr>
<tr>
<td>Nuessle, W., 266, 281</td>
<td></td>
</tr>
<tr>
<td>Nussbaum, J. F., 404, 430</td>
<td></td>
</tr>
<tr>
<td>Nuthall, G., 192, 193</td>
<td></td>
</tr>
<tr>
<td>Nystrand, M., 176, 178, 188, 196, 201, 239</td>
<td></td>
</tr>
</tbody>
</table>

O

O’Brien, D.G., 97, 102, 103, 104, 105, 116, 118, 119, 222, 224, 228, 230, 390, 391, 498, 521
Oakhill, J., 98, 99, 116, 119
Oatley, K., 345, 351
Ochs, E., 32, 80, 291, 300, 312, 318, 398, 404, 418, 419, 421, 429, 431, 434, 435
Ochsman, R., 250, 251, 272, 273, 279
Ogbo, J. U., 191, 196
Ohtsuka, K., 343, 351
Olaniran, B. A., 270, 283
Oldfather, P., 213, 239
Olson, G. M., 333, 337, 338, 350, 354
Olsen-Pulero, L., 399, 401, 435
Olson, D., 130, 133, 159, 162, 333, 337, 338, 350
Olson, G. M., 247, 252, 259, 283
Olson, J. S., 247, 252, 259, 283
Orlando, V., 200, 235
Ormond, J. E., 208, 239
Ortony, A., 3, 22
Oviatt, S., 248, 261, 262, 263, 275, 277, 283

P

Paek, T. S., 138, 161
Palincsar, A. S., 183, 197, 212, 224, 227, 239
Pan, B. A., 396, 400, 435, 436
Papafragou, A., 373, 392
Pappeneheimer, R., 170, 194
Parbery, G., 110, 120
Paris, S. G., 205, 239, 440, 441, 442, 460, 462, 469, 479, 484
Parker, D., 209, 233
Parrish, R., 250, 251, 272, 273, 279
Parsons, G., 397, 436
Passeron, J.-C., 29, 75
Pasupathi, M., 146, 162
Patashnick, M., 208, 239
Patrick, J., 192, 193
Patthey-Chavez, G., 14, 20
Payton, P., 408, 437
Pearson, P. D., 211, 212, 235, 239
Pechnmann, T., 133, 159, 162
Peng, F. C. C., 57, 80
Pennebaker, J. W., 10, 22
Perez, C., 404, 435
Rumelhart, D. E., 3, 22, 85, 110, 111, 119, 461, 484, 493, 494, 519, 521, 522
Rupley, W. H., 211, 238
Russell, A. W., 154, 158, 162
Rutherford, M., 179, 196
Rutter, D., 267, 272, 283, 284
Rutter, M., 248, 265, 267, 271, 276, 284
Ryan, R. M., 208, 235
Ryle, G., 204, 206, 240
Sachs, J. S., 91, 119, 405, 436
Sacks, H., 4, 12, 22, 123, 126, 127, 162, 163
Sadoski, M., 209, 236, 346, 352
Sag, I. A., 98, 120
Samuel, A. G., 132, 150, 162
Sanford, A. J., 92, 94, 97, 98, 103, 114, 116, 120, 509, 519
Sanford, J. A., 493, 498, 509, 517
Santorini, B., 7, 22
Saul, E. U., 13, 20
Saunders, W., 14, 22
Sawyer, D., 185, 194
Scardamalia, M., 215, 240
Scearce, K. A., 99, 116
Schaefer, E. F., 127, 153, 158, 261, 278, 279
Schafer, A. J., 151, 162
Schaffer, H. R., 397, 436
Schallert, D. L., 105, 114, 203, 209, 210, 233, 234, 240
Schank, R. C., 4, 18, 23, 115, 119, 461, 484, 491, 522
Schaw, G., 363, 392
Scheften, A., 32, 81
Schegloff, E. A., 4, 12, 22, 23, 123, 126, 127, 162, 163
Schellings, G. L. M., 214, 240
Schiefele, U., 209, 213, 235, 240
Schieffelin, B. B., 68, 81, 287, 318, 398, 418, 434, 436
Schiffman, D., 126, 163, 411, 436
Shlitz, K., 94, 95, 118
Shug, J., 181, 196
Scholey, S., 400, 436
Schmalhofer, F., 112, 117, 120, 493, 500, 520
Schmidt, S. J., 324, 326, 327, 328, 329, 336, 354
Schober, M. F., 126, 127, 128, 129, 130, 131, 132, 133, 134, 142, 143, 154, 157, 158, 159, 162, 163
Schober-Peterson, D., 398, 436
Schraw, G., 202, 209, 213, 240, 241
Schreier, M., 328, 352
Schreiner, M. E., 513, 523
Schreuder, R., 138, 159
Schulze, S. K., 211, 233
Schunk, D. H., 207, 239
Schwaneveldt, R. W., 87, 118, 510, 522
Schwarz, T., 363, 377, 391
Schwarz, H., 249, 260, 264, 273, 278, 282, 283
Schwarz, N., 156, 163
Scabassi, R., 260, 282
Scott, J. C., 60, 81, 442, 470, 484
Scribner, S., 223, 235
Scrimshaw, P., 184, 197
Searle, J. R., 29, 81, 127, 163, 362, 392
Secco, T., 103, 108, 120, 507, 522
Sedivy, J., 139, 150, 152, 162, 163
Seidenberg, M. S., 494, 522
Sellman, U., 341, 342, 353, 355
Sell, M. A., 490, 511, 519
Selln, A., 246, 247, 248, 253, 254, 256, 263, 264, 273, 280, 282, 283, 284
Sengul, C. J., 91, 115
Sereno, M. I., 16, 23, 89, 117
Sereno, S., 89, 119
Seron, X., 133, 160
Sethna, B., 270, 280
Seymour, H., 300, 307, 315
Sfard, A., 210, 240
Shah, P., 502, 521
Shank, D. M., 105, 119
Shannon, P., 293, 316
Shapiro, A. M., 92, 118
Shapiro, B. P., 54, 103, 115
Shapiro, L. R., 296, 317
Sharkey, N. E., 494, 522
Sharp, D., 293, 316
Shat, M., 134, 163, 401, 418, 436
Shaughnessy, M., 224, 240
Sheeran, Y., 167, 196
Sheldon, A., 422, 423, 436
Shen, Y., 377, 392
Sheridan, D., 295, 318
Shinjo, M., 90, 101, 118, 363, 393
Shook, D., 255, 283
Shope, S., 512, 520
Short, J., 244, 248, 251, 265, 267, 268, 269, 271, 272, 274, 276, 284
Shroyer, S., 89, 116
Shua, A. M., 294, 308, 318
Shultz, J., 31, 49, 50, 77
Shuy, R. W., 32, 49, 50, 81
Sidner, C. L., 249, 277, 285, 440, 441, 443, 449, 451, 470, 471, 481, 484, 496, 520
Siegel, J., 129, 161, 244, 247, 259, 267, 270, 272, 281, 282, 284
Silverman, D., 29, 40, 79, 81
<table>
<thead>
<tr>
<th>Author</th>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silverstein, M.</td>
<td>40, 52, 81, 311, 318, 319</td>
</tr>
<tr>
<td>Simon, H.</td>
<td>88, 116, 205, 239</td>
</tr>
<tr>
<td>Simon, T.</td>
<td>134, 159</td>
</tr>
<tr>
<td>Simons, W. H.</td>
<td>99, 121</td>
</tr>
<tr>
<td>Sinatra, G. M.</td>
<td>210, 240</td>
</tr>
<tr>
<td>Sinclair, J. M.</td>
<td>3, 23, 67, 81, 172, 196, 224, 241</td>
</tr>
<tr>
<td>Singer, I. B.</td>
<td>306, 319</td>
</tr>
<tr>
<td>Singh, S.</td>
<td>61, 81</td>
</tr>
<tr>
<td>Skelly, J. M.</td>
<td>31, 80</td>
</tr>
<tr>
<td>Slobin, D. L.</td>
<td>31, 80, 405, 409, 429, 436</td>
</tr>
<tr>
<td>Smagorinsky, P.</td>
<td>305, 319</td>
</tr>
<tr>
<td>Smallwood, L.</td>
<td>264, 273, 279</td>
</tr>
<tr>
<td>Smith, E. E.</td>
<td>491, 522</td>
</tr>
<tr>
<td>Smith, R. C.</td>
<td>418, 435</td>
</tr>
<tr>
<td>Smith, R. E.</td>
<td>41, 81</td>
</tr>
<tr>
<td>Smitherman, G.</td>
<td>169, 196</td>
</tr>
<tr>
<td>Snow, C. E.</td>
<td>174, 196, 396, 398, 399, 400, 401, 403, 404, 430, 434, 435, 436</td>
</tr>
<tr>
<td>Söderbergh, R.</td>
<td>398, 436</td>
</tr>
<tr>
<td>Soep, E. M.</td>
<td>182, 196</td>
</tr>
<tr>
<td>Solsken, J.</td>
<td>299, 300, 319</td>
</tr>
<tr>
<td>Sontag, S.</td>
<td>322, 355</td>
</tr>
<tr>
<td>Sotillo, C.</td>
<td>132, 139, 143, 147, 150, 157, 264, 279</td>
</tr>
<tr>
<td>Spears, R.</td>
<td>268, 282</td>
</tr>
<tr>
<td>Speer, S. R.</td>
<td>151, 162</td>
</tr>
<tr>
<td>Sperber, D.</td>
<td>136, 163, 370, 383, 387, 388, 392, 393</td>
</tr>
<tr>
<td>Sperl, C. T.</td>
<td>201, 233</td>
</tr>
<tr>
<td>Sperry, L.</td>
<td>103, 120, 423, 433</td>
</tr>
<tr>
<td>Spilich, G. J.</td>
<td>3, 23, 491, 508, 522</td>
</tr>
<tr>
<td>Spiro, R. J.</td>
<td>3, 19, 215, 239, 330, 331, 355</td>
</tr>
<tr>
<td>Spitzer, S.</td>
<td>310, 316</td>
</tr>
<tr>
<td>Spivey-Knowlton, M. J.</td>
<td>138, 163, 378, 383, 388, 391, 392, 393</td>
</tr>
<tr>
<td>Springerst, F. J.</td>
<td>96, 97, 120</td>
</tr>
<tr>
<td>Sproul, L.</td>
<td>244, 248, 255, 267, 268, 270, 284</td>
</tr>
<tr>
<td>Stallworth, L. M.</td>
<td>146, 162</td>
</tr>
<tr>
<td>Stanfield, R. A.</td>
<td>114, 120, 121</td>
</tr>
<tr>
<td>Stanovich, K. E.</td>
<td>187, 196, 203, 241</td>
</tr>
<tr>
<td>Steen, G.</td>
<td>42, 43, 81, 353, 391</td>
</tr>
<tr>
<td>Steffenson, M. S.</td>
<td>399, 437</td>
</tr>
<tr>
<td>Stefik, M.</td>
<td>245, 284</td>
</tr>
<tr>
<td>Stein, N. L.</td>
<td>120, 408, 413, 436, 436, 437</td>
</tr>
<tr>
<td>Steinberg, M. W.</td>
<td>34, 81</td>
</tr>
<tr>
<td>Steiner, G.</td>
<td>39, 81</td>
</tr>
<tr>
<td>Steinfield, C.</td>
<td>271, 284</td>
</tr>
<tr>
<td>Stellmann, P.</td>
<td>131, 163</td>
</tr>
<tr>
<td>Stenton, P.</td>
<td>261, 285</td>
</tr>
<tr>
<td>Stephen, M.</td>
<td>225, 235</td>
</tr>
<tr>
<td>Stephenson, G. M.</td>
<td>248, 267, 269, 272, 282, 284</td>
</tr>
<tr>
<td>Stevens, K. V.</td>
<td>105, 114</td>
</tr>
<tr>
<td>Stone, C.</td>
<td>182, 193, 440, 441, 477, 485</td>
</tr>
<tr>
<td>Stornetta, S.</td>
<td>278, 281</td>
</tr>
<tr>
<td>Straus, S. G.</td>
<td>270, 274, 276, 284</td>
</tr>
<tr>
<td>Streby, W. J.</td>
<td>84, 112, 117</td>
</tr>
<tr>
<td>Street, B.</td>
<td>295, 318</td>
</tr>
<tr>
<td>Strom, L.</td>
<td>378, 381, 391</td>
</tr>
<tr>
<td>Suchman, L.</td>
<td>245, 284</td>
</tr>
<tr>
<td>Suessbrick, A. L.</td>
<td>154, 163</td>
</tr>
<tr>
<td>Suh, S. Y.</td>
<td>88, 104, 120, 408, 437</td>
</tr>
<tr>
<td>Summers, C. R.</td>
<td>403, 437</td>
</tr>
<tr>
<td>Summers, M.</td>
<td>403, 437</td>
</tr>
<tr>
<td>Suttles, G. D.</td>
<td>48, 81</td>
</tr>
<tr>
<td>Swaim, J.</td>
<td>170, 194</td>
</tr>
<tr>
<td>Swales, J. M.</td>
<td>205, 241</td>
</tr>
<tr>
<td>Swamer, S. S.</td>
<td>490, 494, 511, 517, 519, 520</td>
</tr>
<tr>
<td>Swanson, G.</td>
<td>277, 285</td>
</tr>
<tr>
<td>Sweetser, E.</td>
<td>369, 378, 386, 390, 393</td>
</tr>
<tr>
<td>Swezy, L.</td>
<td>134, 161, 248, 282</td>
</tr>
<tr>
<td>Sykes, J.</td>
<td>380, 391</td>
</tr>
<tr>
<td>Symons, S.</td>
<td>212, 240</td>
</tr>
<tr>
<td>Syrota, A.</td>
<td>16, 22</td>
</tr>
</tbody>
</table>

**T**

<table>
<thead>
<tr>
<th>Author</th>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tabors, P.</td>
<td>169, 193</td>
</tr>
<tr>
<td>Tabossi, P.</td>
<td>366, 389</td>
</tr>
<tr>
<td>Takaki, R.</td>
<td>303, 319</td>
</tr>
<tr>
<td>Tang, J.</td>
<td>245, 247, 248, 252, 254, 255, 256, 273, 281, 284</td>
</tr>
<tr>
<td>Tannen, D.</td>
<td>126, 176, 196</td>
</tr>
<tr>
<td>Tasman, N.</td>
<td>292, 317</td>
</tr>
<tr>
<td>Tatar, D.</td>
<td>245, 284</td>
</tr>
<tr>
<td>Taylor, C. E.</td>
<td>193, 194, 418, 435</td>
</tr>
<tr>
<td>Taylor, M.</td>
<td>309, 319</td>
</tr>
<tr>
<td>Teasley, S. D.</td>
<td>181, 196</td>
</tr>
<tr>
<td>Temple, J.</td>
<td>364, 365, 393</td>
</tr>
<tr>
<td>Terry, R.</td>
<td>321, 355</td>
</tr>
<tr>
<td>Terveen, L.</td>
<td>247, 277, 285</td>
</tr>
<tr>
<td>Tessler, M.</td>
<td>420, 437</td>
</tr>
<tr>
<td>Tharp, R.</td>
<td>14, 23</td>
</tr>
<tr>
<td>Throndyke, P. W.</td>
<td>508, 520</td>
</tr>
<tr>
<td>Thorne, J.</td>
<td>67, 81</td>
</tr>
<tr>
<td>Thuiborn, K. R.</td>
<td>90, 115</td>
</tr>
<tr>
<td>Thurlow, R.</td>
<td>508, 522</td>
</tr>
<tr>
<td>Tierney, R.</td>
<td>219, 241</td>
</tr>
<tr>
<td>Till, R. E.</td>
<td>106, 120, 495, 500, 522</td>
</tr>
<tr>
<td>Tobias, S.</td>
<td>213, 241</td>
</tr>
<tr>
<td>Todorov, T.</td>
<td>125, 164, 291, 319</td>
</tr>
</tbody>
</table>
Toma, C., 224, 240
Tomasello, M., 397, 399, 403, 417, 428, 433, 437
Torode, B., 29, 40, 81
Townsend, D. T., 84, 120
Trathen, W., 202, 241
Traugott, E., 52, 81
Tremblay-Leveau, H., 403, 437
Trevarthen, C., 397, 437
Trew, T., 29, 31, 77
Troicki, M., 132, 150, 162
Trollip, S. R., 105, 114
Trueswell, J. C., 138, 144, 152, 160
Tsur, R., 327, 349, 355
Tunstall, P., 178, 191, 196
Turner, A., 492, 522
Turner, J. C., 3, 231, 241
Turner, M., 386, 390
Turner, T. J., 3, 20
Turoll, M., 267, 268, 269, 270, 271, 281
Tuval-Mashiach, R., 287, 317
Twain, M., 179, 196
Twilley, L. C., 18, 20, 338, 352
Tzourio, N., 16, 22, 23
U
Ulchiny, P., 179, 196
Ullman, J., 110, 119
Urban, G., 31, 40, 81
Ury, W., 31, 77
Ugiris, I. È., 418, 421, 430
V
Van Berkum, J. J. A., 89, 120
van den Broek, P., 103, 104, 108, 120, 121, 331, 355, 408, 437, 500, 506, 507, 508, 518, 522
van der Meij, H., 55, 81
van Hout-Wolters, B. H. A. M., 214, 240
Van Meter, P., 204, 209, 237
van Oostendorp, H., 345, 355
Van Peer, W., 327, 346, 347, 348, 355
van Petten, C. K., 89, 117
VanLehn, K., 19, 21
VanSledright, B. A., 202, 203, 218, 225, 232, 241
Varma, K. O., 293, 316
Varma, S., 206, 237, 497, 502, 519
Vermunt, J. D., 214, 240
Vesonder, G. T., 3, 23, 491, 508, 522
Viehoff, R., 325, 327, 328, 353, 355
Vieira, R., 129, 162
Vine, E., 306, 319
Vipond, D., 327, 339, 340, 346, 353, 355, 512, 520
Vogel, D., 276, 282
Voterra, V., 397, 429
Vonk, W., 94, 96, 99, 114, 121
Vorderer, P., 345, 351
Voss, J. F., 3, 23, 491, 508, 522
Vygotsky, L. S., 125, 164, 223, 241, 417, 437
W
Wade, S. E., 202, 204, 209, 213, 220, 221, 241
Waletzky, J., 404, 405, 413, 432
Walker, C. H., 508, 522
Walker, M., 261, 284, 477, 485
Walther, J., 249, 271, 272, 284, 285
Wanner, E., 492, 522
Warren, P., 151, 162
Warren, R. P., 308, 315
Warschauer, M., 186, 196
Wasow, T., 126, 129, 159
Watson-Giego, K. A., 61, 81
Watt, R., 258, 285
Watts, L., 252, 257, 280
Webb, N. M., 183, 197
Weber, K., 110, 119
Weber, S., 4, 20, 293, 318, 319
Weck, G., 250, 251, 272, 273, 279
Wegerif, R., 184, 197
Weinheimer, S., 129, 131, 132, 161, 262, 282
Weintraub, J., 414, 434
Weisband, S., 270, 285
Weisberg, R. W., 84, 121, 129, 130, 154, 160
Weist, R. M., 402, 437
Wellman, B., 256, 282
Wellman, H. M., 398, 437
Wells, A. D., 510, 519
Wells, D., 170, 227, 236
Wells, G., 170, 172, 173, 174, 197, 224, 241
Welsch, D. M., 85, 112, 117, 493, 500, 520
Wendler, L., 305, 315
Wenger, E., 14, 21, 189, 195
Wentzel, K. R., 208, 241
Werner, H., 417, 437
Wertsch, J. V., 182, 197, 418, 421, 430
AUTHOR INDEX

White, C. S., 209, 236
White, D., 403, 433
White, G. M., 61, 81
White, S. D., 206, 211, 214, 238
Whiten, A., 397, 437
Whitenack, J., 183, 194
Whitney, P., 105, 121
Whitten, S. N., 94, 121
Wiemer-Hastings, K., 10, 11, 21, 496, 520
Wiemer-Hastings, P., 10, 11, 21, 496, 520
Wiemer-Hastings, W., 295, 296, 317
Wigfield, A., 209, 213, 231, 235, 237, 239, 241
Wilbur, S., 246, 247, 248, 252, 254, 257, 262, 273, 275, 283
Wiley, A. R., 423, 437
Williams, E., 244, 249, 251, 265, 267, 269, 271, 272, 274, 276, 284, 285
Williams, M., 126, 158
Wills, J., 31, 80
Willson, V. L., 211, 238
Wilson, D., 136, 163, 370, 383, 387, 388, 393
Wilson, S. M., 175, 193
Wimsatt, W. K., 323, 355
Wineburg, S. S., 218, 225, 241
Winner, E., 363, 384, 390
Winograd, T., 4, 23
Winters, L., 266, 280
Wittgenstein, L., 312, 319
Wixson, K. K., 205, 239
Woehler, C., 214, 232
Wolf, D. P., 420, 437
Wolfe, M. B., 54, 75, 506, 513, 523
Wolfram, W., 193, 197
Wollman-Bonilla, J., 176, 197
Wong Fillmore, L., 35, 36, 79
Wood, T., 225, 235
Woods, B. S., 209, 211, 233, 241
Woods, W. A., 4, 23
Woolard, K., 287, 318
Wright, J. W., 142, 159
Wright, L., 54, 81
Wright, P. C., 245, 247, 252, 275, 277, 282
Y

Yackel, E., 225, 235
Yates, J., 96, 97, 115
Yaxley, R. H., 114, 121
Yngve, V., 261, 286
Yore, L. D., 202, 241
Young, M. F., 215, 216, 217, 236, 441, 462, 463, 484, 485
Youniss, E., 221, 236
Yule, G., 27, 75, 129, 164, 397, 401, 429

Z

Zacharski, R., 147, 160
Zajchowski, R., 211, 240
Zajonc, R. B., 349, 355
Zaret, D., 70, 81
Zeitz, C. M., 332, 355
Zhao, Y., 221, 241
Zilber, T., 287, 317
Zimny, S., 112, 117, 493, 500, 520
Zubrow, D., 265, 267, 268, 270, 272, 281
Zwaan, R. A., 6, 21, 32, 78, 91, 92, 93, 94, 114, 120, 121, 325, 326, 330, 331, 332, 334, 335, 336, 343, 345, 350, 351, 353, 355, 493, 510, 523
This page intentionally left blank
Subject Index

A
Act characteristics, 32
Activation measures, 87–89
Adaptation
In spoken discourse, 134–139
Adjustment
Factors of, 141–144
In spoken discourse, 139–141
Adolescent identities in discourse, 424–427
Change with respect to gender, 425
Idiosyncratic language, 425
Affective factors
In literary discourse, 333–334
Affordances
Of mediated communication, 245–247, 275–276
Anaphora, 416
Anchoring, in oral text, 227
Anticipation
In literary discourse, 337–338
Applications
In computational discourse, 473–477
Of discourse models, 512–513
Applications of computational discourse, 473–477
Text summarization, 473–475
Coreference resolution, 475
Text segmentation, 473
Question answering, 476–477
Approaches
Plan based approaches to generating computational discourse, 462–470
Architectures in discourse models, 493–496
Connectionism, 494–496
Production systems, 493
Aspects of discourse modeling, 506–512
Coherence, 506–507
Dialogue patterns, 511–512
Inferential processing, 507–509
Metaphors, 510–511
Situation models, 509–510
Association cues, 359
Attention
Shared in mediated communication, 257–260
B
Bandwidth hypothesis, 249–251
Bauman’s conception of narrative, 299
Berko-Gleason hypothesis, 402
Brain activity measures, 89–90
Brain imaging, 89–90
Bridging inferences, 100–104
C
Capacity Constrained Construction-Integration Model, 497
Cataphora, 417
Chronotopes
In narrative, 297–298
Classroom discourse
Cultural differences, 189–192
Development of students, 192–193
Differential treatment, 187–189
Group discourse, 180–187
With computers, 184–187
With peers, 181–184
Lessons, 170–174
Nontraditional, 174–175
Traditional, 171–174
Sharing time, 168–170
Nontraditional, 169–170
Traditional, 168–169
Variations in discourse features, 175–180
Classroom routines, 179–180

543
Pace and sequence, 179
Speaking rights and listening responsibility, 176–177
Teacher feedback, 178–179
Teacher questions, 177–178
Teaching vs. testing, 180
Classroom discourse and computers, 184–187
Talk at computers, 184–185
Talk in relation to computers, 186–187
Talk through computers, 185–186
Talk with computers, 184
Cognitive cueing, 247, 248, 249, 250
Evidence for, 263–264
Hypotheses, 249–257
Studies in, 252–253
Cognitive processes and power relations
In narrative, 295–297
Coherence, 95–100
Anaphoric resolution, 96–99
Model of, 99–100
And coreference, 95–96
Generating discourse, 459–473
In discourse modeling, 506–507
Sustaining coherent dialogues, 399–400
Collaborative problem solving, 183
Components of literary discourse, 345–349
Foregrounding, 346–347
Imagery, 345–346
Phonetic variation, 348–349
Computational theories of discourse, 444–459
Discourse Representation Theory (DRT), 444–449
Grosz and Sidner Theory, 449–452
Rhetorical Structure Theory, 452–457
Segmented Discourse Representation Theory (SDRT), 457–459
Computers and discourse, in classrooms, 184–187
Conceptual blending theory, 386
Conceptual pacts, 136 see also, Temporary flexible agreements, 136
Connectionism
In discourse models, 494–496
Constraint based account, 138, see also probabilistic account, 138
Construction-Integration Model, 497, 499–500
Constructivist approach to narrative, 410–411
Content hypothesis, 266–270
Negotiation hypothesis, 268–269
Participation and acceptance, 269–270
Content knowledge
Of hypertext, 217–218
Of oral text, 225–226
Content knowledge and oral text, 222–226
Contexts, 33–41
Appropriateness norms, 34
Approaches to context, 39–41
Mundane context, 36–39
Text and situation, 35–36
Contradiction, 359–360
Conversation
And narrative, 415–420
Development of, 396–397
Repair, 399
Conversation, in children, 396, 397, 402–406
Development beyond primary care dyad, 402
Berko-Gleason hypothesis, 402
Research trends, 404
Sibling bridge hypothesis, 403
Conversational feedback, 262, 263
Spoken vs. typed interactions, 263
Conversational repair, 399
Conversations in mediated communication, 255–257
Coreference
Resolution of, 475
Cueing
Cognitive cueing, 247
Cue choice when generating computational discourse, 470–473
Social cueing, 248
Cultural differences in classrooms, 189–193
Cultural characteristics, 191–192
Development of Initiations, 398
Dialogue
Interactional vs. transactional, 401–402
Dialogue patterns
Gender factors in discourse, 511–512
Differences
Gender factors in discourse, 421
Differential treatment in classrooms, 187–189
Discourse
Definition of, 395
Computational theories of, 444–449
Cognitive model of, 488–490
Current applications in computational discourse, 473–477
Generating discourse in computational discourse, 459–473
In classrooms, 165–193
Integration with social/psychological theories, 30–32
Modeling, 490–499
Approaches to, 491–499
Aspects of, 506–512
Models, 499–506
Structure of, 440–444
Attentional structure, 441 see also, Grosz and Sidner Theory
Informational structure, 441
Intentional structure, 440 see also, Grosz and Sidner Theory
Discourse development
Definition of, 395–396
In students, 192–193
Discourse modeling, 490–510
Approaches to, 491–499
Architectures, 493–496 see also, architectures in discourse models
Latent semantic analysis, 496–497
Representation, 491–493 see also, representation models in
Applications of, 512–513
Aspects of, 506–512
Capacity constrained construction integration model, 497, 500–502
Construction-integration model, 497, 499–500
Hybrid models, 497
Latent semantic analysis, 496–497, 502–506
Discourse processes
In literary discourse, 330–334
Affective factors, 333–334
Discourse structures, 330–332
Expert-novice differences, 332–333
Discourse processing
Future, 14–19
Computer technology, 16–18
Integrating Neuroscience, 15–16
Interdisciplinary Research, 18–19
History of, 1–5
Artificial intelligence, 4
Education, 3
Psychology, 3–4
Sociology and communication, 4
Text linguistics, 3
Trends, 5–14
Approaches to analysis, 13–14
Computational discourse, 9–11
Conversational analysis, 12–13
Corpus analysis, 7–9
Culture and discourse, 14
Discourse technology, 11–12
Psychology, 6–7
Discourse Representation Theory (DRT), 444–449
Discourse structures
In literary discourse, 330–332
Discourse, dictionary definitions of, 25–28
Discourse, types of, 28
Dual process account, 138 see also, Constraint based account, 138
Dynamical systems, 498

E

Eclectic/Combinatorial view of narrative, 413–414
Empirical Science Literature (ESL), 328
Endotropic and exotropic theories of narrative discourse, 287–288
Euphemisms, 360
Event Related Potentials (ERP), 89
Expert-novice differences
In literary discourse, 332–333
Eyetracking, 86–87

F

Face-to-face interactions, 261–262, 274
In mediated communication, 261–262, 274
Pace and sequence, 179
Routines, 179–180
Speaking and listening, 176–177
Teacher feedback, 178–179
Teacher questions, 177–178
Teaching vs. testing, 180
Variations in classrooms discord (features), 175–180
Fidelity and vraisemblance
In narrative, 290–292
Foregrounding
In literary discourse, 346–347
Formation of identities, 420–427
Frameworks in literary discourse, 334–340
Anticipation, 337–338
Literary meaning, 339–340
Polyvalence, 336–337
Rereading effects, 338–339
The literary perspective, 334–336
Functional magnetic resonance imaging (fMRI), 89–90
Phases, 406–407
Research paradigms, 407–415
Narrative text, 290–302
Chronotopes, 297–298
Cognitive processes and power relations, 295–297
Fidelity and, 290–292
Social relationships and identities, 292–295
Nonliteral speech acts
Defence of literal meaning, 363–368
Definition of literal meaning, 368–370
Early empirical studies, 362–363
Metaphors in, 373–382
New theories, 386–389
Pragmatics in processing, 370–373
Traditional view of, 361–362
Variety of, 358–361

Online measures of text comprehension, 85–91
Activation measures, 87–88
Brain-activity measures, 89–90
Information content measures, 88–89
Processing load measures, 85–87
Oral text, 222–231
Anchoring, 227
Content knowledge, 225–226
Role of content knowledge, 225
Motivation, 230–231
Organization of content, 459
Strategies, 226–227
Social negotiation, 228–230
Text-structure knowledge, 223–225
Overstatements, 359
Overview of, 396–416
Conversation, 396–397, 402
Development of narrative, 404–420
Identities, 420–427
Initiations, 398
Turn taking, 397–398
Responses, 398–399
Conversational repair, 399
Dialogue, 401–402
Sustaining coherence, 399–400

Interactive approach to narrative, 411–413
Psycholinguistic perspectives of narrative, 409–410
Symbolic/Cultural considerations in narrative, 414–415
Peer teaching, 181–182
Personal issues in literary discourse, 340–345
Feeling-based understanding, 344–345
Implicating the self, 340–342
Types of feeling, 342–344
Phonetic variation
In literary discourse, 348–349
Plan based approaches to generating computational discourse, 462–470
Polyvalence
In literary discourse, 324, 336–337
Positioning strategies, 419
Positron-emission topography (PET), 89–90
Pragmatics, 370–373
Probe recognition, 87–88
Processing load measures, 85–87
Eye-tracking, 86
Moving-window, 86
Self-paced reading, 86
Production systems
In discourse models, 493
Propositional representations in discourse models, 491
Psycholinguistic perspectives of narrative, 409–410

Question answering, 88, 108–113
Memory search in, 109–110
Parallel processes in, 110–111
Stages of, 108
Strategies of, 108–109
Question answering systems, 476–477

Reader’s roles, 322–324
Features of literary texts, 323
Ordinary reader vs. professional reader, 323
Psychoanalytic approach, 323
Reading conventions, 324
Aesthetic, 324
Polyvalence, 324
Rereading effects

Pace and sequence in classrooms, 179
Paradigms
Constructivist approach to narrative, 410–411
Eclectic/Combinatorial view of narrative, 413–414

Tautologies, 359
Teacher feedback in classrooms, 178–179
Teacher questions in classrooms, 177–178
Teaching vs. testing in classrooms, 180
Temporary flexible agreements, 136
Text comprehension
  Activation measures, 87–88
  Lexical decisions, 87
  Naming, 87
  Probe recognition, 87–88
Brain activity measures, 89–90
  Brain imaging, 89–90
  Event Related Potentials (ERP), 89
  Functional magnetic resonance imaging (fMRI), 89–90
  Positron-emission topography (PET), 89–90
Information-content measures, 88–89
  Question answering, 88
Think-aloud protocols, 88
Representations, 91–95, see also, Representations of text comprehension
Research methods, 85–91
Theoretical framework for, 84–85
Text memory, 107–113
  Question answering, 108–113
  Memory search in, 109–110
  Parallel processes in, 110–111
  Stages of, 108
  Strategies of, 108–109
Recall and recognition, 111–113
Text summarization systems, 473–475
Text, constituents of 200–210
  Learning dimensions, 202–203
  Motivations, 207–208
  Salient attributes, 201–202
  Types of knowledge, 203–204
    Content knowledge, 204–205, 217–218, 225–226
    Text-structure knowledge, 205–206, 215–217, 223–225
Text, recall of, 111–113
Text, recognition of, 111–113
Texting, 417
Themes in models of discourse, 497–499
  Cohesive mental representations, 498–499
Hybrid architecture, 497–498
Theoretical approaches
  In conversation development, 404
  In discourse
    Constructivist approach to narrative, 410–411
    Eclectic/Combinatorial view of narrative, 413–414
    Interactive approach to narrative, 411–413
    Paradigms on narrative, 407–415
    Psycholinguistic perspectives of narrative, 409–410
    Story grammars, 408–409
    Symbolic/Cultural considerations in narrative, 414–415
Theoretical distinctions
  In spoken discourse, 139–144
Theoretical issues
  In literary discourse, 327–330
    Empirical Science Literature (ESL), 328
    Socially determined conventions vs. innate proclivities, 329
Theories
  New approaches in nonliteral discourse, 386–389
    Conceptual blending theory, 386
    Relevance theory, 387–388
Think-aloud protocols, 88
Thought
  Metaphors of, 378
  Traditional text, 210–214
    Knowledge, 210–211
    Motivation, 213–214
    Construct of interest, 213
    Strategies, 211–213, 218–220
    Strategic processing and comprehension, 211
TRS coding category, 171
Turn taking, 253–255
Types of narrative, 299

U

Understatements, 359
Unity, 307–309

V

Vagueness, 360
Vraisemblance
  In narrative, 290–292