

Psycholinguistics, formal grammars, and cognitive science

FERNANDA FERREIRA

Abstract

In the 1980s, Charles Clifton referred to a “psycholinguistic renaissance” in cognitive science. During that time, there was almost unanimous agreement that any self-respecting psycholinguist would make sure to keep abreast of major developments in generative grammar, because a competence model was essential, and the linguistic theory was the proper description of that competence. But today, many psycholinguists are disenchanted with generative grammar. One reason is that the Minimalist Program is difficult to adapt to processing models. Another is that generative theories appear to rest on a weak empirical foundation, due to the reliance on informally gathered grammaticality judgments. What can be done to remedy the situation? First, formal linguists might follow Ray Jackendoff’s recent suggestion that they connect their work more closely to research in the rest of cognitive science. Second, syntactic theory should develop a better methodology for collecting data about whether a sentence is good or bad. A set of standards for creating examples, testing them on individuals, analyzing the results, and reporting findings in published work should be established. If these two ideas were considered, linguistic developments might once again be relevant to the psycholinguistic enterprise.

1. Introduction

Perhaps the closest partnership in cognitive science is the one between psycholinguistics and formal linguistics. The two disciplines were born together and have grown up essentially side-by-side as sister disciplines. But like many family relationships, the dynamics between these fields have been complicated. In this article, I will describe the pattern of attraction and withdrawal that has been evident for the last fifty years. At present, psycholinguistics and formal

linguistics are not nearly as close as they once were, and it appears that scholars from both fields have decided they can do their work better independently rather than cooperatively. I will try to explain what has led to this rift and what can be done to bring the fields closer together once again.

2. The psycholinguistic enterprise

Psycholinguists who study adult processing are interested in how people understand and produce language. In the sub-area of comprehension, their aim is to develop theories that explain how listeners understand utterances in real time, even in the face of massive ambiguity and indeterminacy in the input. For production, the goal is to capture how speakers move from a communicative intention to a series of articulatory gestures, which results in utterances that are reasonably fluent and typically comprehensible to others. Psycholinguistic investigations focus on the constraints associated with real time processing. People understand language at the rate of about 300 words per minute, which implies that lexical retrieval, syntactic parsing, and semantic interpretation all occur in a matter of a few hundred milliseconds. Considering the size of the databases that must be consulted during comprehension, the speed and accuracy of human processing is truly astonishing. In the area of language generation, research has established that lemma access, tree building, and phonological /phonetic processing all happen simultaneously and at the rate of about one word per half-second. The fact that linguistic information is accessed so quickly in both comprehension and production has important implications for how that knowledge should be formally represented and how it should be made available to other cognitive systems.

The other key question about language processing that psycholinguists are obliged to address is what enables the linguistic system to work so seamlessly with the rest of the cognitive architecture. For example, visual input is often the trigger for speech (e.g., if while I was teaching a class a bat were to fly into the room, I would be likely to say something about that event), and it can help to shape interpretations (e.g., a lexically ambiguous word such as *ball* likely has only one meaning in the context of a scene that includes a game of soccer). Conversely, language can also guide performance in visual tasks such as visual search. If I am in an office and someone says *Can you pass me the stapler*, my scan patterns will immediately adjust so that I look only at the appropriate sorts of horizontal surfaces. This tight link between vision and language suggests that the two systems work closely together through some type of interfacing representation (for some ideas about the nature of this interface, see Jackendoff 1983). Another example of a system that interacts with language is the one that engages in complex problem solving. It has long been known that

the way a person characterizes a problem to him- or herself affects the range of solutions that are considered (for a recent discussion of this phenomenon, see Li and Gleitman 2002). In addition, language processing calls on mechanisms of attention, memory, and executive control. The attentional system becomes engaged when linguistic processing takes place, and at the same time many aspects of lexical retrieval and even parsing are highly automatic and routinized. Information from long-term memory has to be combined quickly with linguistic input in order for rapid comprehension to occur. All of this processing takes place in a mental “workspace” that is severely limited in capacity. Most people can hold only between about three and seven unstructured pieces of information before they have to relate them in some way. A major job for psycholinguists, then, is to figure out how the linguistic system works in real time and interacts with the rest of the cognitive architecture. Their focus is language, but it is language as it is actually produced and understood by normal adult users of the language.

3. Why have psycholinguists cared about formal linguistics?

Given this description, it should be clear why linguistic theory has been of interest to psycholinguists. (It is less obvious, of course, why generative linguistics is the particular approach that has been dominant.) Linguistic theory has played a major role in psycholinguistic research for many of the reasons that Jackendoff gives in his book *Foundations of Language* (2002). Psycholinguists are in the business of explaining linguistic performance, so it is in their interests to make sure they have a good understanding of the knowledge base that is manipulated during the performance of linguistic tasks. For example, no reasonable scientist could hope to develop a theory of speech perception without knowing about phonological features, segments, and the phonological/phonetic processes that can affect the pronunciation of words. Similarly, a model of sentence comprehension must be based on an understanding of how words are put together compositionally to build interpretations, which means that it must assume some type of model of syntax (although not necessarily the one found in modern generative grammar). Even more fundamentally, as Jackendoff has stressed, psycholinguists are partners in the broader linguistic enterprise. Their focus is on processing, but the representations presumably being generated are linguistic. Therefore, it would be foolish to ignore insights from linguistic theory about the nature of those structures (although psycholinguists might end up disagreeing with those proposals, of course).

There is no one theory of any linguistic domain, let alone a unifying approach to syntactic, semantic, and phonological systems. (Jackendoff’s work,

especially as represented in *Foundations of Language*, might be one important exception.) Linguists have long debated what the best system for describing and explaining structures is. The generative model is only one of the approaches on offer, but it is unquestionably the one that has most influenced psycholinguistics. In some ways, it is ironic that other schools of linguistics that explicitly refer to themselves as “cognitive” (e.g., Taylor 2002) or that have been specifically designed to be representationally compatible with real-time processing algorithms (e.g., Head-Driven Phrase Structure Grammars; Pollard and Sag 1994) have actually had much less impact on those who study comprehension and production. Not just for psycholinguists, but for most cognitive scientists as well, linguistics *is* generative grammar.

4. History of the relationship

The history of the connection between psycholinguistics and linguistic theory reveals both the attraction and the tensions between the two fields. In the early days of generative grammar, there was something of a mutual infatuation. Psychologists were thrilled at the prospect of being able to explain linguistic performance algorithmically and as a type of information processing, and linguists assumed that one of the best tests of the adequacy of their theories was whether they provided workable objects for computer scientists and psychologists trying to implement linguistic structures in machines and in human cognitive systems. Unfortunately, by the early 1970s, the relationship between the two fields was already souring, in part because the Derivational Theory of Complexity (DTC; Foss and Hakes 1978), which assumed that the difficulty of processing a sentence was related directly to the number of transformations that had been applied to it to yield the surface string, did not survive tough experimental scrutiny. This episode in the history of cognitive science has been thoroughly described in other works (e.g., see Fodor et al. 1974), so there is no need to provide a lot of details here. Suffice it to say that when psychologists became disenchanted with the DTC, which was based tightly on what was then current linguistic theory, they concluded that all attempts to link linguistics and psychology were doomed, because linguists were concerned with inventing abstract structures whereas psychologists were interested in something they called “psychological reality” (see, for example, Halle et al. 1978). Of course, this simple-minded view is clearly wrong: Linguists are deeply interested in the issue of psychological reality (as is evident from their attention to the problem of language acquisition), and psychologists need to make assumptions about the types of structures being processed. Nevertheless, this view of the division of labor became the dominant one in cognitive science for about a decade. The two fields went their separate ways, and they remained apart

until the early 1980s, when what Charles Clifton termed a “psycholinguistic renaissance” took place (Clifton 1981).

A number of influences came together to reunite the two fields at this stage in the history of cognitive science. Psycholinguists who remained close to formal linguistics were intrigued by the new approach to grammar known as Government and Binding (GB) (Chomsky 1981), and they pointed out to their colleagues that the representations proposed in GB were enormously promising for capturing certain psycholinguistic phenomena (Frazier et al. 1983). The best example was the new treatment of movement offered in GB. Rather than assuming a set of derivational rules, the theory postulated the existence of elements that might have been moved from some canonical position (the “filler” – J. D. Fodor 1978), along with a representation of the constituent’s original location (the “gap”). Thus, in processing, the problem of understanding certain complex sentences such as the passive could be described not as one of decoding the derived form back to its original kernel form (as was proposed in the DTC, for example; see Fodor et al. 1974), but rather as one of relating the moved constituent and its trace. This view seemed eminently suited to the processing problem of left to right comprehension. The perceiver must identify the filler, hold it in working memory, find the gap, and then relate the filler to that gap. In addition, linguists were addressing phenomena such as lexical argument structure and constraints on syntactic form (e.g., “subjacency”), both of which are clearly relevant to psycholinguistic research. A great deal of productive work was done during the 1980s, and teams involving a linguist and a psychologist became the collaborative ideal (e.g., Lyn Frazier and Charles Clifton; Greg Carlson and Michael Tanenhaus).

But, alas, this rapprochement would not last long. Even as the 1980s were drawing to a close, psycholinguists were once again questioning the usefulness of linguistic theory. Probably the most powerful influence at this stage was the emerging popularity of what became known as “connectionist” or parallel distributed processing (PDP) explanations of cognition (McClelland and Rumelhart 1986). Developers of these models explicitly rejected the standard linguistic methodology and style of argumentation. One of the most famous battlegrounds was over the status of inflectional rules such as the one for past tense formation (Marcus et al. 1992). PDP advocates argued that even something that looked as obviously like a rule as “add -ed” could be described in a much simpler way – as a pattern of activation over a set of simple units based on a reinforcement history of encountering words with particular forms. Linguists and other cognitive scientists raised concerns about the adequacy of the simulations that purported to demonstrate the viability of the PDP approach (Pinker and Prince 1988), but the overall PDP research program itself continued to gain practitioners and has ultimately had a strong influence on the study of language in the cognitive sciences.

This is still true today. There is now a group of psycholinguists who do not believe formal linguistics has much to contribute to their work, because their concern is to discover how patterns in the input lead to complex performance patterns that are merely suggestive of sophisticated knowledge systems, but can be explained far more simply. Some would even now argue that the “poverty of the stimulus” argument has really turned out to be more about the inability of armchair theorists to imagine how much structure and regularity might actually exist in the environmental input than about the lack of such information. Although this coalition is perhaps still the minority among psycholinguists, it is now clear that no one interested in human performance can ignore the possible effects of things such as frequency and exposure on ease of processing (see Bybee & McClelland, this issue). Fortunately, most practicing psycholinguists do not completely dismiss insights from formal linguistics, but the link between generative grammar and processing is not nearly as close as it was in the early 1980s.

Why haven’t the fields reestablished their closer relationship? I would argue that a fundamental reason is the theoretical shift in formal syntax from GB to the Minimalist Program (MP). Unfortunately, the MP is highly unappealing from the point of view of human sentence processing (but see Phillips 2003); this point will be discussed further in the next section. Moreover, the empirical foundation for the MP is almost exclusively intuition data obtained from highly trained informants (i.e., the theorists themselves). Data from other areas such as neurolinguistics, computational linguistics, and psycholinguistics were not taken into account at all, nor were any insights from the rest of the cognitive sciences. Of course, this was true of other theoretical shifts in generative grammar, but it is particularly striking today given the broad range of methods now available for studying language.

5. Formal linguistics and linguistic methodology today

As mentioned above, the MP as a syntactic theory appears to be a step backwards for psycholinguistics (although perhaps not for syntacticians, of course). One of the fundamental problems is that the model derives a tree starting from all the lexical items and working up to the top-most node, which obviously is difficult to reconcile with left-to-right incremental parsing (but see Phillips 2003, for an attempt to deal with this challenge). Also awkward is the notion of “spell-out”. The basic idea behind spell-out is that, under some conditions, all syntactic information within a subpart of a sentence is purged before syntactic analysis of the entire sentence is complete. Because of the way these subparts are defined, this proposal predicts that reanalysis of a garden-path structure such as *While Mary was mending the sock fell off her lap* would be impossible,

because the syntax of the adjunct clause would have been deleted by the time the processor received the error signal *fell* (Weinberg 1999). However, there is evidence that these structures can be successfully reanalyzed under a wide variety of circumstances. Relevant factors include the transitivity of the verb as well as the degree to which the ambiguous noun phrase is plausible as a preposed direct object versus subject of main verb (for example, consider how easy *While John was eating the doorbell rang* is to process). In addition, our work has demonstrated that ease of reanalysis depends on the syntactic characteristics of the ambiguous NP (Ferreira and Henderson 1991; see also Bailey and Ferreira 2003, and Van Dyke and Lewis 2003). Specifically, when that NP contains prenominal modifiers, reanalysis is as easy as it is with just a simple determiner-noun NP; but when the NP contains postnominal modifiers (e.g., a relative clause), reanalysis is far more difficult. Thus, reanalysis processes are sensitive to the position of the head noun in the utterance, a fact that is impossible to explain if spell-out has taken place. Here we see, then, a case where a basic mechanism of minimalism is completely incompatible with known facts about human processing (which were published in mainstream journals more than a decade ago).

Another concern psycholinguists have had about generative grammar almost since its earliest days has to do with the basic methodology for collecting data. When linguistics began, it made a great deal of sense that the primary data would be intuitions about whether sentences were grammatical or ungrammatical. The field needed to get off the ground, and the techniques used in other areas of cognitive science were hardly more sophisticated. Moreover, the contrasts were extremely clear – for example, no experiment is required to show that a sentence such as *John sold the car at Bill near forty dollars* (Akmaijan and Heny 1975) is bad, as its ungrammaticality is obvious. But today the situation is different. Other areas of cognitive science have moved on to far more powerful methodologies, including psychophysics, which relies heavily on converging evidence from other cognitive science methodologies, including computational and mathematical modeling and functional neuro-imaging techniques. (Indeed, the field “psychophysics” no longer really exists, but has been supplanted by a broader discipline known as vision science and visual cognition.) In addition, in formal syntax, the intuitions are no longer uncontroversial. Is a sentence such as *Which car did John ask how Mary fixed* bad, and is it worse than *Who did John ask which car fixed?* (As most readers no doubt know, these are both supposed to be bad because they violate the ECP, but the second is predicted to be worse because it involves a subject extraction.) These judgments are far less straightforward. The result is that there is now a large “disconnect” between linguistic theory and the rest of cognitive science. Other areas of cognitive science increasingly rely on a varied array of sophisticated methodologies that provide far more detailed and accurate data than what was

possible 25 years ago. But linguistics continues to insist that its method for gathering data is not only appropriate, but is superior to others. Occasionally a syntactician will acknowledge that no one type of data is privileged, but the actual behavior of people in the field belies this concession. Take a look at any recent article on formal syntax and see whether anything other than the theorist's judgments constitute the data on which the arguments are based.

Unfortunately, even if we were to accept that judgments of grammaticality are the best data source to use for developing theories of linguistic competence, serious questions can still be raised about the way those data are collected. Judgments are typically gathered as follows. An example sentence that is predicted to be ungrammatical is contrasted with some other sentence that is supposed to be similar in all relevant ways; these two sentences constitute a "minimal pair". The author of the article provides the judgment that the sentence hypothesized to be bad is in fact ungrammatical, as indicated by the star annotating the example. But there are serious problems with this methodology. The example that is tested could have idiosyncratic properties due to its unique lexical content. Occasionally a second or third minimal pair is provided, but no attempt is made to consider the range of relevant extraneous variables that must be accounted for and held constant to make sure there isn't some correlated property that is responsible for the contrast in judgments. Even worse, the "subject" who provides the data is not a naïve informant, but is in fact the theorist himself or herself, and that person has a stake in whether the sentence is judged grammatical or ungrammatical. That is, the person's theory would be falsified if the prediction were wrong, and this is a potential source of bias. Consciously guarding against any possible influence is not sufficient, for research has shown that expectations affect judgments implicitly and unconsciously. Occasionally theorists seem to be aware enough of this problem that they decide to check judgments with a colleague down the hall (sometimes called the "Hey Sally" method). But there is no agreed upon procedure for collecting these additional judgments, and no policies for reporting and reconciling contradictory opinions. I myself have been in the situation of providing a judgment to a linguistics colleague, only to be looked at with an expression of incredulity and to be asked, "Really? Are you sure you have the right reading?". One occasionally even feels badgered into acceding that the data are in fact as the theorist wants them to be. Clearly, this is no way for a modern science to proceed. (For more discussion of problems with the judgment-elicitation methodology, see Edelman and Christensen 2003, and a response to these points from Lasnik and Phillips 2003.) Effects of theorist/experimenter bias have been known in the social psychology literature for decades (the so-called "Hawthorne Effect"). And usually when data are gathered that require some type of subjective decision even from a naïve investigator, a second or even third set of decisions is collected so that reliability statistics can be calculated. The problem is not

that intuitions are bad data; it is that the way they are collected by most formal linguists is problematic. I will discuss this point further below.

6. Current issues and methods in psycholinguistics

To appreciate how wide the separation is now between the study of formal grammar and the field of psycholinguistics, it might help to see a survey of hot topics and approaches in the latter area. What will become clear, I believe, is that few psycholinguistic problems are directly relevant to current theorizing in formal syntax, and vice versa. This survey is based largely on presentations at the most recent major psycholinguistics conferences (viz., the CUNY Sentence Processing Conference and AMLaP, the Conference on Architectures and Mechanisms in Language Processing), and must by necessity be brief and non-technical.

The study of how syntactically ambiguous structures are comprehended continues to be a major area of investigation, because the original logic of the approach is still impeccable: by seeing what the parser does at choice points, we can uncover its decision principles (Frazier et al. 1983). Moreover, because certain “garden-path” effects are so robust, they can be used as a tool to study the influences of nonsyntactic sources of information on parsing. For example, the potential influence of aspectual information is an emerging area, and some researchers have begun to investigate whether progressive forms are easier to process as intransitive, resulting in reduced garden-path effects. The use of argument structure information linked particularly with verbs is also a major research area. In addition, reanalysis itself is a significant topic (see Fodor and Ferreira 1999). The fundamental question is, when the parser encounters a syntactic dead-end, how does it repair the structure it has built? The answer to this question can help to determine whether the system originally builds analyses serially or in parallel, and can also provide information about the time-course of incremental tree creation (both forwards and backwards).

A major weakness in the field of psycholinguistics is that it has focused too heavily on written language, when the spoken medium is much more commonly used and obviously is ontogenetically primary. The reason for this reliance on reading has been convenience: It is easier to present stimuli to participants on a computer monitor than to try to record speech files and play them out, and more importantly, until recently, no sensitive online measures for recording moment-by-moment processing were available for auditory language. Reading can be studied through the use of eye movement monitoring systems that record exactly where people direct the foveal part of their eyes as they read text, which means that sentence comprehension can be studied word by word (or even character by character, if one desired that level of preci-

sion). Recently, this technique has been adapted for use with spoken language (Tanenhaus et al. 1995). Naïve subjects listen to utterances while looking at objects (the so-called “visual world”), some of which are referred to in the sentences. For example, a person might hear *Pick up the candy* while looking at a variety of objects on a table, including candies, candles, cookies, and so on. The eyetracker allows one to measure how quickly the subject looks at the candy in response to the instruction – only at its offset, at its onset, or even before the word begins? This technique has generated an enormous amount of interest, with psycholinguists using it to examine not just word recognition processes but also parsing and language production (for a review, see Henderson and Ferreira 2004).

To appreciate how useful this visual world/eyetracking technique can be, consider this experiment, previously being run by Ming Xiang in my laboratory. (Ming Xiang was a graduate student in Michigan State University’s Linguistics program.) Thanks to linguistic analysis, we know that determiner phrases must be structured so that any numeral phrase precedes any adjectives: *three red squares*, not **red three squares*. Now imagine that people are looking at a number of objects clustered into groups, including three red squares, five red squares, and two red triangles. If they can quickly use information about the linguistic constraint on ordering within DPs, then we might expect that if they hear *Pick up the red . . .* – they will look right away at the triangles, because that cluster does not need to be distinguished by number. Preliminary data indicate that this is exactly what happens, revealing that this linguistic constraint is not only obeyed during language generation, but is also used essentially immediately to guide interpretations during comprehension. Notice too the role that linguistic theorizing has played in this work: Without basic linguistic research, we would probably not know about the constraints on ordering within DPs. But on the other hand, it also does not really matter exactly how the various elements within the DP are structured with respect to each other; all that we are concerned with is that there is a constraint regarding linear order, and that it is highly reliable. In future work Xiang would like to look at adjective ordering, contrasting *the big red squares* versus **the red big squares*. This contrast is not nearly as strong as the one involving numbers and adjectives, and so people’s ability to use the adjective ordering constraint online and immediately could very well be attenuated.

The visual world/eyetracking paradigm is a powerful tool for studying a topic that Jackendoff has emphasized for decades, namely the nature of the interface between language and vision. In addition, it will allow psycholinguists to investigate another important topic, which is the relationship between phonology and syntax during online processing. It certainly seems as if the spoken versions of sentences that are supposed to produce reanalysis failure are quite easy to understand. For example, *While Mary was mending the sock*

fell off her lap would normally be spoken with an intonational phrase boundary at *mending*, and so *the sock* would likely not be taken as its object. But although the intuition is clear, not much is known about how prosodic information is actually used and under what circumstances. And an important problem from the point of view of keeping generative grammar and psycholinguistics together is the problem of spell-out referred to earlier: If syntactic structure is purged for command units, how can prosodic structures that are dependent on syntax and that span an entire sentence be generated?

A new problem that has recently received some attention in psycholinguistics is disfluencies. Researchers are now asking what causes a speaker to say *uh* and *um*, and what is involved in a linguistic repair (e.g., *turn right I mean left at the light*). In our work, we have focused on the comprehension of sentences with disfluencies. We have argued that if the parser hears something like *Put the red ball – the blue ball in the box*, it is faced with a problem of syntactic ambiguity similar to the one that has been the focus of parsing studies. When a person encounters *Because the man drinks beer is never kept in the fridge*, the parser does not know when it first encounters the NP *beer* that it should be treated as the subject of an upcoming clause rather than the object of *drinks*. Similarly, in a repair, the parser does not know when it hears *the red ball* that it will be overwritten with *the blue ball*, and so it initially misanalyzes *the red ball* as object of *put* and then somehow revises that structure. We have argued that this process is similar to garden-path repair, except that a more powerful operation we term *Overlay* must be applied which allows one tree to be put on top of another (when certain specific conditions obtain), creating a layered syntactic representation. The top layer is the one that ultimately should determine the interpretation, but because the reparandum portion is “underneath”, it can continue to exert some influence (Ferreira and Bailey 2004). We believe, then, that disfluencies are not just performance problems; and even if they were, it would certainly fall within the psycholinguist’s job description to study them, as we are supposed to be the ones who study performance as opposed to pure competence.

One important moral of the disfluency project is that generative grammar has not been useful to us in trying to understand the representational or processing issues. Instead, we have found ourselves turning increasingly to the insights from computational linguistics, which I believe is a general trend in the field. Computational work has helped us in two ways. First, statistical investigations of large corpora have yielded a great deal of information about the distribution of disfluencies, and the ideas that have been proposed for getting machines to handle them are excellent starting points for thinking about psychological hypotheses. Second, a computational formalism such as Tree-Adjoining Grammar (TAG) (Joshi and Schabes 1997) turns out to be particularly well suited to capturing human language processing in general and the comprehension of

utterances with disfluencies in particular. Our Overlay operation alluded to earlier is based on TAG,¹ and others in the field have made use of TAG to explain other aspects of comprehension and production.

Finally, it is important at least to mention briefly current research investigating language production. There has been an explosion of interest in this topic over the last 20 years. The earliest work relied exclusively on analyses of speech errors (Garrett 1975), but ingenious techniques have recently been developed for studying experimentally how people generate utterances. Some of the issues that psycholinguists are trying to understand include the following: What determines a speaker's choice of syntactic form? For example, in English a person could express essentially the same idea using an active or passive structure. Earlier work focused on the idea that the two alternatives had different focus structures, which is of course true; but more recent experiments have demonstrated that syntactic options are used to help the speaker manage the activation levels of words and concepts to produce utterances fluently. Thus, if a concept corresponding to a theme is highly activated, it might lead to retrieval of a passive syntactic frame, because that form allows the theme to be plugged into the earliest syntactic position in the tree (Bock 1987). Another issue is incrementality, which has to do with the amount of planning speakers engage in before they begin to speak. Although some investigators argue that the system is highly incremental, others have proposed that the production system attempts to plan about one clause at a time before initiating speech (Garrett 1975; Christianson and Ferreira *in press*; Ferreira and Swets 2002).

7. What would it take for the fields to reunite?

There is no doubt that psycholinguists need to understand the structures that they study when they investigate human language processing. But for reasons I hope I have made clear in this article, it is difficult for psycholinguists to work with generative grammar in its current incarnation. What would make the situation more conducive to collaboration? I have three recommendations.

1. For those who might be interested in the details, the idea is that the parser uses an L-TAG (Lexicalized Tree Adjoining Grammar) to retrieve trees, and those elementary trees are pasted together using the standard TAG operations of substitution and adjoining. But when the parser encounters a disfluency repair, it is faced with a situation in which neither is possible. What it does then is to try to identify root nodes that are identical, and then it places the newer tree on top of the older one, anchored at the root node site. In this way the old structure is overruled, but because it is not actually erased, it can continue to influence processing in the ways we have discovered in our experiments.

The first has to do with the data that forms the basis of generative grammar. It would be good if generative grammarians relied on more than just grammaticality judgments when developing theories of structure, and it would also help if the methods for collecting those judgments were improved. On the first point, in all other areas of cognitive science, the facts on which theories are based are derived using a variety of techniques, ranging from intuitive judgments to reaction times to monitoring of eye movements to computational and mathematical modeling to neural imaging. It sometimes seems as if only generative grammar relies exclusively on one type of data. On the second point, it is important to appreciate that there is nothing wrong with grammaticality judgments as *one* of many types of data. But the way they are collected is problematic. A wide range of examples should be evaluated, along with appropriate controls for each. The examples should be buried in a long list containing filler sentences, because otherwise priming effects and other influences related to exposure could potentially contaminate the results. The items as a group should be presented in more than one random order, to insure that the pattern of responding is not due to sequencing effects. If a context is required for getting the intended reading, it should also be systematically developed and carefully provided along with the critical sentences. And the subject of the experiment should not be the theorist himself or herself; the people providing judgments should be naïve informants who do not know the theories. This suggestion is sometimes dismissed with the argument that expert linguists are needed to provide judgments because they know how to ignore irrelevant aspects of the sentence when making their judgments. But this assumption does not give enough credit to, for example, college undergraduates. They are used to performing all sorts of arbitrary tasks in which they are required to pay attention to some aspects of the stimuli they are presented with and ignore others. Moreover, if linguists adopted a methodology where they used a variety of sentences in more than one condition, computed both means and some type of measure of variability, and used inferential statistics (e.g., analysis of variance), they could take into account those sorts of nuisance effects, because both parametric and nonparametric statistical tests are designed to separate systematic from error variance.

Second, it would be very helpful if the entire field of formal linguistics attempted to re-integrate the studies of syntax, semantics, and phonology. In the 1980s it appeared that syntax and phonology were much closer than they are now, perhaps because of the widespread interest in suprasegmental phonology that emerged during that time. Today, the two fields seem to have separated again. Similarly, semantics is not particularly closely tied to theories of syntax. This is problematic for psycholinguists because even if the system is representationally modular, at some point all sources of information are integrated and form the basis for the comprehender's interpretation of an utterance. For example, we know that listeners build both a syntactic and a prosodic representation

for an utterance, but we do not know how one is generated from the other, nor do we understand how they are integrated or how the two together inform decisions about meaning. Because the field of generative grammar has ignored this question, the vacuum has been filled by psycholinguists, especially those who take a connectionist approach. This is not necessarily bad, but there are reasons to worry that the architecture of such models requires certain simplifying assumptions to be made about representations (e.g., recursion is problematic², as is any type of identity rule), which could lead to misleading ideas about the nature of the system. Therefore, it would be useful if those whose primary interest is language focused more on the relations among phonology, syntax, and semantics.

Finally, returning to the overall theme of this volume, it would be helpful if generative grammar paid more attention to what is happening in the rest of cognitive science. At the moment, the study of formal syntax in particular seems to be off in its own entirely separate world, which makes it very difficult to connect it to the broader enterprise of trying to understand human and other minds. The study of language should be at the core of human cognitive science. Moreover, because we now have ways of looking at human brains when they perform tasks without causing people any pain or harm, it is now possible to develop a biologically plausible model of language. Linguists should join this enterprise; it is a lot of fun, and results have been impressive.

Michigan State University

References

- Akmajian, Adrian, and Frank Heny (1975). *An Introduction to the Principles of Transformational Syntax*. Cambridge, MA: MIT Press.
- Bailey, Karl G. D. and Fernanda Ferreira (2003). Disfluencies affect the parsing of garden-path-sentences. *Journal of Memory and Language* 49: 183–200.
- Bock, J. Kathryn (1987). An effect of the accessibility of word forms on sentence structure. *Journal of Memory and Language* 26: 119–137.
- Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht, Holland: Foris Publications.
- Christianson, Kiel, and Fernanda Ferreira (in press). Planning in sentence production: Evidence from a Free Word Order Language (Odawa). *Cognition*.
- Clifton, Charles Jr. (1981). Psycholinguistic renaissance? *Contemporary Psychology* 26: 919–921.

2. In this context, recursion refers to the embedding of one structure inside another (especially of the same type), and this phenomenon of natural languages is known to be problematic for connectionist theories, which assume that parsing is essentially pattern recognition. Some connectionists have dealt with this weakness by asserting that recursive structures are rare in natural speech (e.g., St. John and McClelland 1992), a claim which is certainly false.

- Edelman, Shimon, and Morten H. Christiansen (2003). How seriously should we take Minimalist syntax?: Comment. *Trends in Cognitive Sciences* 7: 60–61.
- Ferreira, Fernandes and Karl G. D. Bailey (2004). Disfluencies and human language comprehension. *Trends in Cognitive Sciences* 8: 231–237.
- Ferreira, Fernanda, and John M. Henderson (1991). Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language* 31: 725–745.
- Ferreira, Fernandes and Benjamin Swets (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language* 46: 57–84.
- Fodor, Jerry A., Thomas G. Bever, and Merrill F. Garrett (1974). *The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar*. New York: McGraw-Hill.
- Fodor, Janet D. (1978). Parsing strategies and constraints on transformations. *Linguistic Inquiry* 9: 427–473.
- Fodor, Janet D., and Fernanda Ferreira (eds.) (1999). *Renanalysis in Sentence Processing*. Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Foss, Donald J., and David T. Hakes (1978). *Psycholinguistics: An Introduction to the Psychology of Language*. Englewood Cliffs, N.J.: Prentice-Hall.
- Frazier, Lyn, Charles Clifton and Janet Randall (1983). Filling gaps: Decision principles and structure in sentence comprehension. *Cognition* 13: 187–222.
- Garrett, Merrill F. (1975). The analysis of sentence production. In *The Psychology of Learning and Motivation*, Gordon H. Bower (ed.), 133–175. San Diego: Academic Press.
- Halle, Morris, Joan Bresnan, and George A. Miller (1978). *Linguistic Theory and Psychological Reality*. Cambridge, MA: MIT Press.
- Henderson, John M., and Fernanda Ferreira (eds.) (2004). *The Interface of Language, Vision, and Action: Eye Movements and the Visual World*. New York: Psychology Press.
- Jackendoff, Ray (1983). *Semantics and Cognition*. Cambridge, MA: MIT Press.
- (2002). *Foundations of Language*. New York: Oxford University Press.
- Joshi, Aravind K. and Yves Schabes (1997). Tree-adjoining grammars. In *Handbook of Formal Languages*, Grzegorz Rosenberg and Arto Salomaa (eds.), 69–123. Berlin: Springer.
- Li, Peggy, and Lila Gleitman (2002). Turning the tables: Language and spatial reasoning. *Cognition* 83: 265–294.
- Marcus, Gary F., Steven Pinker, Michael Ullman, M. Hollander, T.J. Rosen, and F. Xu (1992). *Overregularization in Language Acquisition*. *Monographs of the Society for Research in Child Development* 57 (4, Serial No. 228). Oxford: Blackwell Publishers.
- McClelland, James and David Rumelhart (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press.
- Phillips, Colin (2003). Linear order and constituency. *Linguistic Inquiry* 34: 37–90.
- Pinker, Steven and Alan Prince (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* 28: 73–193.
- Pollard, Carl J., and Ivan A. Sag (1994). *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- St. John, Mark F. and James McClelland (1992). Parallel constraint satisfaction as a comprehension mechanism. In *Connectionist Approaches to Language Processing*, R. Reilly and N. Sharky (eds.), 97–136. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Tanenhaus, Michael K., Michael Spivey-Knowlton, Kathleen Eberhard, and Julie Sedivy (1995). Integration of visual and linguistic information in spoken language comprehension. *Science* 268: 1632–1634.
- Taylor, John R. (2002). *Cognitive Grammar*. Oxford, U.K.: Oxford University Press.
- Van Dyke, Julie A. and Richard L. Lewis (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language* 49: 285–316.

Weinberg, Amy (1999). Minimalist theory of human sentence processing. In *Working Minimalism*, Samuel D. Epstein and Norbert Hornstein (eds.), 283–315. Cambridge, MA: MIT Press.