

## Accessing Spoken Words: The Importance of Word Onsets

William Marslen-Wilson  
Max-Planck Institute for Psycholinguistics  
Nijmegen, The Netherlands  
and MRC Applied Psychology Unit  
Cambridge, England

Pienie Zwitserlood  
Max-Planck Institute for Psycholinguistics  
Nijmegen, The Netherlands  
and Interfaculty Research Unit for Language and Speech  
University of Nijmegen, Nijmegen, The Netherlands

Approaches to spoken word recognition differ in the importance they assign to word onsets during lexical access. This research contrasted the hypothesis that lexical access is strongly directional with the hypothesis that word onsets are less important than the overall goodness of fit between input and lexical form. A cross-modal priming technique was used to investigate the extent to which a rhyme prime (a prime that differs only in its first segment from the word that is semantically associated with the visual probe) is as effective a prime as the original word itself. Earlier research had shown that partial primes that matched from word onset were very effective cross-modal primes. The present results show that, irrespective of whether the rhyme prime was a real word or not, and irrespective of the amount of overlap between the rhyme prime and the original word, the rhymes are much less effective primes than the full word. In fact, no overall priming effect could be detected at all except under conditions in which the competitor environment was very sparse. This suggests that word onsets do have a special status in the lexical access of spoken words.

A fundamental property of the speech signal is its intrinsic directionality in time. Spoken utterances are spread out along the time line, moving necessarily from beginning to end, in a way that is not true of written language. This directionality of the speech input is strongly reflected in the claims made by the cohort model of spoken word recognition for the manner in which speech inputs are mapped onto the representations of word forms in the mental lexicon (Marslen-Wilson, 1984, 1987; Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Welsh, 1978; Tyler, 1984; Tyler & Wessels, 1983; Warren & Marslen-Wilson, 1987).

The cohort model of word recognition stresses the sequential and continuous nature of the mapping between the speech input and mental representations of word forms. This emphasis is closely tied up with the concept of a cohort and its implications for the properties of the on-line lexical decision space. In particular, according to the cohort model, the decision space is determined by the beginnings of words. The speech input at the beginning of the word maps onto all lexical items that share the same initial sequence. This initial set of candidates is termed the *word-initial cohort*, and the subsequent process of word recognition is determined by the

way in which different sources of constraint—both sensory and contextual—operate on the membership of this set.

This emphasis on the beginnings of words has been widely disputed. The criticisms have been developed along two fronts. The first kind of criticism is based on the belief that the sensory input in fluent speech cannot guarantee the system reliable information about word onsets. A cohort-based recognition process would run into trouble whenever the information at word beginnings was inadequate, and this would happen far too often for such a system to be workable. A favorite example is the system's difficulties with items like "shigarette" or "dwibble," in contrast to the alleged ability of human listeners to recognize such items with no trouble at all (e.g., Grosjean, 1985; Norris, 1981; Salasoo & Pisoni, 1985). Furthermore, as Marcus and Frauenfelder (1985) pointed out, the mental lexicon may in any case be structured to provide maximum distinctiveness not at the beginnings of words, but rather as more of the word is heard.

The second line of criticism derives from connectionist models of the mapping process, in which the directionality of the mapping is less crucial. In these models, what is most important is the overall goodness of fit between the complete stimulus and a given lexical representation, relative to the input's goodness of fit to other potential candidates. The major representative of these models in the auditory domain is the TRACE model developed by Elman and McClelland (1986; McClelland & Elman, 1986). Directionality, in this model, is not an explicit condition on the possibility of lexical access, although there may be implicit effects of the temporal sequence in which information arrives over time, given that the word onsets are processed first. Such a model, as McClelland and Elman (1986) demonstrated in some detail, will identify an input like "bleasant" as the word "pleasant," because the degree of overlap with the form representation of "pleasant" is high and there is no other word form in the lexicon to which "bleasant" is a better fit.

---

This research emerged out of regular discussions with Uli Frauenfelder and Stephen Marcus, whose stimulus and advice we gratefully acknowledge. We also thank Stef van Halen for his valuable help with many aspects of this research and James Cutting, Jeff Elman, and Aditi Lahiri for their comments on the manuscript.

This work was previously reported at the 28th Tagung experimenteller Psychologen, Saarbrücken, West Germany, March 1986, and at the Joint Conference of the Experimental Psychology Society (United Kingdom) and the Societa Italiana di Psicologia, Padova, Italy, April 1986.

Correspondence concerning this article should be addressed to William Marslen-Wilson, MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, England.

Contrasts between the cohort model and models like TRACE raise general questions about the directionality of lexical access. The purpose of the present research was to answer these questions—to find out whether there is a strong temporal directionality in lexical access and whether the on-line decision process does tolerate the late entry into the decision space of candidates that mismatch early in the word. To answer these questions we evaluated the consequences of partial matches between inputs and representations for two kinds of cases. We contrasted cases in which the partial match included the beginning of the word and cases in which it did not, that is, when there was a rhyme match. For each instance of partial match between an acoustic input and a lexical form representation, our question was the same: To what extent does this partially matching input produce activation of the word in question, and how does this activation compare with the activation of the word when there is a complete match between the stimulus and the relevant lexical form?

### Previous Research

We already have information from our earlier research (Marslen-Wilson, Brown, & Zwitserlood, 1989; Zwitserlood, 1985, in press) about the effects of word-initial partial matches, that is, cases in which the input matches a given lexical representation from word onset. This research, as well as the present experiment, used a cross-modal priming task (Swinney, 1979) to measure the degree of lexical activation during the processing of spoken words. In this task subjects listen to spoken words, heard either in isolation or in sentence contexts, and simultaneously see visual probe words, which stand in some relation (usually semantic in nature) to the spoken word of interest in the experiment (e.g., Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Swinney, 1979). The spoken word is intended to prime the visual probe word. The subjects' task is to make some response to this visual probe—usually a lexical decision judgment—and the timing of this response is compared with a control or baseline condition.

Our reasons for using this task were twofold. First, the use of a visual probe allowed us to test during the processing of the spoken word, whereas an auditory probe could be presented only after the offset of the word. Second, the use of a cross-modal semantic probe ensured that what we were testing was activation at the lexical level, and not activation at sublexical levels of representation. For example, in the phonological priming study reported by Ślowiacek, Nusbaum, and Pisoni (1987), in which stimuli like *dress* were used as primes for words like *dream* in a perceptual identification task, any resulting priming effect could be located either lexically or sublexically. Because the initial auditory sequence for the prime word is identical to the initial auditory sequence for the probe word, any facilitation in response to the probe could be due to repetition priming at the phoneme or feature level, without implicating the lexical level at all. In contrast, if an auditory prime, such as *dog*, facilitates responses to a visual probe, such as CAT, this can only be because of phenomena occurring at the lexical semantic level, because it is only at this level that *dog* and *cat* are related; they have no form

parameters in common that might cause priming at a sublexical level. It is worth noting that, although the TRACE model has no semantic level of representation, McClelland and Elman (1986) recommended the cross-modal priming paradigm as a means of tracing the activation of lexical elements in a model of the type they described.

In our earlier research investigating word-initial partial matches, we used two different versions of the cross-modal priming task. In the first sequence of experiments (Marslen-Wilson et al., 1989), we presented subjects with complete spoken words, such as *kapitein* (meaning *captain* in Dutch) which shared a large initial overlap with words like *kapitaal* (meaning *capital*). The mean amount of overlap, estimated by counting the number of shared initial segments, was 3.3.

Concurrent with these spoken words, the subjects also saw visual probes, to which they were required to make lexical decision responses. In the critical conditions, the visual probes were presented at a point at which the input was still compatible with both possible words, for example, before the release of the [t] in *kapitein* or *kapitaal*. The probe was either associatively related to *kapitein* (e.g., BOOT, meaning *ship*) or to *kapitaal* (e.g., GELD, meaning *money*). Reaction times to make lexical decisions to these visual probes were compared with response times to the same probes presented in the middle of control words to which they were not associatively related. In addition, the probes were also presented at the ends of the critical words and their controls.

The results showed that both probes were facilitated, relative to the control condition, when they were presented in the middle of the spoken word. When either GELD or BOOT was presented during the [t] of *kapitein*, not only was the probe related to *kapitein* facilitated, but so also was the probe (GELD) related to the other word that was still possible at the moment that the probe was presented (i.e., *kapitaal*). In contrast, when the probes were presented at the end of the word, only the probe related to the word actually presented was facilitated. Thus, at the end of *kapitein*, BOOT was facilitated but not GELD.

This research (Marslen-Wilson et al., 1989) was originally conducted to provide evidence for the multiple activation of semantic codes early in the word recognition process, as specified in the cohort model (c.f. Marslen-Wilson 1984, 1987). From the perspective of the current research, what this research demonstrates is that a partial word-initial match between an input and a lexical representation is sufficient to activate this representation to the degree necessary to produce priming in a cross-modal task of this nature. A second set of experiments (Zwitserlood, 1985, in press) makes this point even more clearly.

These experiments also used cross-modal semantic priming to investigate multiple lexical activation during the processing of a spoken word. But they differed from the first study by presenting the spoken words in a sentence context and by presenting only fragments of these words. Subjects would hear a fragment, for example, of the word *kapitein*, preceded either by a constraining sentential context or by a neutral carrier phrase. We cite here only the results for the carrier phrase condition for cases in which the fragment heard (e.g., [kapi:]) was still short enough to be compatible with two

possible words (such as *kapitein* and *kapitaal*).<sup>1</sup> The average overlap between input and target lexical representation was 2.9 segments, counting from word onset. Under these conditions, Zwitserlood obtained a mean facilitation effect of 29 ms for the probe related to the word actually occurring (BOOT, in the *kapitein* case) and 36 ms for the probe (e.g., GELD) related to its close competitor (e.g., *kapitaal*). This compares favorably with the effect of 40 ms for probes occurring after all of the word had been presented, in which only the probe related to the word actually heard was facilitated.

We take Zwitserlood's result as confirming the effectiveness of word-initial partial matches in activating lexical representations, and we will take the size of the effect she obtained as an appropriate estimate of the sort of effect that should be looked for in a cross-modal priming experiment involving partial matches between inputs and lexical representations. We turn now to the description of such an experiment.

### Experimental Considerations

The question at issue is whether directionality of mapping—that is, whether or not inputs match from word onset—is as important in lexical access as the total amount of input that matches a given lexical form representation. We tested this in the present study in three ways. Let us take as an example the word *honing* (meaning *honey*), which has as a close associate the word *bij* (meaning *bee*). In the first experimental condition, called the original-word condition, there was a complete match between acoustic stimulus and lexical element: Subjects heard the word *honing*, and at its offset saw the visual probe word BIJ, to which they made a lexical decision response. The degree of facilitation of this response—that is the amount to which the spoken word cross-modally primed the visual probe word—was assessed relative to response times to BIJ when it followed a word that was both semantically and phonetically unrelated to *honing*, for example, the word *pakket* (meaning *parcel*). This was the real-word control. The amount of facilitation induced by the original word, relative to this control word, was the baseline for assessing the effects in the partial match conditions.

In the first partial match condition, called the real-word rhyme condition, the target word BIJ was presented following a spoken word that rhymed with *honing*, in this case the word *woning* (meaning *dwelling*). The question here concerns the relation between this rhyme prime and the form representations of *woning* and *honing*. The representation for *woning* will be fully activated. But will *honing*, which matches the input perfectly except for the initial segment, be activated as well? If *honing* is activated, then it will prime the visual probe BIJ and therefore facilitate the lexical decision response to this probe.

For the real-word rhyme condition as a whole, the average amount of overlap between the stimulus word and the word related to the visual probe was 4.5 segments, counting from the second segment on. If the simple amount of matching input is the critical variable in determining amount of activation, then a rhyming stimulus like *woning* should facilitate responses to BIJ at least as much as a stimulus like [kapi:] should facilitate responses to GELD and BOOT. The com-

parison condition for the real-word rhyme was, as before, responses to BIJ following the real-word control.

The experiment also needed to include a second partial match condition, using rhyme primes that are not real-words. This is because models like TRACE allow elements at the same level of the system to inhibit each other. When the lexical nodes in the system are activated on the basis of input from the phoneme level, and when this level of activation exceeds a certain threshold, they begin to inhibit their neighbors in the network. The more strongly an item is activated, the more it inhibits the level of activity of its closest competitors. This means that if a stimulus like *woning* is heard, it will inhibit competitors like *honing*. Even if the input to the lexical level is sufficient to activate the lexical nodes corresponding to both *honing* and *woning*, the inhibition from the item that best fits the input will reduce the level of activation of its competitors. This means that BIJ would only be weakly facilitated, if at all, following *woning*.

Because of this possibility of interlexical inhibition, we needed to include a nonword rhyme condition, in which the rhyming stimulus is itself not a real-word. In the *honing* case, for example, we used the nonword rhyme *foning*. Because by definition there is no lexical node corresponding to a nonword, and there is no inhibition between levels, the presentation of *foning* cannot inhibit the bottom-up activation of *honing*. If lexical access is nondirectional, then a nonword rhyme like *foning* should facilitate responses to visual probes like BIJ, given the amount of matching input involved.

Two additional points need to be made. The first concerns the role of mismatch information in lexical access. A prime like *foning* not only partially matches a word like *honing* but also partially fails to match with it, since it begins with an [f] rather than an [h]. We assumed in the present study (as in TRACE and in recent versions of the cohort model, Marslen-Wilson, 1987) that this mismatch has no negative or inhibitory consequences for the representation of *honing*. Note that when there is mismatch late in a word, as in the *kapitein/kapitaal* cases, the mismatch does not itself directly inhibit or suppress the mismatching candidate. When the final syllable of, say, *kapitein* is heard, its effect is to boost the activation of *kapitein*, whereas the activation of *kapitaal* immediately starts to decay. In a model like TRACE, lateral inhibition from *kapitein* may then accelerate the decay of activation for *kapitaal*, but this is an effect that is mediated through the lexicon. It is not a direct effect of the mismatch on the representation with which it mismatches. What these assumptions meant for the present experiment was that nonword rhyme primes like *foning* were directly comparable, as far as the activation of their target representations was concerned, with word-initial primes like *kapi*.

The second point is that it cannot be the case, given the results we already have, that partially matching nonword inputs like *foning* can cause sufficient mutual inhibition between competing candidates (such as *honing* and *woning*) to

<sup>1</sup> In Zwitserlood's version of the task, the spoken input terminated at the end of the word fragment. In different conditions, of course, subjects heard longer or shorter fragments (cf. Zwitserlood, 1985, in press).

prevent priming of semantically related visual probes. The fact that an input like [kapi:] can prime the visual probes associated with both *kapitein* and *kapitaal* means that representations activated by partial inputs do not cancel each other out at the lexical level. If the generation of inhibition is based on a threshold, then partial inputs may not be sufficient to take the activation level above this threshold. There is therefore no reason, for a primarily nondirectional model, why a nonword rhyme should not prime the semantic fields associated with the lexical forms with which it partially matches.

Finally, the use of a nonword prime raises the possibility that the real-word control will not be the appropriate baseline for evaluating effects in the nonword rhyme condition. Hearing a nonword may have effects on performance over and above those of partial matching with lexical representations. To allow for this possibility, we included an additional baseline condition, the nonword control condition. Subjects also saw visual probes like BIJ presented at the offset of nonwords, such as *dakket*, which have no phonetic similarities with their paired nonword rhymes.

In summary, the present experiment contrasted the hypothesis that lexical access is strongly directional, so that the beginnings of words determine the entry of word candidates into the decision space, with the hypothesis that word onsets do not have a special status and that overall goodness of fit is the most important determinant of the system's behavior. To do this, we examined the priming of a visual probe in three prime conditions: the original word condition, in which the match between the input and the prime word was complete, and two partial match conditions, called real-word rhyme and nonword rhyme. Lexical decision times in these three conditions were compared with two baseline conditions (real-word control and nonword control) in which there was no match at all with the original word.

## Method

### Subjects

Sixty subjects participated in the experiment. They were all native speakers of Dutch and were paid for their participation. The subjects were randomly assigned to one of the experimental versions, six subjects per version.

### Materials

**Test items.** The materials were constructed around 50 sets of rhyming pairs (such as *honing/woning*) and their associated visual probes. These 50 pairs were selected from a larger set of 264 words (132 pairs) for which association norms were collected in a pretest. The criteria for the selection of these materials will be described first.

The original pool of 132 pairs of rhyming words was constructed on the basis of an exhaustive dictionary search for rhyming pairs that met the following criteria: The two words were at least two syllables in length and differed only in their first segment, both members of the pair had the same form class, neither word was a recent loan word or was judged very uncommon, and both words were matched for estimated frequency of occurrence (Uit den Boogaard, 1975).

Because association norms were not available for these words, it was necessary to conduct a pretest, in order to have a basis for

selecting visual probes for use in the cross-modal priming task. The 264 words (132 pairs) were presented in written form to 60 subjects in a free-association test. The subjects were asked to write down as a response the first word that came to mind when they read each word.

Given the resulting associations, we could find 50 rhyming pairs that had appropriate associates. The criteria used for selecting the associates were (a) an overall associative strength of 10 or more responses out of 60 (i.e., more than 16%), (b) similar associative strength for both words in a pair, and (c) no semantic overlap between the associates chosen for each member of a rhyming pair. A probe like BIJ was selected not only because of its association with *honing*, but also because it was never given as a response to *woning*, the other member of the rhyming pair. On the basis of these criteria, 100 associates (one for each member of 50 pairs of rhyming words) were selected to serve as cross-modal visual probes. The mean association strength between these probes and the spoken primes was 23.6 (39%). The mean phonetic overlap between pairs, in terms of number of shared segments, was 4.5.

The 50 rhyming pairs had the following syntactic properties. In 39 pairs, both words were nouns; 9 pairs were verbs (in their infinitive form), and 2 pairs were adjectives. The nouns and adjectives were all two syllables in length, with one three-syllable exception. The verbs were all three syllables in length.

These rhyming pairs, with their associated visual probes, formed the basis for 100 sets of five spoken primes, with each set associated with one visual probe. These five primes corresponded to the five experimental conditions, as illustrated in Table 1.

In Condition 1 (original word), the subject heard one member of a rhyming pair (e.g., *honing*), followed by its associated visual probe (e.g., BIJ). In Condition 2 (real-word rhyme), the subject heard the other member of the rhyming pair (e.g., *woning*), followed by the visual probe (e.g., BIJ) associated with the original word.

In Condition 3 (nonword rhyme), the subject heard a rhyming nonword (e.g., *foning*), followed by the same visual probe. The initial segment of these nonword rhymes was made as phonetically distinct as possible from the initial segment of the original word and the real-word rhyme, within the constraints imposed by other real-words that might also rhyme with the original word. The "nonword point" for these stimuli (the point at which they diverged from existing Dutch words) was set as early as possible, and always at least two segments before the end of the stimulus ( $M = 2.9$  segments). This ensured that the lexical item to which the nonword rhyme was the closest match was always the original word and not some other word belonging to the same cohort as the nonword rhyme. Although the initial segments of a nonword prime, such as *foning*, might match the beginning of a word like *fonetiek*, the complete nonword would be a better match to *honing* (or *woning*) than to any of its word-initial cohort members.

Conditions 4 and 5 (real-word control and nonword control, respectively) contained stimuli matched in length and syllabic structure to the test stimuli, but had no segments in common and (in the case of the real-word control) had no associative or semantic relation to the visual probe.

Table 1  
Sample Stimulus Set

Experimental condition	Spoken prime	Visual probe
Original word	<i>honing</i>	BIJ
Real-word rhyme	<i>woning</i>	BIJ
Nonword rhyme	<i>foning</i>	BIJ
Real-word control	<i>pakket</i>	BIJ
Nonword control	<i>dakket</i>	BIJ

Each of the 50 such sets was matched to a parallel set (assigned to a different version of the experiment) in which the other member of the rhyming pair became the original word. *Woning*, for example, would become the original word, with the associate *HUIS* (meaning *house*) now being used as the visual probe, so that *honing* would now function as the real-word rhyme, with *foning* remaining as the non-word rhyme.

**Fillers.** One hundred fifty filler items (spoken word/visual probe pairs) were also constructed. Ninety of the spoken fillers were real-words, and 60 were nonwords, reflecting the distribution of spoken words and nonwords in the test conditions. Most of the real-word fillers were verbs and adjectives, to compensate for the fact that most spoken test words were nouns. Since most of the test words were two syllables long, the filler stimuli were either one or three syllables long. Of the 150 visual probes for the filler targets, 100 were orthographically legal nonwords and 50 were real-words. The lengths of the word and nonword filler targets matched the length of the test targets.

An additional 48 item sets were chosen for use as practice. This practice list was made up of 24 real-word and 24 nonword visual targets, in combination with spoken real-words and nonword primes.

### Design and Procedure

The 100 stimulus sets (see Table 1) were split into two versions with one member of each probe pair assigned to each version. Within each version, the stimuli were rotated across five tapes, so that each visual probe (such as *BIJ*) appeared only once in each tape and appeared in conjunction with all five conditions across the five tapes. The 50 test and control trials on each tape were pseudorandomly interspersed with the 150 fillers, with the fillers occurring in the same position in each tape.

The 10 experimental tapes were constructed by cross-recording from two master tapes, of the filler and test items. The master tapes were recorded by a woman native speaker of Dutch. On the second channel of each experimental tape, a timing pulse was set, concurrently with the acoustic offset of the spoken primes. This timing pulse served to trigger the exposure of the target word on a CRT screen, and to start the counter modules that registered the subjects' reaction times. The presentation time for the visual probes was 50 ms (unmasked), with a 6-s interval between trials.

The subjects were tested in groups of 2–4, each seated in separate carrels. They were instructed to listen carefully to the spoken materials on tape and to decide, as quickly as possible, whether the string of letters that was presented after each spoken stimulus was a real-word or not. On each trial, subjects heard a spoken word or nonword binaurally over closed-ear headphones, followed immediately by the presentation of the visual target for lexical decision. They responded by pressing the JA (yes) or NEE (no) button on the response box in front of them. Each test session lasted about 35 min.

### Results and Discussion

Of the 3,000 lexical decision times we collected, 212 (7.1%) were errors (answering no to a real-word). These error responses were deleted from the matrix and replaced by the mean of the item and subject means for the relevant condition. A further 2.7% extreme values, defined as reaction times that were more than two standard deviations outside the relevant subject and item means, were replaced using the procedure recommended by Winer (1971). Item and subject means were then calculated for the five experimental conditions and entered into two one-way analyses of variance (ANOVAS), on items and on subjects, with the single factor of

prime type. The  $F'_{min}$  ratio for prime type was then computed,  $F'_{min}(4, 597) = 3.423, p = .009$ .

The overall results, summarized in Figure 1, are straightforward. As reflected in a Newman-Keuls test (using an error term derived from the  $F'_{min}$  ratio), only the original-word condition, in which there was complete match between the input and the word associated with the visual probe, produced significant cross-modal semantic priming effects. Comparing the original word with the real-word control condition, we obtained a facilitation effect of 32 ms, which is significant at the .01 level in the Newman-Keuls and is comparable to the effects we have obtained in other cross-modal priming studies. Zwitserlood (1985, in press), for example, obtained a 40-ms facilitation for end-of-word probes and a mean of 33 ms for midword probes.

In neither of the two partial match conditions was there any significant facilitation. For the real-word rhyme condition there was an 11-ms difference. For the nonword rhyme, depending on which baseline condition was chosen, the difference was either 10 ms (relative to the real-word control) or 4 ms (relative to the nonword control). Both real-word and nonword rhyme conditions were significantly slower than the original-word condition ( $p < .05$ ), but did not differ from each other. In fact, the absolute reaction times for the two rhyme overlap conditions, were almost identical (547 vs. 548 ms). Given the similarity between the real-word and nonword rhyme conditions, and in order to simplify the subsequent analysis, we used the same baseline (real-word control) for both of the partial match conditions in the analyses described below.

These overall means show that the directional story is broadly correct. Primes that do not share word onsets with the relevant lexical form representations are much less effective than primes that do—whether compared with the complete match conditions in this experiment or with the word-initial partial primes used in earlier research. Other aspects of the results, however, suggest that this is not the whole story and that rhyme primes can have consequences for the activation states of the words that they partially map onto. One

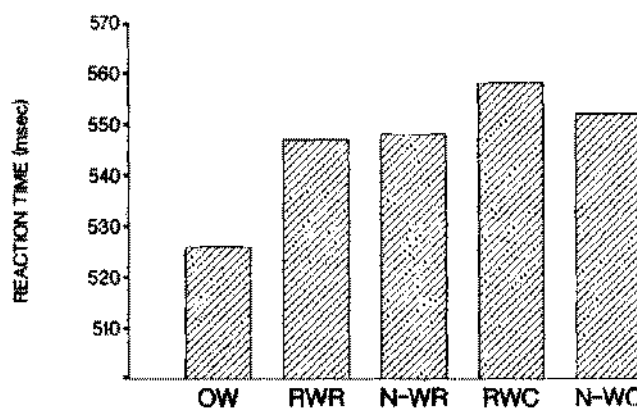


Figure 1. Mean lexical decision times to the visual probes in the five priming conditions: original word (OW), real-word rhyme (RWR), nonword rhyme (N-WR), real-word control (RWC), and nonword control (N-WC).



hint that this might be so comes from an analysis of the correlations between response times across the different test conditions.

Looking at the overall item means, there were significant correlations between the two partial match conditions (real-word rhyme and nonword rhyme) and the complete match (original word) condition, of .38 and .47, respectively. These correlations survived in a partial correlation analysis, which factored out the component of shared variance due to the properties of the visual probe. This analysis was run using the real-word control item means as the basis for partialing out the relevant components of the shared variance. The rationale for doing this was that the real-word control primes had no phonetic or semantic properties in common with the three test conditions. Therefore, any variance that this control condition had in common with the test conditions should have been due to the properties of the visual probe. With this aspect of the shared variance factored out, there was a correlation of .28 ( $p = .002$ ) between the real-word rhyme and the original-word conditions and of .34 ( $p < .001$ ) between the nonword rhyme and original-word conditions. This indicates that rhyme primes do have some effect on the lexical representations of items with which they rhyme, even if this produces only a small overall increment in facilitation.

We now turn to some subsidiary analyses, designed to find out whether there were any specific conditions under which rhyme primes were more or less effective in activating lexical representations. We look first at the effects of amount of overlap, measured in segments, between the rhyme primes and the original words. Even though the average number of overlapping segments (4.5) was greater than in either of the word-initial partial match experiments described earlier (Marslen-Wilson et al., 1989; Zwitserlood, 1985, in press), it could be argued that the rhyme primes did not match lexical representations quite as well as the simple count of segments would suggest. In particular, the coarticulation of the initial (mismatching) consonant with the following vowel might make at least the onset of the vowel a poor fit to the relevant lexical representation. This is consistent with some recent results by Warren and Marslen-Wilson (1987) and with the version of TRACE described by Elman and McClelland (1986).

#### Amount of Segmental Overlap

We were able to evaluate the role of amount of overlap by reanalyzing the results we already had, given the large number of items used and the range of different lengths that they covered. In fact, it was possible to subdivide the stimulus sets into four major overlap categories, ranging from three segments overlapping (as in *herrie/merrie/lerrie*) to six or more overlapping (as in *handelen/wandelen/jandelen*). The number of items in each overlap group varied as follows: For the three, four, five, and six-or-more groups, the number of items was 16, 48, 12, and 16, respectively. Even on the most conservative estimate of the amount of effective overlap, the large-overlap groups would provide at least as much matching input as the word-initial partial primes in earlier experiments, which averaged between 2.9 and 3.3 segments overlap.

The test-control difference scores for the items in the four groups were entered into an unequal- $N$  ANOVA, with the factors length (three, four, five, and six or more shared segments) and prime type (original word, real-word rhyme, nonword rhyme). In this and all subsequent post hoc analyses, subsets of items are compared. Since the original design was fully counterbalanced, subjects did not contribute to each item in each condition. Therefore, all analyses were based on the item means only, normalized for subject variation. We excluded from this data set eight items that had high error rates, leaving 92 sets. As Figure 2 shows, the results were clear-cut. There was a significant effect of prime type,  $F(2, 176) = 8.16$ ,  $p < .001$ , with the amount of facilitation averaging 20 ms more for the original word than for either of the two rhyme primes. But there was no effect of length,  $F(3, 88) = 1.33$ , and no interaction with prime type ( $F < 1$ ). The two partial-match conditions, which behave in very similar ways across duration groups, are collapsed into one group in Figure 2. The amount of facilitation did vary across length groups, but this pattern seemed to have nothing to do with the length variable per se, and did not interact with type of prime.<sup>2</sup>

These results make it unlikely that the failure of rhyme primes to produce facilitation was because they did not provide amounts of matching input comparable to the word-initial primes discussed earlier (such as the [kapi] example).

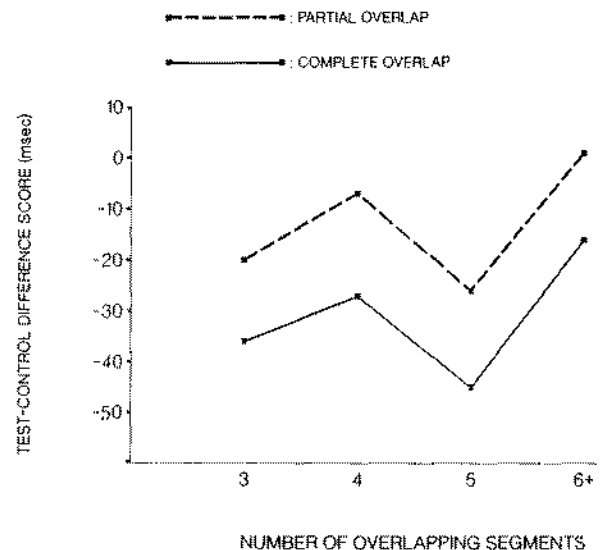


Figure 2. Mean facilitation effects for full primes and rhyme primes as a function of the amount of overlap (in number of segments) between the prime and the original word. (The results for the rhyme primes are collapsed together into the partial overlap group [broken lines].)

<sup>2</sup> If we look at the absolute means for the complete and partial overlap conditions, then we see these did not vary at all across conditions. Going from small to large overlap, the means for the original-word conditions were 524, 523, 523, and 528 ms, respectively, and for the two partial overlap conditions (collapsed together), the means were 541, 542, 545, and 545 ms.

The longer rhyme primes, with five or six segments of overlap, provide ample amounts of matching input and yet are no more facilitating, and no closer to the original word, than the rhyme primes with only three segments of overlap.

It is also apparent from these results that the degree of activation of a lexical representation seemed to be independent of the simple amount of matching input involved. This showed up in two ways. First, there was no general increase in the overall level of facilitation for the longer overlap groups, either for the original-word group or for the two rhyme groups. Second, there was no evidence of any convergence between the complete and partial match groups as the amount of overlap increased. On a straightforward activation story, the level of activation of a lexical node (or of the array of nodes corresponding to a lexical entry) is related to the amount of excitatory input that this node receives. This predicts that the relative disparity between rhyme primes and full primes should diminish as a function of the increase in the amount of positively matching input that they had in common. We find instead that the difference between the full prime and the partial rhyme primes remained constant as the number of overlapping segments increased from three to six or more.

### *The Competitor Environment*

In current models of lexical access, the competitor environment within which a lexical representation is activated is an important factor in determining the response of the system. There are two aspects of this competitor space that might affect the effectiveness of a rhyme prime: (a) the acoustic-phonetic distance of the initial segment of the prime from the initial segment of the word with which it rhymes and (b) the number of near competitors.

We cannot evaluate here the effects of acoustic-phonetic distance, because the rhyme stimuli were selected so as to reduce to a minimum any variation along this dimension. As far as possible, we chose the stimuli to maximize the distinctiveness of the onsets of the rhyme primes and of the original words. For almost all the item sets, the initial segment of the nonword rhyme differed from the original word in both manner and place of articulation. A similar pattern held for the real-word rhymes and the original words, with nearly 80% of the pairs having initial segments that differed in both manner and place.

The second competitor variable, which the current stimulus set does allow us to evaluate, is the overall number of competitors that a prime word had to contend with. The original stimulus sets varied in terms of the number of words in the language with which the original word rhymed. We shall refer to these words as *rhyme competitors*. Thirty of the sets had only 1 rhyme competitor (i.e., the matched real-word rhyme). For the other 70, the number of competitors varied from 2 to 23. For example, the word *kettel* (meaning *kettle*) had three rhyme competitors (*zetel*, *betel*, and *netel*).

In a model in which there is simultaneous propagation of bottom-up activation to all appropriate lexical nodes, it is possible for the number of nodes that are good matches to the input to affect how any individual node will respond. This might especially be the case for the nonword rhymes, for

which the goodness of fit will be more nearly equal for all rhyme competitors. In fact, it is only for cases in which there are just two rhyme competitors that the nonword rhyme will be, strictly speaking, in the same position as the word-initial partial match stimuli used in our earlier research. (The stimuli for the word-initial overlap experiments were selected, as far as possible, so that there would be only two candidates left at the midword probe point.) Perhaps, then, this is the appropriate condition for assessing the effects of partial overlap, and not conditions in which there are many more lexical nodes competing (the mean number of rhyme competitors for the entire set was 4.5).

We therefore selected three subsets of the data: the 30 sets in which the original word had only one competitor (the real-word rhyme), the 30 sets that had the largest number of competitors (ranging from 6 to 23 and averaging 10.0), and an intermediate set of 30, in which the number of competitors ranged from 2 to 4 ( $M = 3.1$ ). The test-control difference scores for the three sets were entered into a two-way ANOVA, with the factors competitor (low, intermediate, or high) and prime type (original word, real-word rhyme, or nonword rhyme). The three competitor-size groups were well matched for association value (i.e., the strength of the prime-target associative relation), with means of 23.0%, 22.4%, and 21.0% going from low to high competitor, as well as for amount of overlap (number of segments shared with the original word), with means of 4.9, 4.5, and 4.3, respectively.

There was a strong main effect of prime type,  $F(2, 174) = 11.612$ ,  $p < .001$ , with the difference scores being 20 ms larger for the original-word condition. There was also a consistent competitor effect, with the advantage of test over control averaging 30 ms in the low-competitor group, and falling to 9 ms in the intermediate group and 10 ms in the high-competitor group. Although this effect was not significant on the overall analysis,  $F(2, 87) = 1.935$ ,  $p > .10$ , subsequent post hoc tests (Newman-Keuls), showed significant differences between facilitation effects in the low-competitor condition and those in the other two conditions, both for the original word and for the rhyme primes (Wilcox, 1987). There was, however, no sign of an interaction between competitor and prime type,  $F < 1$ . The two rhyme conditions (real-word and nonword) behaved in very similar ways across competitor groups.

This overall pattern of results is illustrated in Figure 3, with the two rhyme conditions collapsed into a single partial overlap group. The results show, first, that variations in the number of competitors did not affect the advantage of full primes over rhyme primes, confirming the overall correctness of the directionality hypothesis. On the other hand, the competitor environment did seem to matter. When the spoken input was a good match to only two lexical representations, then the overall level of activation (as measured in these cross-modal probes) was higher for both the complete match and partial match conditions. For both the original-word and combined-rhyme groups, the facilitation effects differed significantly from zero. But for conditions in which the number of close competitors was larger than two, there was no effect of partial match at all. Only the scores for the original word differed significantly from zero.

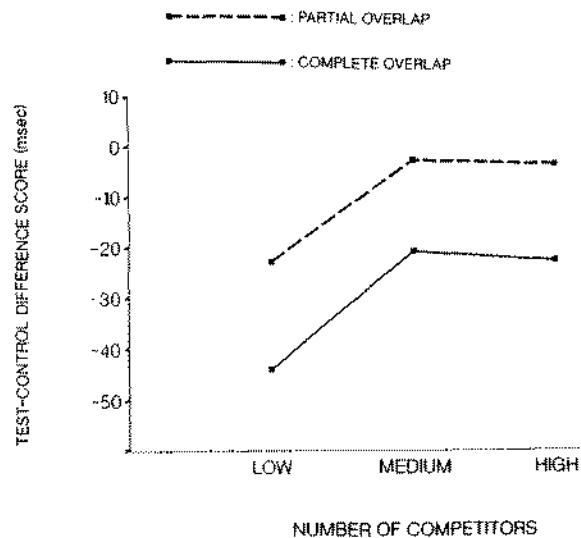


Figure 3. Mean facilitation effects for full primes and rhyme primes as a function of the number of rhyme competitors (low, medium, or high). (The results for the rhyme primes are collapsed together into the partial overlap group [broken lines].)

### Conclusion

The dominant effect in the present results is directionality of mapping. Primes that mismatched at word onset with the relevant lexical form representation were always less effective than primes that did not mismatch word-initially—whether compared with the complete match conditions in this experiment or with the word-initial partial primes used in earlier research. This held true independently of the amount of overlap between a partial prime and the relevant original word. Even with words three syllables long, in which only the first consonant was mismatching, we still saw no significant overall priming effect.

This result is inconsistent with any simple global notion of the nondirectional mapping of stimulus information onto lexical representations. On any story of this sort, there should be not only an effect of rhyme primes at all levels of overlap, but also a general increase in the level of facilitation induced by rhyme primes as the amount of overlap increases. There was no sign of this in the present data; nor was there any sign that the difference in effectiveness between a rhyme prime and the original word diminished with increasing overlap.

Furthermore, there was no detectable effect of the lexical status of the rhyme prime. In terms of their effectiveness or ineffectiveness as partial primes, the real-word and nonword rhyme stimuli behaved in the same way throughout. Although a model like TRACE could account for the absence of any partial match effects for the real-word rhymes (via the mechanisms of lateral inhibition), it would have difficulties with the absence of any effect for the nonword rhymes as well.

In general, the results suggest that word onsets do have a special status in spoken word recognition. Perceptually, sequences that mismatch at word onset do not seem to enter normally into the on-line decision space. Even in the maximal

overlap conditions, the rhyme prime never catches up with the original word. It is never, apparently, treated as if it were the original word. Although the listener may be able to determine, on-line, that the nonsense word *bleasant* is intended as a token of the string *pleasant*, this does not mean that the perceptual experience of *pleasant* is substituted for the percept of the mispronounced token; nor does it mean that this overriding of the original mispronunciation is without cost for the processor.

### Evidence for Late Entry

Accompanying the dominant effect of directionality there were, nonetheless, signs that under certain conditions rhyme primes could have detectable priming effects, that word-initial mismatch does not provide an absolute block to entry into the decision space. This showed up, first of all, in the small but significant correlations between responses to the original word and responses to the rhyme primes, even after components of shared variance due to the visual probe had been factored out in a partial correlation analysis. The subsequent analyses suggest that this effect was strongest in the stimulus conditions in which the original word had only one rhyme competitor. Under these conditions, there was an across-the-board increase in the amount of facilitation of the visual probe, so that there was a significant effect not only for the original word but also for the two rhyme primes (see Figure 3).

The current data do not tell us exactly how to interpret this evidence for late entry nor why it interacts with the density of the competitor environment. One possibility is that the enhancement in sparse competitor environments reflects some form of gang effect (McClelland & Rumelhart, 1981), where the overall level of inhibition in the system rises as a function of the number of candidates that are active at any given moment. A second possibility focuses on the phonological similarity between the onsets of rhymes and targets. In particular, the increased effectiveness of the nonword rhymes in the low-competition environments may reflect the increased likelihood that the rhyme prime will be more similar to the original word than to any of its rhyme competitors.

Neither of these accounts is completely successful, however, in accounting for the full pattern of results. Although it is likely that the effects at least partially reflect the relative phonological distance between rhymes and original words, this is a possibility that will have to be evaluated in future research, using stimuli explicitly designed to vary along this dimension.

### Information and Representation

Finally, we should consider three further accounts, each of which requires some reconsideration of the basic assumptions on which we conducted this research. The first of these concerns the ways in which different systems handle the occurrence of a mismatch between an input and a lexical representation. We assumed here that a mismatch has no negative or inhibitory effects, so that a word-initial mismatch



simply means the absence of positive input for the segment involved. There are some grounds, however, for treating the mismatch between an input and a representation as more than just the absence of a match and allowing it, instead, to have long-term negative effects. This would simplify the interpretation of the current results.

The available data for mismatches late in a word suggest that this information has a strong and immediate effect on the status of the word in question as an active candidate. In research using pairs of the *kapitein/kapitaal* type (Marslen-Wilson et al., 1989; Zwitserlood, in press), facilitation of a cross-modal semantic probe dropped off rapidly as soon as disconfirming information became available. Although [kapi] facilitated probes related to both *kapitein* and *kapitaal*, the sequence [kapitey] facilitated only the matching candidate (*kapitein*); the competitor appeared to be immediately dropped. Analogously, Slowiczek and Pisoni (1986) found that auditory primes like *dress*, followed at a 50-ms delay by auditory targets like *dread*, had no facilitatory effects on their targets (although in this case, of course, the research was looking at phonological rather than semantic priming).

These results are consistent with the view that mismatch information has a directly inhibitory effect. What the current research suggests is that if these effects are inhibitory, they remain in effect for the full duration of the word. Even for very long, multisyllable rhyme primes, there was no sign that the system was losing sight of the fact that the initial segment did not match. We might envisage, then, a processing system that, on the one hand, allows the mapping of the input onto all representations to which it matches, even if the match is only a partial one. On the other hand, however, because the system maintains a representation of the complete pattern of matches and mismatches between the input and the set of lexical representations that are currently active, it will not lose track of the fact that a mismatch has occurred. It does not confuse *mobility* with *nobility*, nor does it perceive *dribble* as a token of *dribble*. This allows it to generate the results we have observed here.

The second possibility concerns the extent to which a nondirectional model, such as TRACE, could in fact account for the present results. Although TRACE is not explicitly directional in its approach to lexical access, it does have an intrinsic directionality, in the sense that speech is ordered in time, so that word onsets are heard first. The presentation of TRACE in McClelland & Elman (1986) may underestimate the significance of this temporal ordering in determining the behavior of the system. Elman (personal communication, October 29, 1987) has since run some additional simulations using non-word inputs like [biss] and [wist], which partially match the real-word *beast* either word-initially (as in [biss]) or word-finally (as in [wist]). In each case, there is three segments overlap with the relevant lexical representations. Elman finds that the word-initial partial input strongly activates *beast*, whereas the rhyming partial input is very ineffective in activating this word, because of competition from other real-words that are activated word-initially by the sequence [wi].

This raises the interesting possibility that directionality effects can be accommodated in a processing model simply as a consequence of the temporal order in which information

arrives. This would reinforce the attractiveness of TRACE as a realization of a cohortlike recognition process. To evaluate this properly, however, we need a systematic investigation of how TRACE performs over the full range of stimulus types used in this experiment, with the appropriate contrasts in amount of overlap and types of competitor environment. In particular, it will be important to establish whether TRACE shows the same lack of sensitivity to the amount of stimulus overlap as we found in the present experiment for human listeners (see Figure 2).

The final possibility concerns the properties of the lexical representations themselves. The assumption of nondirectionality in processing requires the assumption of nondirectionality in representation. In common with almost all other research on speech processing and lexical access, the present experiments were based on the assumption that the abstract representation of lexical form is a sequence of independent representational units (such as segments), strung together like a string of beads. The representation itself, on this assumption, has no intrinsic sequential structure, so that any directionality in access is imposed from outside, by the properties of the processing system.

But one should also consider the possibility that the mental representation of lexical form does have intrinsically directional properties, so that there are sequential dependencies in the representation. Independent of any assumptions about the properties of the access process, these dependencies will impose a penalty on any stimulus input that does not enter the representational space at word onset. We have suggested that the processing system may need to operate in terms of the complete patterning of the input over the representational space. If this representation in some way captures the abstract properties of lexical forms as dynamic patterns in time, then it may be that these patterns can only be fully satisfied by inputs that share the same complete trajectory. In other words, we should think at least as carefully about the properties of representations as we do about the processes that map onto them.

## References

- Elman, J. L., & McClelland, J. L. (1986). Exploiting lawful variability in the speech waveform. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 360–385). Hillsdale, NJ: Erlbaum.
- Grosjean, F. (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception & Psychophysics*, 28, 267–283.
- Marcus, S. M., & Frauenfelder, U. H. (1985). Word recognition—uniqueness or deviation? A theoretical note. *Language and Cognitive Processes*, 1, 163–169.
- Marslen-Wilson, W. D. (1984). Function and process in spoken word-recognition. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 125–150). Hillsdale, NJ: Erlbaum.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25, 71–102.
- Marslen-Wilson, W. D., Brown, C. M., & Zwitserlood, P. (1989). *Spoken word-recognition: Early activation of multiple semantic codes*. Unpublished manuscript. Max-Planck-Institute, Nijmegen.

- Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, 8, 1-71.
- Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word-recognition in continuous speech. *Cognitive Psychology*, 10, 29-63.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18, 1-86.
- McClelland J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review*, 88, 375-407.
- Norris, D. (1981). Autonomous processes in comprehension. *Cognition*, 11, 97-101.
- Salasoo, A., & Pisoni, D. B. (1985). Interaction of knowledge sources in spoken word identification. *Journal of Memory and Language*, 24, 210-231.
- Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, 14, 489-537.
- Slowiaczek, L. M., Nusbaum, H. C., & Pisoni, D. B. (1987). Phonological priming in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 64-75.
- Slowiaczek, L. M., & Pisoni, D. B. (1986). Effects of phonological similarity on priming in auditory lexical decision. *Memory and Cognition*, 14, 230-237.
- Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 18, 645-659.
- Tyler, L. K. (1984). The structure of the initial cohort. *Perception & Psychophysics*, 36, 415-427.
- Tyler, L. K., & Wessels, J. (1983). Quantifying contextual contributions to word-recognition processes. *Perception & Psychophysics*, 34, 409-420.
- Uitden Boogaard, P. C. (1975). *Woordfrequenties*. Utrecht: Oosthoek, Scheltema & Holkema.
- Warren, P., & Marslen-Wilson, W. D. (1987). Continuous uptake of acoustic cues in spoken word recognition. *Perception & Psychophysics*, 41, 262-275.
- Wilcox, R. R. (1987). New designs in analysis of variance. *Annual Review of Psychology*, 38, 29-60.
- Winer, B. J. (1971). *Statistical principles in experimental design*. New York: McGraw-Hill.
- Zwitserslood, P. (1985, November). *Activation of word candidates during spoken-word recognition*. Paper presented at the annual meeting of The Psychonomic Society, Boston.
- Zwitserslood, P. (in press). The locus of effects of sentential-semantic context in spoken-word processing. *Cognition*.

Received October 29, 1987

Revision received August 15, 1988

Accepted August 17, 1988 ■

## Low Publication Prices for APA Members and Affiliates

**Keeping You Up-to-Date:** All APA members (Fellows; Members; and Associates, and Student Affiliates) receive--as part of their annual dues--subscriptions to the *American Psychologist* and the *APA Monitor*.

High School Teacher and Foreign Affiliates receive subscriptions to the *APA Monitor* and they can subscribe to the *American Psychologist* at a significantly reduced rate.

In addition, all members and affiliates are eligible for savings of up to 50% on other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the British Psychological Society, the American Sociological Association, and Human Sciences Press).

**Essential Resources:** APA members and affiliates receive special rates for purchases of APA books, including the *Publication Manual of the APA*, the *Master Lectures*, and *APA's Guide to Research Support*.

**Other Benefits of Membership:** Membership in APA also provides eligibility for low-cost insurance plans covering life; medical and income protection; hospital indemnity; accident and travel; Keogh retirement; office overhead; and student/school, professional, and liability.

For more information, write to American Psychological Association, Membership Services,  
1200 Seventeenth Street NW, Washington, DC 20036, USA.