

Delayed commitment in spoken word recognition: Evidence from cross-modal priming

PAUL A. LUCE and MICHAEL S. CLUFF
State University of New York, Buffalo, New York

Using the cross-modal priming paradigm, we attempted to determine whether semantic representations for word-final morphemes embedded in multisyllabic words (e.g., /lak/ in /hɛmlak/) are independently activated in memory. That is, we attempted to determine whether the auditory prime, /hɛmlak/, would facilitate lexical decision times to the visual target, KEY, even when the recognition point for /hɛmlak/ occurred prior to the end of the word, which should ensure deactivation of all lexical candidates. In the first experiment, a gating task was used in order to ensure that the multisyllabic words could be identified prior to their offsets. In the second experiment, lexical decision times for visually presented targets following spoken monosyllabic primes (e.g., /lak/-KEY) were compared with reaction times for the same visual targets following multisyllabic pairs (/hɛmlak/-KEY). Significant priming was found for both the monosyllabic and the multisyllabic conditions. The results support a recognition strategy that initiates lexical access at strong syllables (Cutler & Norris, 1988) and operates according to a principle of delayed commitment (Marr, 1982).

Despite the large number of representations of spoken words stored in memory, listeners are able to discriminate among these items rapidly and usually quite accurately in normal language understanding. The nature of the processes by which the listener converges on a single stored item has elicited much research and theory in the area of spoken word recognition.

The cohort theory of spoken word recognition (Marslen-Wilson, 1987, 1989) was one of the earliest models intended to account for the complex process of lexical discrimination. In this model, word-initial information (i.e., the first two or three phonemes; Marslen-Wilson & Welsh, 1978) activates a cohort of words that correspond with the acoustic input (see also Cole & Jakimik, 1980). As bottom-up input accrues across time (or, metaphorically, from left to right), the cohort members that match the input remain activated while "the level of activation immediately starts to decay" (Marslen-Wilson, 1989, p. 7) for the items that do not match. According to this model, the precise moment at which a word is recognized (the recognition point) occurs when a single word diverges from all other candidates in the lexicon (the divergence or uniqueness point). For example, a word such as /hɛmlak/ may be recognized at the point at which it diverges from its last remaining competitor, /hɛmlaɪn/, at the final vowel. When a word is produced in isolation, the divergence point and the recog-

nition point should be identical, although within a sentence, semantic information may move the recognition point farther to the left of the divergence point, thus enabling earlier recognition.

According to the cohort model, spoken word recognition

is based on a process of successive reduction of the active membership of the cohort of competitors. As more of the word is heard, the accumulating input pattern will diverge from the form specifications of an increasingly high proportion of the cohort's membership. This process of reduction continues *until there remains only one candidate* that still matches the sensory input. (Marslen-Wilson, 1989, p. 7; emphasis added).

Thus, the model proposes a strategy (at least implicitly) of strong commitment to lexical hypotheses. Although alternatives to the target word are activated prior to the decision point, once a decision is made, activation levels of alternatives begin to fall. Indeed, alternatives inconsistent with the sensory input may actually be inhibited (see Marslen-Wilson, 1989, note 3). Marslen-Wilson (1987) argues that such a strategy avoids computationally wasteful activation of alternative hypotheses.

Others have challenged the utility of a strategy of strong commitment, however. For example, behavioral evidence that words may be recognized only after their acoustic offsets (Bard, Shillcock, & Altmann, 1988; Grosjean, 1985) and computational evidence that uniqueness points prior to the offsets of words are relatively rare (Luce, 1986) indicate that a maximal commitment strategy may be less than optimal in most circumstances.

Marr (1982; see also Klatt, 1989) proposed that a principle of least (or delayed) commitment enables perceptual systems to be more robust. The withholding of all-or-none decisions means that minimal backtracking is required in the

This research was supported (in part) by Research Grant 1 R01 DC 0265801-A1 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health. We would like to thank Mara Goodman for her comments. M.S.C. is now at AT&T Laboratories. Correspondence concerning this article should be addressed to the first author at Language Perception Laboratory, Department of Psychology, State University of New York, Buffalo, NY 14150 (e-mail: paul@deuro.fss.buffalo.edu).

case of degraded or ambiguous input. Although delaying commitment to a word candidate may require a higher initial investment of resources through prolonged activation of alternative candidates, such a strategy also minimizes a costlier reinvestment of resources if the selected candidate eventually proves to be inconsistent with the incoming acoustic-phonetic information or the pragmatic or sentential context.

One of the advantages of delaying commitment to lexical hypotheses is that later occurring information may be used to resolve earlier ambiguous information. For example, Cluff and Luce (1990; see also Charles-Luce, Luce, & Cluff, 1990) demonstrated that disambiguating information acquired during the processing of the end of a word affects the processing of the beginning of the word. Specifically, they studied the perceptual identification of spondees, compound bisyllabic words in which each syllable receives strong stress. Cluff and Luce computed the similarity neighborhoods (see Luce, Pisoni, & Goldinger, 1990) of the component syllables of a large set of spondees. From this set, they chose spondees containing *easy* syllables (those in relatively noncompetitive neighborhoods) and *hard* syllables (those in relatively competitive neighborhoods). Maximal commitment models with a left-to-right processing emphasis suggest that spondees with an easy–hard syllable pattern should be recognized more easily than those with hard–easy patterns, because a first syllable that is easy to recognize should be more likely to aid in the resolution of a more difficult second syllable. A delayed commitment model would predict that later occurring information may be used to identify a difficult first syllable, thus boosting performance for spondees with a hard–easy pattern. Cluff and Luce found evidence supporting the latter prediction, but not the former.

Connine, Blasko, and Hall (1991), using monosyllabic words, found further evidence for the disambiguation of early information by later information. They presented to subjects ambiguous stimuli constructed on a voicing continuum from /dent/ to /tent/ within sentences. Disambiguating context biasing either /dent/ or /tent/ occurred either three or six syllables following the ambiguous token. Even after six syllables, commitment to an interpretation of the ambiguous input was delayed, suggesting that contextual information after the offset of the word can disambiguate alternative lexical hypotheses.

This evidence, taken together with evidence that lexical decisions are not made at divergence points (Goodman & Huttenlocher, 1988) and that word-final overlap affects same–different judgments (Luce, 1991), suggests that processing of spoken words is not strictly left-to-right. To some degree, information occurring later in time can affect processing of information that occurs earlier.

Other approaches, while not specifically designed according to a principle of delayed commitment, are less strict with respect to left-to-right processing. For example, the metrical segmentation strategy (MSS; Cutler, 1989,

1990; Cutler & Norris, 1988; Norris, McQueen, & Cutler, 1995) is particularly amenable to a strategy of delayed commitment. According to this model, lexical access is not necessarily initiated from word onsets but instead from strong syllables containing unreduced vowels. In later versions of the model (Norris et al., 1995), MSS does not *determine* lexical access but instead affects the degree to which candidates with strong onsets are aligned with the sensory input. The later model is still entirely consistent with the principle of delayed commitment, however. Evidence for this approach comes from word-spotting tasks (Cutler & Norris, 1988), as well as “slips of the ear,” in which the majority of misperceptions tend to reflect incorrect placement of word boundaries at strong syllables (as in *it was an eagle* for *it was illegal*; Cutler & Butterfield, 1992). This approach differs from typical left-to-right approaches because it is hypothesized that words with weak initial syllables are not resolved until information from a later occurring strong syllable is taken into account.

In short, the cohort model of spoken word recognition and the metrical segmentation strategy differ on two key sets of assumptions. First, they differ in their assumptions about the point at which lexical hypotheses are initiated. Second, they differ in their assumptions as to how strong the commitment to these hypotheses will be. The cohort model proposes that lexical hypotheses are initiated at word onsets and that strong commitments to a hypothesis are made at or before the divergence point. The MSS proposes that lexical hypotheses are initiated at metrically strong syllables and that commitment to lexical hypotheses may be delayed.

The problem of embedded words (e.g., *boy* and *cot* in *boycott*) may provide further insights into the nature of the process of lexical activation from spoken input. For example, Swinney (Prather & Swinney, 1977, cited in Swinney, 1981) found that initial embedded words (e.g., *boy* in *boycott*) produced facilitation in a cross-modal priming task, whereas final embedded words did not (e.g., *cot*). On the basis of these findings, Swinney proposed that the recognition system operates according to a “minimal accretion principle” that minimizes the number of hypotheses activated. According to this principle, the first potential word (e.g., *boy*) will be activated, although subsequent hypotheses (such as *cot*) will not be entertained if the ongoing information continues to support previously activated hypotheses. The minimal accretion principle is consistent with a strict left-to-right activation process and potentially incompatible with MSS.

Shillcock (1990) claimed that the minimal accretion principle underestimates the activation of embedded words. Using the cross-modal priming paradigm, he presented an auditory prime (e.g., *trombone*) followed by a visual probe that was either related to the homophone of the second syllable of the spoken prime (RIB for *bone*) or was unrelated (BUN). Lexical decisions to the probe revealed that the embedded morpheme, *bone*, primed RIB. In contrast to Swin-

ney's findings, these results argue against strict left-to-right processing models, demonstrating activation of lexical hypotheses of words embedded at the ends of carrier words.

Although compelling, Shillcock's (1990) study has potential shortcomings. In particular, Shillcock did not determine divergence points for his stimuli. In order to test a maximal commitment model properly, the divergence points for the stimuli must occur before offset, which would ensure that a word decision should have been made (according to the cohort model) and that activation levels of competing lexical candidates should be decaying or inhibited (Marslen-Wilson, 1989). In addition, because Shillcock used stimuli with weak-strong metrical stress patterns (all stimuli were British English), the finding that embedded final words were activated may have been due, in whole or part, to the fact that the final syllable was the only strong syllable in the carrier word. Predictions for MSS in these circumstances are clear. If the MSS is correct, lexical access could be initiated at the onset of *bone*, which would place the divergence point at the end of the word. This would predict priming of RIB by *bone*. Precise predictions for a strict left-to-right processing model, however, are not as straightforward when divergence points may or may not occur prior to the offset of the carrier word.

In order to further investigate the activation of embedded lexical items in spoken word recognition, we conducted a cross-modal priming experiment in which we used stimuli with strong-strong metrical patterns. For these particular stimuli, predictions from a left-to-right processing model are somewhat clearer: If a stimulus word begins with a strong syllable, activation of lexical hypotheses should start at the beginning of the word. Moreover, even if lexical hypotheses are activated at the first *and* second syllables, the presence of a divergence point prior to offset of the second syllable should result in identification of the carrier word and should cause deactivation of all other lexical hypotheses. This deactivation would eliminate embedded word-final morphemes from contention. However, a delayed commitment model using the MSS would predict that each strong syllable would initiate separate lexical searches, thus providing activation of an embedded word-final morpheme. Therefore, use of strong-strong stimuli (or subword morphemes beginning with strong syllables) provides a more direct test of the role of divergence points in the deactivation of lexical candidates.

The present study constitutes an attempt to expand Shillcock's (1990) paradigm to test in greater detail the predictions that differentiate strict left-to-right processing models from delayed commitment models. First, we attempted to determine whether noninitial subword units would be activated in multisyllabic words with strong-strong metrical patterns. Second, we attempted to determine the degree to which activation levels of alternative hypotheses would be maintained after the divergence point of the word. To test these hypotheses, we selected multisyllabic words with three requirements: (1) each morpheme began with a metrically strong syllable (such as /hem/ from /hemlak/

and /rækst/ from /rækətbəl/), (2) the final morpheme had at least one meaning unrelated to the word as a whole (e.g., /lak/ has a meaning unrelated to /hemlak/), and (3) the divergence point occurred during the final syllable.

Stimuli fitting these criteria were tested twice. In Experiment 1, they were presented to subjects in a gating task (Grosjean, 1980) to ensure that the words could be identified before their offsets. In Experiment 2, the same stimuli were presented as auditory primes in a cross-modal priming paradigm to separate groups of subjects. Targets were presented for each of the primes that were related to the alternative meaning of the final morpheme, were unrelated to the final morpheme (and the word as a whole), or were phonologically permissible nonwords. For comparison, in Experiment 2, the monosyllabic equivalents to the second morpheme in the multisyllabic stimuli were presented to other subjects in the same fashion.

EXPERIMENT 1

Method

Subjects. Twenty-six undergraduates attending the State University of New York at Buffalo participated for course credit. No subjects reported a speech or hearing disorder.

Materials. The stimuli in this experiment were auditorily presented multisyllabic words. Each word contained two morphemes and began with a metrically strong syllable. For example, the two morphemes in /rækətbəl/ each begin with strong syllables. In addition, the second morpheme of each stimulus word had at least one meaning unrelated to the stimulus word as a whole. For /rækətbəl/, /bəl/ can refer to a formal dance as well as to a spherical object.

In order to ensure that each stimulus word could, at least in principle, be uniquely identified during the second morpheme, we used a computerized version of Merriam-Webster's Pocket Dictionary (1964) to verify that each stimulus word diverged from all other lexical items prior to the end of the word. In addition, to ensure that the stimulus words and their component morphemes would be familiar to subjects, the stimulus words as well as their component morphemes were screened for familiarity. Both had subjective familiarity ratings of 5 or above on a 7-point scale (7 meaning *familiar with the word and its meaning*; Nusbaum, Pisoni, & Davis, 1984).

The spoken words were recorded in a sound-attenuated room by a male native speaker of English using an Electro-Voice D054 microphone. The stimuli were low-pass filtered at 4.8 kHz and were digitized at a 10-kHz sampling rate using a 12-bit analog-to-digital converter. Stimulus files were spliced and placed in individual files with the use of a digital waveform editor (Luce & Carrell, 1981).

Procedure. A gating task (Grosjean, 1980) was performed in order to obtain behavioral evidence that the multisyllabic stimuli could be recognized before their offsets. Stimulus presentation was controlled by a PDP-11/34 minicomputer, which presented stimuli via a 12-bit digital-to-analog converter over matched and calibrated TDH-39 headphones. The subjects were tested in a sound-treated room in individual booths equipped with Microterm 5510 terminals. Each multisyllabic word was presented auditorily to 26 subjects in successive 50-msec increments. After each stimulus increment, the subjects typed on a keyboard what they thought the word was. Presentation of increments continued until the entire word was presented.

Results and Discussion

On the basis of responses in the gating task, recognition points were determined for each stimulus. The recognition point was defined as the average gate at which sub-

Table 1
Examples of Stimuli Presented in Experiment 2

Auditory Prime	Visual Target		
	Related	Unrelated	Nonword
Monosyllabic Primes			
/læk/	KEY	DANCE	CHAGE
/bɔl/	DANCE	KEY	PID
/tɛlɚ/	BANK	BLUE	FACH
Multisyllabic Primes			
/hɛmlæk/	KEY	DANCE	CHAGE
/ɹækɛtbɔl/	DANCE	KEY	PID
/stɔɪtɛlɚ/	BANK	BLUE	FACH

jects correctly identified the stimulus word. In addition, the gate corresponding to the onset of the final morpheme was determined from a waveform display.

On the average, the onset of the final morpheme occurred at gate 6. Recognition points occurred, on the average, at gate 9. The average of the differences between the onset of the final morpheme and the recognition point—computed on an item-by-item basis—was 3.04. Thus, subjects correctly identified the target word three gates (150 msec) after the onset of the final morpheme. All stimulus words were recognized before their offsets.

Based on both computerized and behavioral measures, therefore, the stimulus words could be identified before the offset of the final morpheme. This ensured that any activation of alternative lexical hypotheses elicited by the second morpheme should occur after the multisyllabic word had diverged from all other lexical candidates.

EXPERIMENT 2

In Experiment 2, we used a cross-modal priming task with the stimuli that had been selected and tested in Experiment 1. Table 1 illustrates the design used in Experiment 2. In the monosyllabic prime condition, the second morpheme from the multisyllabic words tested in Experiment 1 served as auditory primes. Thus, for the multisyllabic word /hɛmlæk/, the morpheme /læk/ was paired with the related target KEY, which is related only to the second morpheme in /hɛmlæk/. The visual targets were related to the monosyllabic word as described above, were unrelated (DANCE for /læk/), or were phonologically permissible nonwords. In the multisyllabic prime condition, the words that had been tested in Experiment 1 served as primes. The targets corresponded exactly to those used for the monosyllabic primes. Here, the multisyllabic prime /hɛmlæk/ was paired with the target related to the second morpheme KEY, the unrelated target DANCE, or a phonologically permissible nonword.

We predicted that if the principle of maximal commitment should hold, we would observe a priming effect for monosyllabic but not for multisyllabic primes, because alternative hypotheses would not be activated or would have markedly reduced levels of activation once the recognition point of the multisyllabic prime had passed. Alternatively, we predicted that a delayed commitment strategy would

allow priming for both types of primes, because a final commitment to the identity of the stimulus would occur after the offset of the stimulus word.

Method

Subjects. Fifty undergraduates attending the State University of New York at Buffalo participated for course credit. All subjects reported no speech or hearing disorder at the time of participation.

Materials. The materials in this experiment consisted of auditorily presented primes and visually presented targets. The primes were the multisyllabic tokens used in Experiment 1.

Targets for the 54 sets of primes were divided into three conditions, corresponding to their relatedness to the monosyllabic prime: related, unrelated, or nonword. Related targets were associated with the alternative meaning of the second morpheme in the multisyllabic prime. Thus, for the multisyllabic prime /ɹækɛtbɔl/, the morpheme /bɔl/ was paired with the target DANCE, which is related *only* to the second morpheme. As a result, in the multisyllabic condition, /ɹækɛtbɔl/ preceded DANCE, whereas in the monosyllabic condition, /bɔl/ preceded DANCE. Unrelated targets were selected by randomly assigning the related target words to other unrelated multisyllabic words. For example, DANCE from the previous example was used as a target for /hɛmlæk/ as well as for /læk/ in the unrelated condition. Note that the targets were related only to the final morpheme and *not* to the multisyllabic prime as a whole. That is, /bɔl/ and DANCE are related, whereas /ɹækɛtbɔl/ and DANCE are not. Nonword targets were phonologically permissible letter strings and were, on the average, approximately equal to the target words in length.

Because the words in the monosyllabic condition overlapped with those in the multisyllabic condition (e.g. /ɹækɛtbɔl/ and /bɔl/ overlap), the morpheme conditions were tested between subjects. Thus, one set of groups received multisyllabic primes, whereas the other received monosyllabic primes.

In order to maintain a 50% word/nonword ratio in the lexical decision task, another set of 54 multisyllabic words (and their component word-final morphemes) meeting the same familiarity criteria were used as filler items. Thus, each block of stimuli consisted of 72 trials: 36 nonword trials, 18 related word trials, and 18 unrelated word trials. Three blocks of 72 trials were presented to the subjects; each contained all the auditory primes. The blocks differed only in the specific targets presented after the prime. For example, in one block, related stimuli (e.g., /ɹækɛtbɔl/ and DANCE) were presented. The other two blocks contained the same prime followed by a different target (e.g., /ɹækɛtbɔl/ and KEY in one, and /ɹækɛtbɔl/ and GEEM in the other). The order of presentation of blocks was counterbalanced.

Procedure. Presentation of the stimuli was controlled by a PDP-11/34 minicomputer. The stimuli were presented via a 12-bit digital-to-analog converter at a 10-kHz sampling rate, low-pass filtered at 4.8 kHz, and output binaurally over matched and calibrated TDH-39P headphones.

The subjects were tested in a sound-treated room in individual booths equipped with Microterm 5510 terminals. Before each trial, a row of asterisks presented on the terminal screen for 1 sec signaled the beginning of a new trial. Following offset of the prompt, the auditory prime was output over headphones. Immediately following the offset of the auditory stimulus, the visual target was presented in the center of the terminal. The visual target remained on the screen until all subjects completed their responses.

The subjects entered their lexical decision responses on a computer keyboard by pressing keys labeled “W” for “word” responses and “N” for “nonword.” They were asked to keep one finger from each hand on each button. The responses and response latencies were recorded by a PDP-11/34 computer and were stored for further analysis.

A block of 10 practice trials was presented to the subjects first in order to familiarize them with the task. They were then presented with the 72 experimental trials.

Table 2
Latencies (RT, in Milliseconds),
With Percentages Correct, From Experiment 2

	Syllable Condition			
	Monosyllabic		Multisyllabic	
Relatedness	RT	% Corr.	RT	% Corr.
Related	579	97	562	97
Unrelated	596	96	574	96
Difference	17		12	

Results

Separate analyses were performed on the latency results and accuracy scores. Table 2 shows the mean latency and accuracy results.

A 2 (syllable condition: monosyllabic vs. multisyllabic) \times 2 (relatedness) analysis of variance (ANOVA) was performed on the latency scores. Relatedness was treated as a within-subjects variable, whereas syllable condition was a between-subjects variable. No main effect was found for syllable condition ($F < 1.0$). Targets following monosyllabic primes were not responded to significantly differently than targets following multisyllabic primes. A significant main effect was obtained for relatedness [$F(1,46) = 14.99, p < .05$], indicating that visual lexical decisions were faster for targets following related monosyllabic and multisyllabic primes. Separate post hoc analyses revealed significant effects of priming in both the monosyllabic [$F(1,23) = 8.09, p < .05$] and multisyllabic [$F(1,23) = 7.10, p < .05$] conditions. That is, both /bɔl/ and /æktɔbɔl/ primed DANCE. The syllable \times relatedness interaction was not significant [$F(1,46) < 1.0$].

The items analysis paralleled the analysis by subjects: Targets following monosyllabic primes were not responded to significantly differently than targets following multisyllabic primes [$F(1,106) < 1.0$]. However, targets following related primes were responded to significantly more quickly than targets following unrelated pairs [$F(1,106) = 6.47, p < .05$]. The interaction of syllable condition and relatedness was not significant [$F(1,106) < 1.0$].

The accuracy results are also shown in Table 2. As with the latency scores, a 2 \times 2 ANOVA was performed on the accuracy scores. The main effect for number of morphemes was not significant by subjects [$F(1,46) = 1.0$] or by items [$F(1,106) < 1.0$]. The main effect for relatedness was significant in the subjects analysis [$F(1,46) = 5.30, p < .05$], but not in the items analysis [$F(1,106) < 1.0$]. Lexical decisions for visual targets were more accurate (by 1%) when primed by related words. The syllable \times relatedness interaction was not significant [$F(1,46) < 1.0$].

GENERAL DISCUSSION

In two experiments, we investigated the effects of embedded words on recognition. Specifically, words with multiple morphemes were selected so that their divergence points occurred during the second morpheme. In Experiment 1, we examined subjects' performance in a

gating task, to ensure that the stimuli could be identified before the offset of the word as a whole. The results indicated that the words could be identified after the onset and before the offset of the second morpheme, which ensured that any activation of alternative lexical hypotheses elicited by the second morpheme would occur after the multisyllabic word had diverged from all other lexical candidates. In Experiment 2, we demonstrated that alternative lexical hypotheses remain active after the divergence point of the multisyllabic word had passed. That is, /hɛmlak/ primed KEY, even though KEY was presented at the end of the multisyllabic prime, well after the divergence point.¹

These results are consistent with the principle of delayed commitment and with models that propose that lexical access is initiated at each strong syllable (Cutler, 1989, 1990; Cutler & Norris, 1988; Grosjean & Gee, 1987; Norris et al., 1995). According to these models, lexical access for the multisyllabic prime, /hɛmlak/, should be initiated for each strong syllable, which would include the final morpheme, /lak/. Our cross-modal priming results demonstrated semantic activation of /lak/ independently of /hɛmlak/. By delaying commitment to lexical hypotheses, /lak/ was activated despite ongoing evidence for the identity of /hɛmlak/. This result stands in contrast to the minimal accretion principle proposed by Swinney (1981), which proposes that lexical hypotheses corresponding to post-onset embedded morphemes (such as /lak/ in /hɛmlak/) are not activated if ongoing information supports already activated hypotheses (e.g., /hɛmlak/). According to this principle, the only embedded word within /hɛmlak/ that should have received activation would have been /hɛm/.

These results, in conjunction with evidence that speeded auditory lexical decision latencies are reduced when one or both syllables of a bisyllabic word correspond to a separate lexical item (Luce & Lyons, in press), provide strong evidence that words embedded within larger words, regardless of their location, are activated in memory (see also Shillcock, 1990). These results may be taken as support for a morphological decomposition model of spoken word recognition, but they do not compel such a conclusion (see Sandra, 1990). Instead of the input's being broken into separate lexical entries, as a decomposition model would suggest, it is possible that all lexical items matching a given stretch of the input (such as a syllable) are activated. By this mechanism, /hɛmlak/, /hɛm/, and /lak/ are activated only because they correspond to substantial portions of the input.

Recent research using the word-spotting paradigm has demonstrated that word-final embedded words are inhibited by their carrier items (Norris et al., 1995; see also Norris, 1994). Although such a finding may seem at first glance inconsistent with our results, inhibition may be a later occurring process that operates on activation of second-syllable embedded lexical items (see Luce & Lyons, in press). Thus, the presence of inhibition may be construed as indirect evidence for the initial activation of embedded lexical items (because there must be some activated entity for an inhibitory process to act upon). In short, the results of

the present study are by no means inconsistent with recent models that emphasize the role of inhibition of sublexical competitors in the recognition process (Norris, 1994).

On the other hand, a strictly left-to-right recognition system depending on strong commitment, such as the cohort model (Marslen-Wilson, 1987, 1989; Marslen-Wilson & Welsh, 1978), may not predict any demonstrable influence of lexically ambiguous input beginning after the onset of a word. Because the cohort model states that lexical hypotheses are only activated at word onset and that these hypotheses receive activation only until the recognition point, activation levels of embedded words beginning after word onset should be markedly attenuated after the recognition point. In the present study, semantic activation from /lak/ in /hɛmlak/ primed the response to KEY even after /hɛmlak/ should have been identified and all other contenders excluded from consideration. Although later versions of the cohort model (see Marslen-Wilson, 1987) may allow for residual activation of alternatives following the recognition point, the finding that the mono- and multisyllabic items produced *equivalent* amounts of priming suggests that the cohort model at least underestimates the degree of activation of alternatives after the recognition point. (See also Marslen-Wilson, 1989, note 3, where it is stated that alternatives may be actively inhibited—a claim that is in apparent contrast to the present findings.)

These results also challenge some theories of the nature of semantic effects during word recognition. Zwitserlood (1989), for example, proposed that context prior to a given stimulus word should bias selection of potential word candidates, thus limiting the range of alternatives to the stimulus word. The present study suggests that such a strategy, if it occurs, is limited in scope, because biasing context from the initial morpheme in the multisyllabic words should rule out (or at least minimize) alternative interpretations of the second morpheme. Thus, during the activation of /hɛmlak/, the context from /hɛm/ should rule out any semantic activation's priming KEY.

The present results suggest a broader role for context in the process of spoken word recognition than simple predictive decision-space reduction. Rather, the activation of subword morphemes strongly suggests that semantic processing becomes a factor before a final commitment to lexical identity is made. The notion of a "magic moment" at which semantic information becomes available only after the word is recognized is untenable (see Balota, 1990).²

The mechanism that we propose to account for these findings involves a modified, less binding commitment strategy. Rather than assume an all-or-none commitment during the processing of a word (e.g., at the divergence point), we propose that a dominant hypothesis is entertained, but that semantically and phonetically plausible alternatives remain activated for a period of time, allowing for rapid error recovery if the present hypothesis proves to be erroneous. In the case of /hɛmlak/, while /hɛmlak/ becomes the dominant hypothesis during the second morpheme, alternative hypotheses including /lak/ are activated

as well. This allows the listener to select an alternative hypothesis if the initial hypothesis does not fit the incoming signal. Such a scenario mirrors Grosjean's (1985) distinction between strict divergence points and "total acceptance points," the latter being the point where the listener makes an all-or-none commitment to a lexical hypothesis. This strategy also shares key characteristics with the "sausage machine" sentence-parsing model (Frazier & Fodor, 1978), which analyzes sentences by hypothesizing the simplest possible structure (minimal attachment strategy) while entertaining alternative hypotheses.

In summary, the present study provides evidence that the spoken word recognition system entertains multiple candidates even after the stimulus information has surpassed its divergence point. By delaying commitment to a lexical hypothesis, the system maintains activation of alternative candidates in memory that may enable robust error recovery in the face of ambiguous stimulus input.

REFERENCES

- BALOTA, D. A. (1990). The role of meaning in word recognition. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 9-32). Hillsdale, NJ: Erlbaum.
- BARD, E. G., SHILLCOCK, R. C., & ALTMANN, G. T. M. (1988). The recognition of words after their acoustic offsets in spontaneous speech: Effects of subsequent context. *Perception & Psychophysics*, **44**, 395-408.
- CHARLES-LUCE, J., LUCE, P. A., & CLUFF, M. S. (1990). Retroactive influence of syllable neighborhoods. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 173-184). Cambridge, MA: MIT Press.
- CLUFF, M. S., & LUCE, P. (1990). Similarity neighborhoods of spoken two-syllable words: Retroactive effects on multiple activation. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 551-563.
- COLE, R. A., & JAKIMIK, J. (1980). A model of speech perception. In R. A. Cole (Ed.), *Perception and production of fluent speech* (pp. 133-163). Hillsdale, NJ: Erlbaum.
- CONNINE, C., BLASKO, D., & HALL, M. (1991). Effects of subsequent sentence context in auditory word recognition: Temporal and linguistic constraints. *Journal of Memory & Language*, **30**, 234-250.
- CUTLER, A. (1989). Auditory lexical access: Where do we start? In W. D. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 342-356). Cambridge, MA: MIT Press.
- CUTLER, A. (1990). Exploiting prosodic possibilities in speech segmentation. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 115-123). Cambridge, MA: MIT Press.
- CUTLER, A., & BUTTERFIELD, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory & Language*, **31**, 218-236.
- CUTLER, A., & NORRIS, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 113-121.
- FRAZIER, L., & FODOR, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, **6**, 291-325.
- GOODMAN, J. C., & HUTTENLOCHER, J. (1988). Do we know how people identify spoken words? *Journal of Memory & Language*, **27**, 684-698.
- GROSJEAN, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, **28**, 267-283.
- GROSJEAN, F. (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception & Psychophysics*, **38**, 299-310.
- GROSJEAN, F., & GEE, J. (1987). Prosodic structure and spoken word recognition. *Cognition*, **25**, 135-155.
- KLATT, D. H. (1989). Review of selected models of speech perception.

- In D. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 169-226). Cambridge, MA: MIT Press, Bradford Books.
- KORIAT, A. (1981). Semantic facilitation in lexical decision as a function of prime-target association. *Memory & Cognition*, **9**, 587-598.
- LUCE, P. A. (1986). A computational analysis of uniqueness points in auditory word recognition. *Perception & Psychophysics*, **39**, 155-158.
- LUCE, P. A. (1991, July). *Models of spoken word recognition and the temporal processing window*. Paper presented at Psychophysics of Speech Perception II, Utrecht.
- LUCE, P. A., & CARRELL, T. (1981). *Creating and editing waveforms using WAVES* (Research on Speech Perception, Progress Report No. 7). Bloomington: Indiana University, Department of Psychology, Speech Research Laboratory.
- LUCE, P. A., & LYONS, E. A. (in press). Processing lexically embedded spoken words. *Journal of Experimental Psychology: Human Perception & Performance*.
- LUCE, P. A., PISONI, D., & GOLDINGER, S. (1990). Similarity neighborhoods of spoken words. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 122-147). Cambridge, MA: MIT Press.
- MARR, D. (1982). *Vision*. San Francisco: W. H. Freeman.
- MARSLLEN-WILSON, W. D. (1987). Functional parallelism in spoken word recognition. *Cognition*, **25**, 71-102.
- MARSLLEN-WILSON, W. D. (1989). Access and integration: Projecting sound onto meaning. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 3-24). Cambridge, MA: MIT Press, Bradford Books.
- MARSLLEN-WILSON, W. D., & WELSH, A. (1978). Processing interactions during word-recognition in continuous speech. *Cognitive Psychology*, **10**, 29-63.
- Merriam-Webster's Pocket Dictionary* (1964). Los Angeles: Library Reproduction Service.
- NORRIS, D. (1994). SHORTLIST: A hybrid connectionist model of continuous speech recognition. *Cognition*, **52**, 189-234.
- NORRIS, D., MCQUEEN, J., & CUTLER, A. (1995). Competition and segmentation in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 1209-1228.
- NUSBAUM, H. C., PISONI, D. B., & DAVIS, C. K. (1984). *Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words* (Research on Speech Perception, Progress Report No. 10). Bloomington: Indiana University, Department of Psychology, Speech Research Laboratory.
- PETERSON, R., & SIMPSON, G. (1989). Effect of backward priming on word recognition in single-word and sentence contexts. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 1020-1032.
- PRATHER, P., & SWINNEY, D. (1977, August 26). *Some effects of syntactic context upon lexical access*. Paper presented at the meeting of the American Psychological Association, San Francisco.
- SANDRA, D. (1990). On the representation and processing of compound words: Automatic access to constituent morphemes does not occur. *Quarterly Journal of Experimental Psychology*, **42A**, 529-567.
- SHILLCOCK, R. C. (1990). Lexical hypotheses in continuous speech. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 24-49). Cambridge, MA: MIT Press.
- SWINNEY, D. (1981). Lexical processing during sentence comprehension: Effects of higher order constraints and implications for representation. In T. Myers, J. Laver, & J. Anderson (Eds.), *The cognitive representation of speech* (pp. 263-272). Amsterdam: North-Holland.
- ZWITZERLOOD, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, **32**, 25-64.

NOTES

1. One possible interpretation of these results is that the visual probe itself causes activation of the unrelated meaning of the final morpheme via backward priming (Koriat, 1981). If this were the case, we would expect stronger facilitation effects from the monosyllabic primes than from the multisyllabic primes, because, if the backward priming account is correct, the monosyllabic stimuli should prime the visual targets both directly and by way of backward priming, whereas the multisyllabic primes should prime the targets only through backward priming. However, the magnitudes of facilitation for the two conditions were virtually identical, thus arguing against an explanation based on backward priming. Moreover, evidence that constraining context (in this case, a sentence) eliminates backward priming effects (Peterson & Simpson, 1989) diminishes the plausibility of this interpretation.

2. This study also responds to Zwitserlood's (1989) contention that previous cross-modal priming studies did not bear on the activation of different lexical entries because of their reliance on using multiple meanings of a single lexical token. Here, the lexical ambiguity of a subword unit was examined, which in turn activated different lexical entries (i.e., those that corresponded to the subword unit).

(Manuscript received May 3, 1993;
revision accepted for publication May 4, 1997.)