

The recognition of words after their acoustic offsets in spontaneous speech: Effects of subsequent context

E. G. BARD, R. C. SHILLCOCK, and G. T. M. ALTMANN
University of Edinburgh, Edinburgh, Scotland

Three experiments are presented that investigated the recognition of words *after* their acoustic offsets in conversational speech. Utterances randomly selected from the speech of 24 individuals (total $N=288$) were gated in one-word increments and heard by 12 listeners each. Of the successful recognitions, 21% occurred after the acoustic offset of the word in question and in the presence of subsequent context. The majority of late recognitions implicate subsequent context in the recognition process. Late recognitions were distributed nonrandomly with respect to the characteristics of the stimulus word tokens. Control experiments demonstrated that late recognitions were not artifacts of eliminating discourse context, of imposing artificial word boundaries, or of repeating words within successive gated presentations. The effects could be replicated only if subsequent context was available. The implications are discussed for models of word recognition in continuous speech.

A decade and a half of research on the recognition of words in fluent speech has demonstrated that under the right conditions listeners can apply acoustic and higher-level information to this task more or less instantaneously (Grosjean, 1980; Marslen-Wilson, 1973; Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Welsh, 1978). Models of word recognition designed around these findings have consequently emphasized the optimal functioning of word recognition processes. The following descriptions of the process of listening to speech appear to describe a device that utilizes preceding context only and that starts a new cycle at the onset of each new word:

Utterances are indeed understood as they are heard; ... the listener constructs a syntactic and semantic interpretation of the input word-by-word as he hears it, and ... he actively uses this information to guide his processing of the subsequent words in the string. (Marslen-Wilson, Tyler, & Seidenberg, 1978, p. 240)

Speech is processed sequentially, word by word ... the words in an utterance are recognized one after another ... listeners know where words begin and end by recognizing them in order. (Cole & Jakimik, 1980, pp. 133-134)

We suggest that in perception of conversational speech under good listening conditions, recognition of one word is

complete before recognition of the following word begins. (Cole & Jakimik, 1980, p. 149)

Increasingly, however, it has been suggested that the human speech recognition mechanism operates in a more varied fashion than these statements imply. Experimental evidence of less than optimal processing dates from the experiments of Pickett and Pollack (1963; Pollack & Pickett, 1963, 1964), who used a gating technique in which listeners were asked to identify all the words contained in successively longer substrings of recorded utterances. Pickett and Pollack found that intelligibility in all word positions in an utterance increased with the length of the substring presented: not only did later words in the string become easier to recognize, but earlier words were often belatedly recognized as subsequent words were added. More recently, Grosjean (1985) presented semantically uninformative sentence onsets followed by 50-msec-gated short, rare monosyllabic nouns and brief sentence completions. Recognition was often achieved some time after the acoustic offset of the nouns, and confidence in those recognitions peaked only within some subsequent word. Furthermore, words often were not recognized in the order in which they had been spoken, so that when some words were being recognized, their subsequent contexts were already available.

More recent models accordingly have offered a number of alternatives to a strictly left-to-right and word-by-word device. Marslen-Wilson (1987) proposed that competing word hypotheses may be actively pursued over a stretch of input corresponding to more than one word, so that recognition may occur well after word offset. Grosjean and Gee (1987) and Cutler and Norris (1988) suggested that one of the mechanisms involved in word recognition waits for syllables of particular clarity and is not set off at every actual or potential word onset. Like

This study was supported by Grant GR/C78377 from the Science and Engineering Research Council (U.K.) to H. S. Thompson and E. G. Bard. Speech materials were drawn from those prepared under the auspices of grants by the Education and Social Science Research Council (U.K.) to J. Laver and E. G. Bard (HR 6130) and to G. Brown (HR 3601). The authors wish to thank Gill Brown for the use of these materials and Henry Thompson, John Laver, and the members of the Centre for Cognitive Science Workshop on Speech Processing for discussion and encouragement. Reprint requests should be sent to E. G. Bard, Centre for Speech Technology Research, University of Edinburgh, 80 South Bridge, Edinburgh EH1 1HN, Scotland.

Marslen-Wilson's, these approaches allow for the recognition of words after their acoustic offsets, although the models differ in the extent to which they specify the cases in which postoffset recognition is most likely and the mechanisms that are involved.

At present, however, there are very few facts about late recognition to which theories such as these might conform. Although Grosjean's (1985) experiment established that late recognition can occur, it did not establish the generality of the phenomenon, because it dealt exclusively with rare lexical items preceded by contexts offering little semantic constraint. The purpose of the present study was to provide a normative account of late recognition. On the basis of a large-scale study of the recognition of words in conversational speech, we sought to determine whether the recognition of words after their acoustic offsets is a marginal phenomenon, such as might be attributed to random error in an essentially efficient left-to-right process, or whether the phenomenon has characteristics that invite further explanation.

The work centers around a number of questions. First, do late recognitions actually occur often enough to be of interest? Should they be extremely rare, they may well be seen as the random error in an otherwise efficient process. Second, is subsequent context implicated in any way in the process of late recognition? Only if it is can an account of the phenomenon be distinguished from one in which the processor works in strict chronological order, but with variable delays for processing. Third, are late recognitions randomly distributed over word tokens, as they would be if a truly random process were involved, or are they associated with particular characteristics of words? Because many characteristics of words have known relationships to the recognition process, the task of relating them to late recognition is a preliminary to modeling the mechanisms involved.

These three questions are pursued in Experiment 1 by means of a word-level gating experiment on a sample of utterances randomly selected from audio recordings of conversations. Both the materials and the technique deserve comment.

The use of spontaneous, conversational speech in this study was motivated by the assumption that such speech is what listeners most often process. Conversational speech has well-documented structural and phonological characteristics that are minimized in read texts (Dalby, 1984; Shockey, 1973) and that are almost certainly not optimal for a strictly left-to-right processor because they obscure the intended relationship between the sound produced and its correct phonological representation. If listeners regularly process less than optimal input, the additional difficulty of preparing conversational materials is worthwhile.

The gating technique (Cotton & Grosjean, 1984; Grosjean, 1980; Pickett & Pollack, 1963; Pollack & Pickett, 1963, 1964; Tyler & Wessels, 1983, 1985) is a well-

established method for examining the point at which word recognition takes place. In the present study, the issue was how often and how far recognition lags behind the completion of the word's acoustic form rather than when within a word its recognition might occur. Accordingly, stimuli were presented in substrings, starting with the first word alone, and with each successive substring one word longer than its predecessor. The early recognition data suggest that a word should often be recognized on its first presentation when prior context and full acoustic form are both available. If a word is not recognized until one or more *following* words have been presented, then a late recognition is recorded.

A fourth question arises from the technique itself. Are late recognitions in the present study due to the effects of subsequent context or to various changes made to the original conversational stimuli when they are gated? Experiments 2 and 3 provided the necessary controls.

EXPERIMENT 1

Method

Subjects. Subjects were 48 members of the University of Edinburgh undergraduate and postgraduate community. All reported having normal hearing. All had been resident in Edinburgh for at least 2 years and could be expected to have encountered the range of accents represented in the corpus used in the experiment.

Materials. The utterances used in the experiment were assembled from two taped corpora of conversational English (Bard & Anderson, 1983; Brown, Currie, & Kenworthy, 1980). Both corpora consisted of lengthy conversational dialogues between naive subjects and an experimenter and were recorded on professional audio equipment in a quiet situation. We selected 12 speakers, 6 male and 6 female, from each corpus, to include a range of the accents encountered in the Edinburgh area. For each speaker, 12 utterances were randomly selected, making a total of 288 utterances.

The decision that a particular string of words constituted a single utterance was made on grammatical and prosodic grounds. Utterances that were indistinct as a result of background noise or overlapping with the conversational turn of the second speaker were dismissed. For each utterance actually used in the experiment, an average of 0.6 utterances were rejected because of background noise, and another 0.95 because of interruptions by the second speaker. Thus the materials represent that part of conversational speech in which perceptual processes do not have to overcome extraneous noise.

Of the 288 utterances, 67 contained disfluencies such as filled pauses (37 instances, not necessarily in different utterances), repetitions of complete words or phrases (13 instances), syntactic/semantic reformulations (22 instances), and false starts to words (17 instances). For purposes of segmentation, filled pauses, false starts, and contracted forms such as *I'll* or *didn't* (112 instances) were all counted as single words. The utterances varied considerably in length, with a range from 1 to 22 words and a mean of 6.7 words.

The utterances were digitized at 10000-Hz through a low-pass (5000-Hz) filter. Gates were set at word boundaries as determined auditorily and visually from the time-amplitude waveform in such a way that each gate finished immediately before the onset of material recognizably belonging to the next word. Gated output automatically generated from the digitized utterances was directed to a tape

recorder. The output began with the first word of the utterance and was incremented by one word per presentation until the whole utterance was included. The time between word presentations began at 3.5 sec and was incremented by 400 msec after each gate up to a maximum of 5.5 sec, to allow subjects more time to make lengthier responses.

Four audiotapes were produced, with all 24 speakers represented on each tape in a different random order. Three different, randomly chosen utterances from each speaker appeared in a block on each of the four tapes; however, repetition of content words by the same speaker in different utterances was avoided as far as possible on any one tape. Each block of three trials was preceded by an ungated adaptation utterance from that particular speaker. The same adaptation utterance for each speaker was played on all four tapes. Each tape began with a practice block, starting with an adaptation utterance from a speaker not involved in the main part of the experiment, followed by three gated utterances from the same individual. These trials were followed by the 72 trials of the experiment proper. In each case, the adaptation utterance was preceded by two tones and the beginning of a gated item by three.

Procedure. The subjects were provided with a complete list of the adaptation utterances and were told that the utterances would be gated in increments of one word, with the exceptions outlined above. They were instructed to listen carefully to the gated presentations of the utterances and at the end of each gate to write down all the words it contained, using prearranged symbols to indicate that the previous transcription stood or that they were unable to produce any guess for a word. The subjects were, however, urged to try to identify each word and were cautioned against tampering with their written answers to earlier presentations rather than adopting a new transcription for the new presentation. The subjects listened to the three practice trials and were given a further chance to ask questions. Responses to each utterance were recorded on a separate sheet of paper containing a box for each word of each presentation, but not indicating how many words any utterance would contain.

The subjects were tested individually or in pairs, in a recording studio. The tape was played over a loudspeaker situated approximately 2 m from the subjects. The experiment was run in two sessions of approximately 45 min, separated by a short rest break.

Results

The 288 utterances in the materials contained a total of 1,930 word tokens. Of these, 288 tokens were utterance-final. Because these were presented only once, there was no experimental trial on which they could have been recognized late. Fifty-five additional tokens were fragments or filled pauses, and it is unclear whether normal word recognition processes should be expected to operate on these items. Consequently, 343 items in all were eliminated from the analyses reported below.

Each of the remaining 1,587 word tokens was presented on at least two trials to each of 12 subjects. This gave 19,044 recognition outcomes (i.e., individual listeners' attempts to recognize individual word tokens over multiple trials) through which the time course of recognition could be traced. Each recognition outcome was classified as a successful recognition on the first presentation with prior context only (*immediate*), a successful recognition in the presence of subsequent context (*late*), or a failure to recognize (*missed*). In the analyses presented below, the original word and its homophones, with or without the correct inflection, were scored as correct recognitions.

Other marking schemes, including the one accepting only perfect orthographic transcription, give substantially the same results.

To justify examining results as a whole without attention to by-subject or by-materials means, it was necessary to demonstrate that the data were homogeneous. One-way analyses of variance (ANOVAs) for experimental groups (4) were performed both by subjects and by materials on the percentages of responses in each of the three recognition categories. Except for a tendency of one group to produce fewer missed recognitions than the others [$F_1(3,44) = 3.14, p = .03; F_2(3,251) = 1.39, n.s.$], there were no differences among groups on any measure. Furthermore, late recognitions, the category of most interest, occurred in all but 12 of the 259 multiword stimuli and in the responses of all subjects. Data were accordingly summed over stimuli and subjects for the analyses of the incidence and patterns of late recognition set out below. For multiple regression analyses that explored the relationship of late recognition to a number of stimulus characteristics, outcome probabilities were calculated by word token.

Frequency of late recognitions. The first question to be answered by this experiment was whether late recognition is only a marginal process. Figure 1, which displays the distribution of outcomes by recognition point, shows that it is not. Although the majority of recognition outcomes (69%) yielded success on the word's first presentation with prior context only, 19% of all outcomes¹ and 21% of all successful outcomes were late recognitions. The mean percentage of late recognitions was significantly different from zero [$F_1(1,44) = 547.45, p < .0001; F_2(1,59) = 376.38, p < .0001; \min F'(1,102) = 223.04, p < .0001$].

These late recognitions were not merely an artifact of the interruption of word-final coarticulation. Approximately 35% of them were identified not at the presentation of the next word, but later still. The mean number of subsequent words needed for late identification was closer to two than to one ($M = 1.69, SD = 1.32$).

Patterns of late recognition. The second question was whether subsequent context appeared to be in any way implicated in the process of late recognition. The answer is derived from a subclassification of late recognitions on the basis of the subsequent words correctly identified at the time of recognition.² Table 1 gives examples of the resulting categories, and Figure 2 displays the distribution of late recognitions among them. The categories are discussed below, in ascending order of evidence for a right-to-left flow of information. A conservative assumption has been adopted in forming these categories, namely, that right context can bear on recognition only if the words in that right context are correctly identified. In reality, even if words following an item are not correctly identified, they may contribute information to the recognition process insofar as the listener correctly hypothesizes their syntactic categories or correctly perceives any of the acoustic information that they contain. Thus the follow-

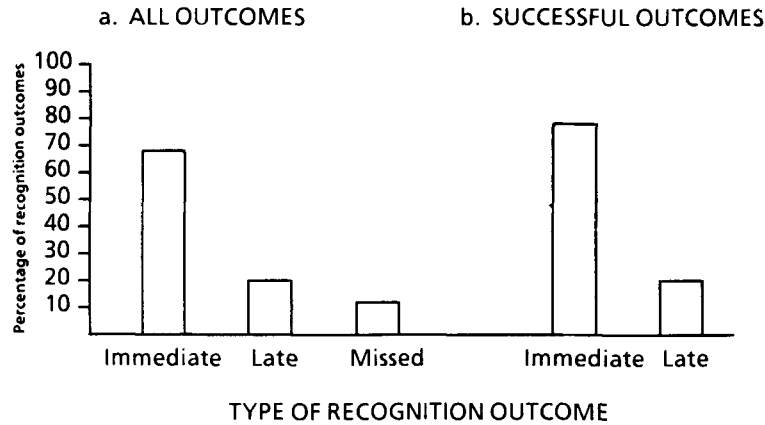


Figure 1. Distribution of recognition outcomes ($n=19,044$) in Experiment 1 by recognition point. Responses for fragments, filled pauses, and utterance-final words are omitted.

ing results are underestimates of the potential extent of right-to-left information flow.

Category 1: Sequential recognitions. Recognitions in this category consist of a single word recognized on a presentation later than its first (i.e., in the presence of subsequent context), but before any of the subsequent words have been correctly identified. Table 1 gives an example of a gated stimulus and a sequential late recognition: the first instance of *the* in the utterance is recognized on its third presentation while the responses *reader* (for *leader*) and *on* (for *of*) are still incorrect.

This pattern provides no direct evidence that any aspect of subsequent context affects the recognition of *the*. When the word is eventually recognized, the listener has not identified any subsequent words that might augment the usable context. Consequently, this sort of late recognition could be interpreted as the delayed arrival of a percept in a temporally sequential system.

Sequential recognitions, however, account for only 9% of all late recognitions.

Category 2: Simultaneous recognitions. In this category of late recognitions, a word is first recognized on the same

Table 1
Temporal Patterns of Late Recognition Found in Experiment 1

Stimulus Sequence		
	HE'S	
	HE'S THE	
	HE'S THE LEADER	
	HE'S THE LEADER OF	
	HE'S THE LEADER OF THE	
	HE'S THE LEADER OF THE LABOUR	
	HE'S THE LEADER OF THE LABOUR PARTY	
Category	Response Sequence	
Sequential	HE'S	
	HE'S <i>a</i>	
	HE'S <i>a</i> reader	
	HE'S <i>THE</i> reader on	
	Complete	Partial
Simultaneous	HE'S	HE'S
	HE'S <i>a</i>	HE'S <i>a</i>
	HE'S <i>a</i> reader	HE'S <i>a</i> reader
	HE'S <i>THE</i> LEADER OF	HE'S <i>THE</i> reader OF
Nonsequential	HE'S	
	HE'S <i>a</i>	
	HE'S <i>a</i> LEADER	
	HE'S <i>a</i> LEADER OF	
	HE'S <i>THE</i> LEADER OF THE	

Note—Uppercase words indicate correct responses. Italics indicate the word whose recognition outcome exemplifies the named pattern.

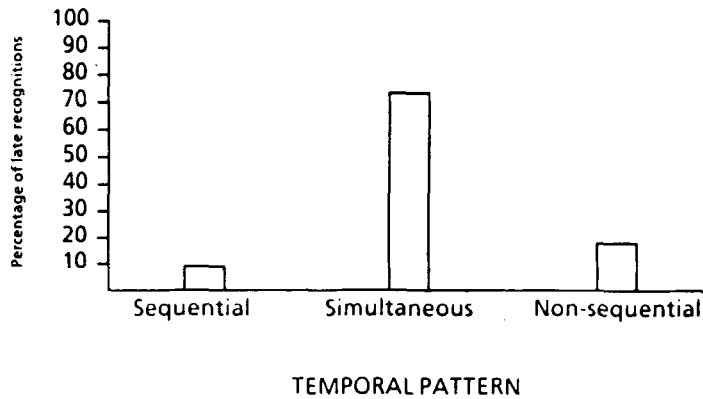


Figure 2. Distribution of late recognitions ($n=3,556$) in Experiment 1 by temporal pattern. Responses for fragments and filled pauses are omitted.

late presentation at which words between it and the end of the gate are also first recognized. A simultaneous recognition may be *complete*, when all the subsequent words are recognized at once, or *partial*, when only some of them are identified. Both types are exemplified in Table 1.

Although the failure of left-to-right processes may account for the formation of a queue of unrecognized words at the end of a stimulus, the resolution of this queue into a string of late recognitions is most simply explained in terms of information becoming available at the right end of the string. The event that we know has occurred at the point at which recognition is achieved is the presentation of a new word at the end of the stimulus. In comparison, purely left-to-right explanations of strings of simultaneous recognitions appear to be more complex, since they ignore this known event in favor of unspecified covert processes involving only the left context of the unrecognized words.

Complete simultaneous recognitions represent 70% and partial simultaneous represent 3% of all late recognitions.

Category 3: Nonsequential recognitions. This category, which represents the strongest evidence for the effects of subsequent context, includes instances where the order in which words are recognized differs from the order in which they were said. In the example in Table 1, the first *the* is recognized only on its fifth presentation, after *leader*, *of*, and the second occurrence of *the* have been correctly perceived. Because these three words have already been recognized, the listener has some basis for generating higher-level information that may be applied to the disambiguation of an earlier word form.

This category accounted for 18% of all late recognitions.

Figure 2, which summarizes these results, shows that the bulk of late recognitions fall in those categories (simultaneous and nonsequential recognitions) which are difficult to account for in terms of left context alone.

Factors associated with late recognition. The third point of interest was whether late recognitions occurred randomly with respect to the characteristics of the words

being recognized, as might be expected if the process of late recognition represented random error in the operation of the speech recognition device. In fact, whether and when a word was recognized depended to a significant degree on the characteristics of that word.

Multiple regression analyses were used to explore a number of potentially relevant characteristics of stimulus words. For these analyses each word was first associated with a number of independent variables relating to it as a lexical item or word type. For each word, *syllable length*, as determined from dictionary forms, classification as a *function-or-content word*, and *frequency of occurrence* (Francis & Kučera, 1982) were recorded. Variables associated with each word as a word token in a particular utterance were also recorded. For each stimulus word, *millisecond length* was calculated from the duration of the gate introducing the word. To reflect the position of each word token within its utterance, the number of *preceding words* and the number of *following words* within the utterance were used. Separate multiple regressions were run by the set-hierarchical method of Cohen and Cohen (1983, chap. 4, pp. 137ff.) to determine the utility of type and token variables as predictors of each of the three classes of recognition outcome: immediate, late, and missed.

In each case, independent variables associated with the words as *tokens* added significantly to the predictive power of equations containing only variables associated with the words as *types*. Table 2 lists standardized correlation coefficients from equations containing both type and token variables.

As this table shows, several of the variables that were examined make independent contributions to the determination of a word's recognition point. Longer words (as measured in milliseconds), content words (nouns, verbs, adjectives, etc.), and words further from the beginning of an utterance were more likely to be recognized on their first presentation. Function words (conjunctions, prepositions, articles, auxiliary verbs, etc.), shorter words (measured in milliseconds), and words closer to the be-

Table 2
Multiple Regression Equations for Each Category of Recognition Outcome in Experiment 1,
Including All Words, Content Words, and Function Words

Dependent Variable	Independent Variables						Intercept	Multiple R^2	F	df
	Functor/ Contentive	Syllable Length	Word Frequency	Preceding Words	Following Words	Length (in msec)				
All Words										
% Immediate	.10‡	-.00	.03	.24§	-.02	.31§	.54	.20	64.10	6,1567
% Late	-.07*	-.02	.00	-.14‡	.08†	-.30§	.26	.16	49.07	6,1567
% Missed	-.08†	.03	-.05	-.22§	-.06†	-.18§	.20	.09	25.38	6,1567
Content Words										
% Immediate		.01	-.11†	.23§	.04	.29§	.52	.17	24.93	5,628
% Late		-.07	.06	-.16§	-.02	-.26§	.29	.13	19.27	5,628
% Missed		.05	.11	-.19§	-.05	-.19§	.19	.08	11.15	5,628
Function Words										
% Immediate		.00	.05	.24§	-.05	.29§	.38	.15	33.42	5,935
% Late		-.04	-.01	-.12‡	.12‡	-.28§	.36	.12	25.07	5,935
% Missed		.04	-.06	-.24§	-.06	-.14‡	.27	.07	14.84	5,935

Note—All F values have $p < .0001$. * $p < .025$. † $p < .01$. ‡ $p < .001$. § $p < .0001$.

ginning of the utterance were more likely to be recognized late or not at all. Those that were closer to the end of the utterance, where the opportunity for late recognition was limited, were more likely not to be recognized at all; those further from the end of the utterance (i.e., in longer utterances) were more likely to be recognized late. Thus length, distance from utterance onset, and form class affect whether or not a word is recognized immediately. If not, whether it is recognized at all depends on the number of additional trials available.

These equations show that the function/content word variable makes a predictive contribution independent of word length and position, but they cannot reveal whether function and content words were affected by the same variables to the same degree. Two further analyses tested for qualitative differences between these two classes of words.

First, separate multiple regression equations were calculated for functors and contentives. There was a significant difference within each pair of equations [for immediate recognitions, $F(6,1563) = 5.95$, $p < .00001$; for late recognitions, $F(6,1563) = 4.86$, $p < .00006$; for missed recognitions, $F(6,1563) = 3.27$, $p < .003$]. As Table 2 shows, however, these differences must have been due largely to a main effect of word class, since this is indicated by the differences between the intercepts of the equations. Interaction between word class and other variables takes the form of differences in the standardized regression coefficients for these other variables within each pair of equations. Yet the differences here are slight. For the strongest predictor variables, millisecond length and number of preceding words, the regression coefficients were virtually indistinguishable. Thus, although the 23% of function words and the 12% of content words that were recognized late were subject to similar effects of word length and position in utterance, the major difference between the word classes is that function words seemed to

be more likely to yield late recognitions simply by virtue of being function words. Similar arguments could be made for the 62% of functors and 81% of contentives recognized immediately and for the 15% of functors and the 7% of contentives that were missed.

The second analysis compared the distribution of function and content words among the three patterns of late recognitions identified above (Figure 2). The distributions did not differ significantly [$\chi^2(2) = 2.32$, $n = 3,556$, n.s.], with 8% of the late-recognized function words and 9% of the late-recognized contentives classed as sequential recognitions, 73% of late function words and 74% of late content words as simultaneous recognitions, and 19% of late function words and 17% of late content words as nonsequential recognitions. As far as these variables are capable of showing, the recognition outcome of all words, functors or contentives, depends on similar processes and follows similar patterns.

Discussion

Experiment 1 was intended to determine whether the late recognition of words in running speech is anything other than an error in the operation of a mechanism that normally decodes word shapes in their chronological order and with the aid of prior context alone. The results suggest that late recognitions are not a marginal phenomenon. First, late recognitions are common enough to warrant theoretical attention. In a sample of 19,044 attempted word recognitions, just over 1 in 5 successful recognitions occurred only when some part of a word's subsequent context was available. Second, late recognitions are not readily accounted for in terms of processes that use only prior context and move in chronological order one word at a time. In fact, only 9% of late recognitions could be explained in this way. Like Grosjean's (1985) findings for monosyllabic nouns, the vast majority of late recognitions conformed to patterns that suggest that the

human word recognition device has access to subsequent as well as prior context. Third, late recognitions are not randomly distributed over spoken words. Significant proportions of the variance in the frequency of each of the three possible types of recognition outcome were accounted for by characteristics of the stimuli. Moreover, these characteristics are known to have psycholinguistic consequences.

The fact that longer words of all categories were more likely to be recognized with left context only and shorter words to be late or missed is readily related to the long-established positive correlation between word length and word intelligibility (Rosenzweig & Postman, 1958). The finding that words close to the beginning of the utterance are more difficult to recognize on their first presentation and more likely to be recognized later or not at all fits quite comfortably with position effects in the word monitoring results of Marslen-Wilson and Tyler (1980) and with the processing difficulties encountered by Woods et al. (1976) within automatic speech recognition. The greater propensity of contentives to be recognized with left context only and of functionals to be recognized late or not at all conforms to what might be predicted from their behavior on other tasks (Bradley, 1978).

The nature of a mechanism capable of creating these effects will be discussed later. First, it is necessary to show that the results of Experiment 1 were not an artifact of the peculiarities of the gating technique as used here. Experiment 2 dealt with the effects of using utterance-length stimuli removed from their natural discourse context. Experiment 3 controlled for the repetition of stimuli within the standard gating technique and for the effects of the artificial word segmentation used in word-level gating.³

EXPERIMENT 2

In Experiment 1, the only prior context available to listeners came from within the gated utterance itself. The rest of the discourse in which the stimulus utterances had been produced was not presented. But if left-to-right processes had been able to operate over the true left context of the stimuli—that is, the full prior discourse—they might have sufficed for first-trial recognition in many more cases.

This explanation seems plausible in view of the finding (Bard, Anderson, & Laver, 1982) that speakers tend to give less careful articulation to content words that have occurred recently in a discourse than to the same words on their first occurrence. For instance, if the recently preceding discourse had mentioned *cars*, subsequent tokens of *car* or *cars* might be less carefully articulated to the point that they could not be deciphered without reference to their extended discourse context. Furthermore, access to the full prior discourse would give listeners considerably more exposure to idiosyncrasies of accent and voice quality than was possible with the adaptation utterances in Experiment 1.

In Experiment 2, subjects were provided with the preceding discourse context, represented by the full taped dialogue up to and including utterances that had elicited substantial numbers of late recognitions in Experiment 1. The utterances themselves were represented by the same gated stimuli as in the earlier experiment. If the late recognitions of Experiment 1 were artifacts of the impoverished discourse environment in which stimuli were presented, then they should be replaced here by immediate recognitions. If dependence on subsequent information is a characteristic of word recognition in natural conditions, however, extending the prior context should have little effect on recognition patterns.

Method

Subjects. The subjects were 48 students at the University of Edinburgh who had not taken part in Experiment 1.

Materials. For 2 male and 2 female speakers, the six utterances producing the greatest number of late recognitions in Experiment 1 were selected. An audiotape was produced for each speaker, including a copy of the original dialogue with the six critical utterances replaced by the gated versions of the same utterances that had been used in Experiment 1. A series of three tones was added to the tape before each gated utterance. Two tones preceded the resumption of the dialogue.

Procedure. The subjects listened to the tapes over headphones. Each of the four tapes was heard by 12 subjects. In each case, the subjects heard all of the taped dialogue containing the gated utterances, up to the end of the last gated item. They were instructed to listen carefully to all of the conversation and to write down the words contained in each gate of the gated utterances. The rest of the instructions were as before.

Results

Although addition of prior discourse context in Experiment 2 reduced the rate of late recognitions and the average delay to recognition, it did not eradicate them. As Figure 3 shows, late recognitions fell from Experiment 1 levels and immediate recognitions rose, whereas the number of missed items remained constant. Although the distributions differed across the two experiments [$\chi^2(2) = 14.51, p < .05, n = 5,496$ for all responses; $\chi^2(1) = 14.33, p < .05, n = 5,037$ for successful recognitions], late recognitions still represented 19% of all outcomes and 21% of all successful recognitions. The mean number of presentations to recognition for words actually recognized fell from 2.69 in Experiment 1 to 1.90 [$t(249) = 2.24, p < .025$].

In other respects, the additional left context did not appear to change the nature of the late recognition process. As Figure 4 shows, the proportions of late recognitions in these materials that could be classed as sequential, simultaneous, and nonsequential were essentially the same as in Experiment 1 [$\chi^2(2) = 0.74, n.s., n = 1,161$]. Also, as Table 3 shows, multiple regression equations for Experiment 2 stimuli failed to differ from equations for the same stimuli heard within Experiment 1 [all $F_s(8,438) < 1$]: the same factors determined recognition outcome. Interestingly, even the detrimental effect of proximity to

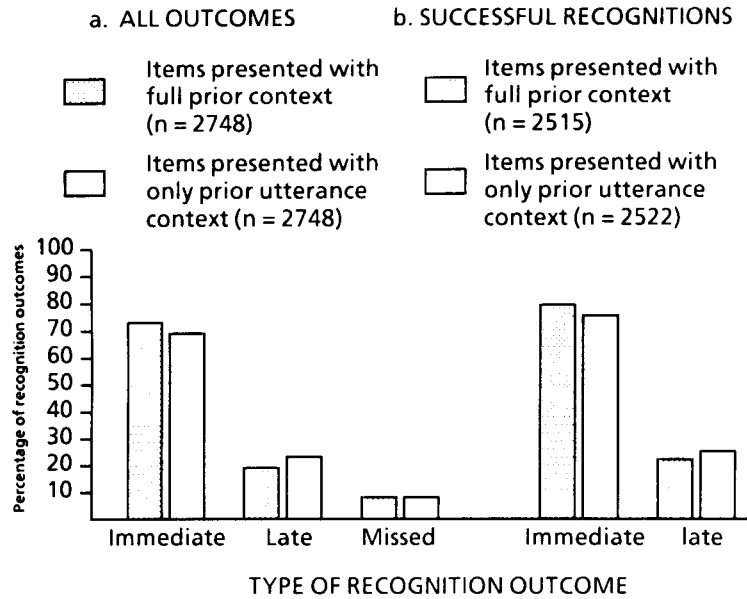


Figure 3. Distribution of recognition outcomes by recognition point for materials presented with full prior discourse context (Experiment 2) or with prior within-utterance context only (Experiment 1).

utterance onset reappeared in this experiment, despite the fact that utterance onset was no longer the onset of left context.

Thus the absence of prior discourse context in Experiment 1 was not responsible for the bulk of the late recognitions recorded there, for the patterns they assumed, or for the stimulus characteristics associated with them. It does seem, however, that the completeness of prior context determines when the late recognition is finally achieved: with more prior context, less subsequent context is needed. The finding that the same words were recognized with different delays in the two experiments is consistent with the view that the recognition device

needs some disambiguating information, but not some particular piece of disambiguating information in each case of late recognition. Such a mechanism cannot depend on a particular word or syllable to enable it to discharge pending recognitions, but will address the task of recognition in a way determined by the combination of prior context and word shape available at the time.

EXPERIMENT 3

Experiment 1 shared two characteristics of conventional gating tasks that might be suspected of contributing to the number of late recognitions that listeners produced. One

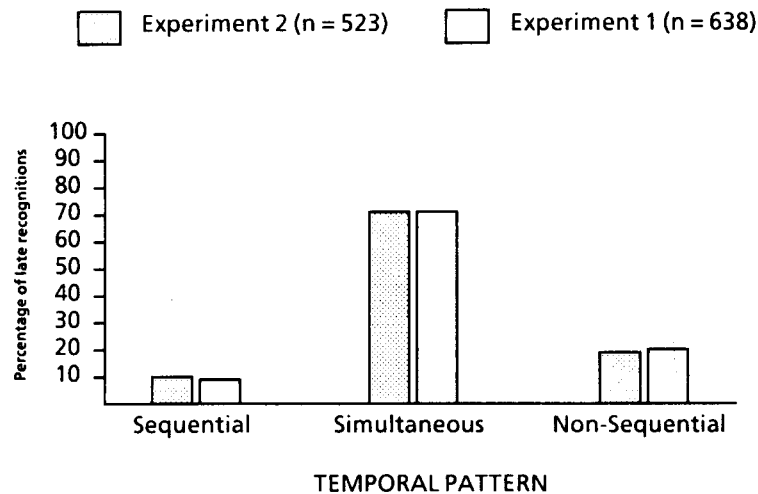


Figure 4. Distribution by temporal pattern of late recognitions for materials presented with full prior discourse context (Experiment 2) or with prior within-utterance context only (Experiment 1).

Table 3
Multiple Regression Equations for Each Category of Recognition Outcome for Words Presented With (Experiment 2) and Without (Experiment 1) Full Prior Discourse Context

Dependent Variable	Independent Variables						Intercept	Multiple R ²	F	df
	Functor/Contentive	Syllable Length	Word Frequency	Preceding Words	Following Words	Length (in msec)				
Full Discourse										
% Immediate	.25†	-.11	.10	.34§	.04	.22†	.33	.20	9.31§	6,220
% Late	-.24†	.02	-.09	-.33§	-.05	-.24†	.53	.23	10.88§	6,220
% Missed	-.13	.16*	-.06	-.16*	.00	-.07	.14	.05	1.90	6,220
Utterance Only										
% Immediate	.19*	-.05	.03	.34§	.00	.26‡	.27	.25	12.39§	6,220
% Late	-.17†	-.02	.02	-.30§	-.05	-.24†	.59	.23	11.12§	6,220
% Missed	-.08	.12	-.09	-.19*	.09	-.13	.14	.09	3.42	6,220

**p* < .05. †*p* < .01. ‡*p* < .001. §*p* < .0001.

characteristic is the artificial segmentation of the utterance imposed by the process of gating. The other is the repetition of stimuli in successive gates. These are examined factorially in Experiment 3.

In Experiment 1, gates ended at word boundaries. The segmentation of utterances into words may supply the listener with information that he/she would otherwise have to derive actively. Though this information might facilitate recognition, the interruptions might just as easily disrupt the timing and continuity of recognition processes. Where the discontinuities in the early part of an utterance have misled the subject into an incorrect interpretation, mistakes may propagate through the utterance until the erroneous interpretation becomes totally untenable. The listener may then abruptly change interpretations and produce numerous late recognitions. Experiment 3 controls for this possibility by presenting one word per utterance under two conditions: in one, the whole utterance is gated, as in Experiment 1, and in the other, the first gate for the utterance contains all the words up to and including that critical word. If segmentation promotes late recognitions, then this *unsegmented* control will produce more immediate and fewer late recognitions than the *segmented* replication of the earlier experiment.

The gating method used in Experiment 1 also leaves open the possibility that late recognitions resulted not from

the processing of additional context with each new gate, but from the additional processing of the original acoustic material which is repeated in all the succeeding gates. Cotton and Grosjean (1984) have shown, however, that responses to individual gates presented to different listeners do not differ substantially from responses to the same gates presented in sequence to a single group of listeners. Experiment 3 provides an additional control for repetition by directly testing the hypothesis that late recognitions are caused wholly or largely by repetition of the stimulus word. Conditions in which an utterance is gated to the end of a particular word and then merely repeated *without subsequent context* are contrasted with those in which the same number of subsequent presentations consist of the original gates *with subsequent context*, as in Experiment 1. If it is repetition and not subsequent context that is responsible for late recognition, both conditions should show the same proportion of late recognitions, because in both the critical word is repeated. If, however, the subsequent context is crucial, then only the *with-subsequent-context* condition should produce many instances of late recognition.

The factorial design of Experiment 3 appears in Table 4. To deal with segmentation, each utterance is presented up to and including its single critical word either segmented, like the examples in the first row of the ta-

Table 4
Stimulus Sequences for Experiment 3

Left Context	Right Context	
	With Subsequent Context	Without Subsequent Context
Segmented	He's	He's
	He's the	He's the
	He's the <i>leader</i>	He's the <i>leader</i>
	He's the <i>leader</i> of	He's the <i>leader</i>
	He's the <i>leader</i> of the	He's the <i>leader</i>
	He's the <i>leader</i> of the Labour	He's the <i>leader</i>
	He's the <i>leader</i> of the Labour Party	He's the <i>leader</i>
Unsegmented	He's the <i>leader</i>	He's the <i>leader</i>
	He's the <i>leader</i> of	He's the <i>leader</i>
	He's the <i>leader</i> of the	He's the <i>leader</i>
	He's the <i>leader</i> of the Labour	He's the <i>leader</i>
	He's the <i>leader</i> of the Labour Party	He's the <i>leader</i>

Note—Only the italicized word is scored.

ble, or unsegmented, like those in the second. Once the gate ending with the critical word has been presented, it is repeated either with subsequent context, as illustrated in the first column of the table, or without subsequent context, as in the second column. If segmentation generates artificially high numbers of late recognitions, then the number of late recognitions will differ across rows. If repetition of stimuli is what allows late recognition to be achieved, then late recognitions will be the same across columns, because both conditions involve repeating the stimulus. If subsequent context is what makes late recognition possible, however, then only the patterns in the first column of the table will produce late recognitions.

Method

Subjects. Subjects were 48 students from the University of Edinburgh: none had taken part in any of the previous experiments. Twelve were run in each condition.

Materials. From the 288 items used in Experiment 1, 60 were selected which contained one word recognized late by at least 3 of the 12 subjects. These 60 reliable instances of late recognition contained 15 from each of the following four categories: (1) *sequential recognitions*; (2) *short complete simultaneous recognitions*, in which two consecutive words were recognized on the first presentation of the second word; (3) *long complete simultaneous recognitions*, in which three or more consecutive words were recognized on the first presentation of the last word in the string; and (4) *nonsequential recognitions*. The *partial simultaneous* category did not provide a sufficient number of reliable instances for inclusion. In 9 of the 15 items in each category, the critical word was utterance-medial; in the remaining 6, it was utterance-initial.

The sequences corresponding to the four cells of Table 4 were made from the original gated stimuli. For each utterance, the gating of subsequent context continued until the last point at which a listener in Experiment 1 had come to recognize the critical word, and the corresponding stimulus in the without-subsequent-context conditions used the same number of repetitions. One audiotope was prepared for each cell of the design.

Procedure. The stimulus materials were presented to the subjects over headphones. Subjects in the segmented/with-subsequent-context condition received the instructions given in Experiment 1. Subjects in the other conditions received appropriate descriptions of their stimulus materials, used appropriate practice items, and, in the case of the unsegmented left context conditions, had inter-stimulus intervals beginning with 12 sec. In other respects, the method was the same as in Experiment 1.

Results

To determine whether the segmented/with-subsequent-context condition replicated the results of Experiment 1, the distribution of recognition outcomes was compared with the distribution for the corresponding items in the earlier experiment. The differences were minimal: percentages of Experiment 1 and Experiment 3 outcomes in each class, respectively, were 27% and 33% immediate recognitions, 60% and 50.4% late recognitions, and 12% and 16% failures to recognize [$\chi^2(2) = 2.17, n = 200, n.s.$].

Next, separate pairs of 2 (segmented, unsegmented) \times 2 (with subsequent context, without subsequent context) ANOVAs were run on numbers of immediate, late, and missed recognitions. Subjects were nested in cells, and

stimuli were crossed with both factors. Figure 5 displays the mean results for each cell.

The segmentation of sentences at word boundaries did not reduce the number of immediate recognitions or inflate the number of late recognitions achieved. In fact, none of the three dependent variables showed a main effect of the segmentation of prior context or any interaction with segmentation (all F values for these effects have $p > .10$).

The only significant effects were due to the presence or absence of subsequent context. The mean number of late recognitions was greater with subsequent context than without [30.79 and 3.08, respectively: $F_1(1,44) = 366.27, p < .0001$; $F_2(1,59) = 231.73, p < .0001$; $\min F'(1,102) = 141.93, p < .001$]. The mean number of missed recognitions was less with subsequent context than without [9.125 and 43.79, respectively: $F_1(1,44) = 814.03, p < .0001$; $F_2(1,59) = 279.55, p < .0001$; $\min F'(1,92) = 208.09, p < .001$]. Because subjects were unable to better their original guesses given repetition alone, the conclusion that subsequent context is responsible for the late recognitions is compelling.

It would appear that listeners could not make use of simple repetition even when they intended to do so. Subjects in the repetition-without-subsequent-context conditions produced fewer immediate recognitions of critical words than those in the with-subsequent-context conditions [13.13 and 20.08, respectively: $F_1(1,44) = 22.55, p < .0001$; $F_2(1,59) = 34.36, p < .0001$; $\min F'(1,91) = 13.61, p < .005$]. Because the prior contexts were matched across these conditions, the effect on first-trial recognitions could not have been due to left context. Nor could it have been attributed to the content of right contexts, for these had not yet been heard when the critical word was first presented.

Instead, subjects in the repetition-without-subsequent-context conditions might merely have been holding back their guesses in the expectation of being able to improve on them in the course of further presentations of exactly the same material. But with no further information forthcoming, the total number of correct identifications scarcely rose (by an average of 3.08 late recognitions) with subsequent presentations. In fact, the average number of recognitions over all trials in these conditions (16.21) is close to what subjects in with-subsequent-context conditions achieved on first trials alone (20.08). Because the without-subsequent-context subjects had no more pertinent information on their last trials than both groups had had on their first presentations of these same critical words, it is not surprising that similar results were obtained. Because the with-subsequent-context subjects produced many late recognitions (30.79 on average), they ultimately recognized a large proportion of critical words (50.87 out of 60, or 85%). This recognition rate is within the reported range of recognition accuracies for careful transcription of dialog (see Note 1). It seems, then, that subjects in the without-subsequent-information conditions employed an ad hoc strategy which failed, whereas sub-

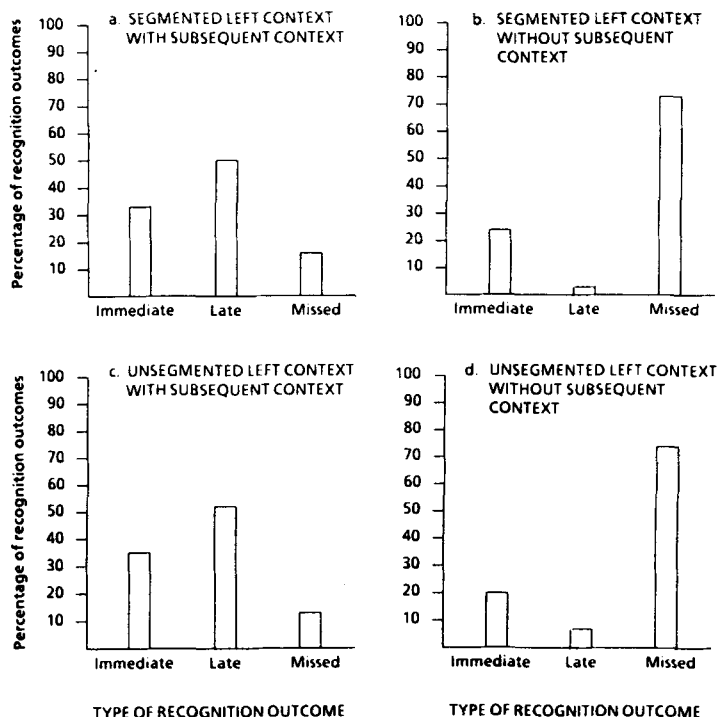


Figure 5. Distribution of recognition outcomes for each cell of Experiment 3. (For each cell, $n=720$.)

jects who heard subsequent context achieved what is normal in the recognition of conversational speech.

GENERAL DISCUSSION

The answers to the four central questions posed in the introduction to this paper all indicate that late recognition is a phenomenon of theoretical interest rather than the product of random error. First, late recognitions were not rare, but accounted for roughly one in five words recognized. Second, the majority of observed late recognitions followed patterns which are not readily explained as the result of left-to-right processes alone, but appear instead to implicate right context in the process of recognition. Third, late recognitions were not distributed randomly: word tokens which were recognized late tended to possess certain characteristics which are known to affect ease of recognition. Finally, late recognitions were not artifacts of the technique: they persisted when all of the preceding discourse was made available to listeners or when earlier parts of the sentence were not segmented into words, and they could not be made to occur merely by offering listeners repeated opportunities to examine the same word and left context. Only the addition of subsequent context made late recognition possible.

The characteristics of word tokens associated with late recognition suggest the sort of processes which may be involved. The word tokens which were recognized late tended to be shorter than those recognized with left context alone, possibly reflecting the fact that shorter tokens are less intelligible, even out of context, than long. The

tokens recognized late were also likely to be those with inadequate support from left context. Higher rates of late recognition were associated with smaller amounts of prior context: words closer to utterance onset and words presented with only utterance context were more often recognized late. Late recognitions depended also on amount of subsequent context: words not recognized on their first trial were more likely to be recognized late if more words followed in their utterance. A final characteristic concerns the syntactic and semantic nature of right contexts and is reflected in the functor/contentive form-class difference.

Function words were recognized late more often than were content words. Although a difference in average length of tokens in the two categories is partly responsible, the multiple regression equations for Experiment 1 showed an additional, independent effect for the functor/contentive distinction itself. Function words do not have straightforward interpretations which might be read without alteration from a conventional dictionary: the interpretation of a function word is often determined by its subsequent context, as example 1 illustrates for the word *to*:

- 1a. to skip
- 1b. to my mind
- 1c. to Edinburgh

The choice of function word, like its interpretation, depends on the construction to which it belongs. Whenever the function word is not the final element in a construction—and in a pre-positional language such as En-

glish, it most often is not—the choice may be constrained by subsequent as well as by preceding context. In example 2 below, for instance, an unintelligible word following *eats* might be any of the function words listed in parentheses. For examples 3a–3b, however, left and right context together restrict the possible choices to a much smaller set.

2. Benjamin eats ____ ... (*with, at, on, under, from, like, the, when, ...*)
- 3a. Benjamin eats ____ a spoon. (*with, from, without*)
- 3b. Benjamin eats ____ the table. (*at, on, under*)

In example 4, the prior context occurs in a separate utterance, example 4a, but the identity of the function word, *the* or *a*, still depends on both prior context and subsequent context. The listener attempting to recognize a short, indistinct token in the positions marked by blanks would not be helped by prior sentential context, because at the point at which the article is said, examples 4b and 4c are identical. The decisive context in each case is the noun *mug* or *cup*, but which article it selects depends on the number of mugs or cups previously mentioned. The asterisked items in examples 4b and 4c are unacceptable.

- 4a. You'll find a cup and some mugs on the shelf.
- 4b. Give Adam ____ mug. (**the, a*)
- 4c. Give Adam ____ cup. (*the, *a*)

These cases all suggest the value of processing subsequent context in instances in which function words cannot be recognized on the basis of their acoustic form alone.

Our findings demonstrate, then, that prior context, acoustic information, and subsequent context all have roles in the processing of connected speech. As we mentioned earlier, models of word recognition have been designed mainly to explain the roles of prior context and acoustic word shape; the role of subsequent context has been largely ignored. Although subsequent context may be processed by nonsequential mechanisms operating more or less in parallel with a left-to-right device (Cutler & Norris, 1988; Grosjean & Gee, 1987), a more parsimonious account involves a single mechanism which allows the interpretation of some part of the input to be affected in similar ways both by what precedes and by what follows it.

An adequate model of word recognition should employ contexts of various kinds to select a word candidate from among several competitors hypothesized on the basis of incomplete acoustic information. Such a general mechanism has already been proposed, for example, in TRACE (McClelland & Elman, 1986) and in the revised cohort model (Marslen-Wilson, 1987). In each case, however, the kinds of context which can affect word recognition are limited.

In TRACE, for example, each feature identified in the input activates all phoneme units having that feature, and these in turn increase the activation for all word units in which they participate. A poorly specified phonetic segment will not prevent the activation of the word which

contains that segment, because both the left and the right phonetic context will contribute to the correct recognition of the word. Although TRACE allows for left and right context effects within words and between adjacent words, it instantiates no higher-level processes which would allow syntactic, semantic, or pragmatic interpretations of the word's surroundings to bear on the process of recognition.

The revised cohort model, like TRACE, allows many word hypotheses to be activated by the same stretch of input. Activation level is determined solely by goodness of fit with the available acoustic input, and it is expected that the presence of a particular word token will usually cause the corresponding word hypothesis to be among the most highly activated. Even when one or more segments of a word are missing or incorrectly identified, the correct hypothesis may still be highly activated, because the correct and incorrect pronunciations are phonologically similar: *shigarette* will activate the representation of *cigarette*. Unlike TRACE, the revised cohort model also contains an interpretative level within which an account of the meaning and structure of the utterance is being constructed. Higher-level processes assess all active word hypotheses in parallel for integration into this developing interpretation; the process of integration constitutes recognition. The model allows for the recognition of a word on the basis of only its earlier segments, that is, for early recognition. The match between the input and the first few segments of the word candidate is sufficient to activate that candidate and permit its assessment by the higher-level processes. As a result, that candidate may be integrated into the utterance interpretation before all the acoustic information is processed which would make it recognizable out of context.

The revised cohort model also allows for postoffset recognition in a particular set of instances: words which do not diverge phonemically from all other words before their own final phoneme. Such words are homophonous with the early parts of longer words: for instance, *boy* may become *boycott* or *boisterous*. When the shorter word is said and prior context of whatever kind does not distinguish it from its competitors before its offset, recognition of the word must wait for that later point at which the acoustic information can no longer support the word's competitors. Luce (1984) has pointed out that the majority of words heard will fall into this category. Luce's account is framed in terms of citation pronunciations. The problem is presumably more widespread for tokens produced in running speech.

The majority of these cases, however, should be resolved within the next word: *boy from* is clearly not a token of *boisterous*, *boyish*, or *boycott*. The late recognitions observed in the present experiments, however, were often so delayed and of such chronological complexity as to make it unlikely that postoffset uniqueness points of this kind were responsible. Some more general account is needed both of how a word token may not be recognized during its acoustic lifetime and of how it comes to be recognized later. In what follows, we will show how

this account might be captured within the framework provided by the revised cohort model.

In instances in which words are not immediately recognized, we assume first that the word token does activate the correct word hypothesis during its acoustic lifetime, albeit weakly and with many competitors. This assumption is necessary to preserve the modular architecture of the revised cohort model, within which only the analysis of current acoustic input can generate word hypotheses. Second, it must be assumed that a late recognized word token should be unintelligible if heard in isolation, for if the token had been a particularly clear one, the priority accorded to acoustic evidence in this model should yield an immediate recognition.

Finally, it must be assumed that the left context (defined at whatever level) of a late-recognized word is insufficiently constraining to select among the weakly active candidates. Whether or not a word hypothesis is incorporated into the interpretation by its offset is, therefore, some function of the constraint offered by its left context and the goodness of fit between the stored representation of the word and the acoustic signal. The revised cohort model currently depends on these two sources of information. And it is well established that when left context offers more constraint, whether syntactic, semantic, or discoursal, then a poorer acoustic specification is sufficient for recognition (Hunnicut, 1985; Lieberman, 1963; Marslen-Wilson & Welsh, 1978). The present suggestion is that the consequences of this complementary relationship include the possibility that no word hypothesis will be selected for integration by word offset, because there are some values of the context/clarity function at which selection is impossible. In a model giving priority to acoustic evidence, clarity is the variable which contributes more strongly to the failure of immediate recognition.

If word recognition has failed to occur by word offset, processing must continue through the input corresponding to the next word. We propose that right context will make a late recognition possible when it offers sufficient acoustic or interpretative constraint to allow the integration of the earlier weakly activated word hypothesis. Whether or not a word is recognized at all becomes a joint function of left context, acoustic specification, and right context. Under current models, however, it is not clear how subsequent words are recognized when parts of their left context remain unresolved.

The present results demonstrate that the incorporation of words into an interpretation need not correspond to the chronological order of those words in the input. Accordingly, a word must be capable of being represented at the level of the interpretation without there existing a complete representation of its left context. In such cases, of course, recognition need not proceed without contextual help. Many of the sources of information available for uninterrupted left-to-right processing remain. The topic of the discourse, the most plausible antecedents of pronouns, the prevailing tense, and so on are still avail-

able. Even without the immediately preceding word or words, syntactic constraints still exist. These aspects of context could function much as they would on uninterrupted interpretations.

Furthermore, the difficulty of recognizing words with incomplete left contexts should diminish where relatively unambiguous stretches of the acoustic input, such as syllables with unreduced vowels, provide highly activated hypotheses which require only minimal contextual constraint for disambiguation. Two approaches already mentioned here suggest active processes which respond to input of this type. Cutler and Norris (1988) proposed that strong syllables, which are often word-initial, are actually used to demarcate new words, a particularly useful function in instances in which the words have incomplete left contexts. Grosjean and Gee (1987) suggest that a process centering on phonologically stressed syllables serves to organize surrounding uninterpreted material, with the stressed syllables keying entries in the mental lexicon. Within the framework adopted here, however, late recognition simply emerges from the regular functioning of a mechanism in which alternative hypotheses are entertained simultaneously until such time as there is sufficient evidence to select one hypothesis above the others. Clearly articulated syllables correspond to "islands of certainty" simply because they would tend to activate a cohort containing relatively few strong competitors to the intended word.

For a substantial proportion of the time, this more passive processor must be pursuing hypotheses across a stretch of input corresponding to more than one word. The processor may, in fact, be assessing against the interpretation not only all of the locally competing word candidates, but also all of the *combinations* of successive candidates. That is, listeners may employ their capacity for dealing with syntactic ambiguity in tandem with their capacity for dealing with phonological ambiguity: they may assess against the existing interpretation every complete syntactic path through the series of competing word candidates.

Whichever mechanism or mechanisms actually serve to deliver the information held in words' right contexts, a full model of word recognition must accommodate to the limits of the processor's performance both in optimal conditions, with maximally clear articulation, and in the usually less than optimal conditions which obtain in normal conversational speech. A picture emerges of the processing of normal conversational speech in which the processor is regularly confronted with stretches of low intelligibility in the speech stream which may only be resolved with reference to subsequent context.

REFERENCES

- BARD, E. G. (1982). *The intelligibility of speech addressed to children*. Unpublished doctoral dissertation, University of Edinburgh, Scotland.
 BARD, E. G., & ANDERSON, A. H. (1983). The unintelligibility of speech to children. *Journal of Child Language*, 10, 265-292.
 BARD, E. G., ANDERSON, A. H., & LAVER, J. (1982). *The intelligibil-*

- ity of speech to children (Economic and Social Research Council [U.K.] Project Report No. HR6130, London).
- BRADLEY, D. (1978). *Computational distinctions of vocabulary type*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge.
- BROWN, G., CURRIE, K., & KENWORTHY, J. (1980). *Questions of intonation*. London: Croom Helm.
- COHEN, J., & COHEN, P. (1983). *Applied multiple regression analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- COLE, R. A., & JAKIMIK, J. (1980). A model of speech production. In R. A. Cole (Ed.), *Perception and production of fluent speech* (pp. 133-163). Hillsdale, NJ: Erlbaum.
- COTTON, S., & GROSJEAN, F. (1984). The gating paradigm: A comparison of successive and individual presentation formats. *Perception & Psychophysics*, **35**, 41-48.
- 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.
- DALBY, J. (1984). *The phonetic structure of fast speech in American English*. Unpublished doctoral dissertation, Indiana University, Bloomington.
- FRANCIS, W. N., & KUČERA, H. (1982). *Frequency analysis of English usage*. Boston: Houghton Mifflin.
- 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.
- HUNNICUTT, S. (1985). Intelligibility versus redundancy—conditions of dependency. *Language & Speech*, **28**, 47-56.
- LIEBERMAN, P. (1963). Some effects of the semantic and grammatical context on the production and perception of speech. *Language & Speech*, **6**, 172-175.
- LUCE, P. (1984). *A computational analysis of optimal discrimination points in auditory word recognition* (Research on Speech Perception Progress Report No. 10, pp. 1-8). Bloomington: Indiana University.
- MARSLÉN-WILSON, W. D. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, **244**, 522-523.
- MARSLÉN-WILSON, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, **25**, 71-102.
- MARSLÉN-WILSON, W. D., & TYLER, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, **8**, 1-71.
- MARSLÉN-WILSON, W. D., TYLER, L. K., & SEIDENBERG, M. (1978). The semantic control of sentence segmentation. In W. J. M. Levelt & G. B. Flores d'Arcais (Eds.), *Studies in the perception of language* (pp. 219-246). London: Wiley.
- MARSLÉN-WILSON, W. D., & WELSH, A. (1978). Processing interactions during word-recognition in continuous speech. *Cognitive Psychology*, **10**, 29-63.
- MCCLELLAND, J., & ELMAN, J. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **18**, 1-86.
- PICKETT, J. M., & POLLACK, I. (1963). Intelligibility of excerpts from fluent speech: Effects of rate of utterance and duration of excerpt. *Language & Speech*, **6**, 151-165.
- POLLACK, I., & PICKETT, J. M. (1963). The intelligibility of excerpts from conversation. *Language & Speech*, **6**, 165-171.
- POLLACK, I., & PICKETT, J. M. (1964). Intelligibility of excerpts from fluent speech: Auditory vs. structural context. *Journal of Verbal Learning & Verbal Behavior*, **3**, 79-84.
- ROSENZWEIG, M. K., & POSTMAN, L. (1958). Frequency of usage and the perception of words. *Science*, **127**, 26-36.
- SHOCKEY, L. (1973). *Phonetics and phonology of conversational speech* (Working Papers No. 17). Columbus: Ohio State University, Department of Linguistics.
- TYLER, L. K., & WESSELS, J. (1983). Quantifying contextual contributions to word-recognition processes. *Perception & Psychophysics*, **34**, 409-420.
- TYLER, L., & WESSELS, J. (1985). Is gating an on-line task? Evidence from naming latency data. *Perception & Psychophysics*, **38**, 217-222.
- WOODS, W. A., BATES, M., BROWN, G., BRUCE, B., COOK, C., KLOVSTAD, J., MAKHOUL, J., NASH-WEBBER, B., SCHWARTZ, R., WOLF, J., & ZUE, V. (1976). *Speech understanding systems: Final technical progress report (Report No. 3438, 5 volumes)*. Cambridge, MA: Bolt, Beranek and Newman, Inc.

NOTES

1. The failures of recognition (12% of outcomes) are approximately what would be expected from reports of disagreements in transcribing spontaneous dialogue (see Bard, 1982, chap. 2, for a review), and do not necessarily indicate poor-quality material.

2. Similar categories were used by Grosjean (1985), but only with respect to the order of recognition of a single pair of adjacent words in each sentence. The current categorization takes into account all the words following each late recognition.

3. Another possible source of artifact, the time available for making recognition decisions, has been dealt with by Tyler and Wessels (1985), who found no difference between timed and untimed responding in gating tasks with short, fixed gate duration. Given the scope of the listener's transcription task in the present experiments, the time pressure is considerable and the task closely approximates a timed gating task. Both because of Tyler and Wessels's results and because of the nature of the task in Experiment 1, the effects of permitted response latencies will not be explored further here.

(Manuscript received September 18, 1987;
revision accepted for publication May 20, 1988.)