Spontaneous segmentation in normal and in time-compressed speech

ARTHUR WINGFIELD and KAREN A. NOLAN
Brandeis University, Waltham, Massachusetts 02254

An experiment is reported in which subjects heard recorded speech passages at normal speaking rates and at various degrees of time compression, with instructions to interrupt the speech input for intermediate report by stopping the tape at points of their choosing. An analysis of subjects' points of interruption as a function of speech rate eliminated temporal duration and number of words per segment as primary determiners of segmentation strategy. Formal linguistic constituents at all speech rates served as a good, but not perfect, predictor of segmentation boundaries. The results are interpreted in the light of on-line interactive and augmented transition network (ATN) models of speech perception.

From the earliest use of sound spectrography in speech research have come arguments that speech must be processed in "units" of certain sizes and that the perception of these units is heavily context dependent. While numerous studies have investigated linguistic elements the size of phonemes and syllables (Cole & Scott, 1974; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Massaro, 1972), it seems likely that studies will reflect processing units on a variety of levels, depending on the demand characteristics of the particular experimental tasks (cf. Foss & Swinney, 1973; Healy & Cutting, 1976; McNeill & Lindig, 1973).

Of immediate interest are a number of studies that have suggested the role of phrases and syntactic clauses as among the higher order processing units on the sentence level (e.g., Bever, Lackner, & Kirk, 1969; Fodor & Bever, 1965; Garrett, Bever, & Fodor, 1966). These early studies traditionally employed the technique of subjective localization of extraneous sounds ("clicks") embedded in recorded sentences. The fact that the accuracy of localizing these sounds and that the distribution of errors could be reliably predicted by major syntactic boundaries led most writers to conclude that formal linguistic constituents have a certain perceptual integrity which resists disruption. Clausal structure was seen as the primary determiner of perceptual segmentation on the sentence level (Fodor, Bever, & Garrett, 1974).

This research was supported by PHS Grant NS-14662 from the National Institutes of Health to the first author. We are grateful to Dr. Emerson Foulke, Center for Rate Controlled Recordings, University of Louisville, for assistance in time-compression of the speech materials, and to Dennis Byrnes and François Grosjean for comments on an earlier draft. A number of critical comments and suggestions offered by our reviewers were incorporated into the final manuscript. These reviewers also have our thanks. Requests for reprints should be sent to Arthur Wingfield, Department of Psychology, Brandeis University, Waltham, Massachusetts 02254.

However ingenuous the technique, and intuitively appealing the conclusions, the procedure of click localization itself became sufficiently controversial to distract from the reasons these experiments were originally performed. Criticisms were related not only to the relative contribution of surface or deep structure boundaries in perceptual segmentation (Chapin, Smith, & Abrahamson, 1972), but also to whether the effects observed were indeed perceptual, as opposed to possible contaminants of attention, memory, or response bias influencing the results (cf. Bertelson, 1972; Ladefoged, 1967; Reber, 1973; Reber & Anderson, 1970).

In spite of these criticisms, there must be some lower and upper bound limits on the size of ordinary processing segments in ongoing speech perception and language comprehension. On the one hand, not only must phoneme-by-phoneme or word-by-word analysis on the sentence level be rejected on formal grounds (Fodor, Bever, & Garrett, 1974), but also the postulate of serial, cumulative processing after each and every utterance would be inefficient or impossible to complete, given the rates at which we ordinarily hear and comprehend speech. By serial, cumulative processing, we mean the analysis of the complete meaning of the first word, storing the outcome of this analysis for the arrival of the next word, holding the combined analysis of the meaning of these two words for the arrival of the third word, and so forth. On the other hand, a passive holding of entire sentences prior to any analytic processing would put inordinate demands on any short-term, transient memory store one might wish to contemplate, whether one conceptualizes this limitation in terms of classical short-term memory (Atkinson & Shiffrin, 1968; Shiffrin, 1976), resource allocation within a limited-capacity, central processor (Norman & Bobrow, 1975), or limits on the speed of "deep" processing (Craik & Lockhart, 1972). Indeed, we already have some evidence that memory span for

connected speech is heavily influenced by its syntactic structure (cf. Aaronson, 1976; Jarvella, 1971; Jarvella & Herman, 1972).

Any performance model of language processing must thus address the question of the presumed size, nature, and, indeed, the existence of ordinary processing units on the sentence level and the extent to which the size of such units might be limited by memory or short-term processing considerations. While it is difficult to relate standard memory-capacity studies to the continuous processing of ongoing speech, we can assume that listeners impose some subjective segmentation on heard speech and that correlates of syntactic structure supply the necessary information for this segmentation. An important step yet to be taken is to determine the extent to which performance limitations in the form of transient memory capacity or duration may influence the size and nature of these processing units.

In the following experiment, subjects heard recorded paragraph-length passages for immediate recall. The passages were specifically designed to exceed short-term memory capacity, such that accurate recall could be accomplished only if subjects were allowed to interrupt the passages periodically and report what was heard on a segment-by-segment basis. There were two aspects to this study. First, we intended to analyze the points at which the subjects spontaneously interrupted the passages for intermediate recall. These segmentation strategies might show some regularity and, in turn, allow us to make some guided inferences about the size and character of the analytic units listeners may ordinarily impose on the processing of ongoing, fluent speech.

The second aspect of this study made use of artificially accelerated, or "time-compressed," speech in order to separate potential effects of performance limitations due to speech content vs. segment duration on the size of processing units. Time compression can increase speech content per unit time while still keeping overall intelligibility of the speech materials relatively high (Foulke, 1971; Wingfield, 1975a). If time per segment were a major determinant of segment size, as might be due to decay time limitations in a short-term holding memory (e.g., Reitman, 1974; Wingfield & Byrnes, 1972), one would expect segment duration to be independent of increasing speech rate and speech content per segment to show a systematic increase with rate. Alternatively, if segment size is determined primarily by speech content, either in terms of lexical content or syntactic constituents, speech content per segment should be relatively independent of speech rate. If this were the case, the average time per segment would systematically decrease with increasing speech rate.

It is certainly the case that caution must always be exercised if we wish to infer, from studies based in whole or part on verbatim memory (cf. Aaronson, 1976; Jarvella, 1971), what might ordinarily be done in

normal speech processing. On the other hand, determination of spontaneous segmentation strategies within these experimental conditions could offer a paradigm from which to build a sufficient data base to guide such inference.

METHOD

Speech Materials

The stimuli consisted of 30 120-word passages selected from recent issues of the *American Scientist* magazine. The passages covered a wide range of topics and were selected to offer understandable, but high-information (low-redundancy), speech samples representing a variety of syntactic forms. Proper names and difficult-to-pronounce technical terms were eliminated from the speech passages, and preliminary studies were used to match the passages for overall difficulty and intelligibility. Each of the 30 passages was recorded by a native speaker of American English in normal intonation at an average speech rate of 185 words/min.

The original master recording of all 30 passages was then time-compressed, using the so-called "sampling method" on an electromechanical compressor of the Fairbanks type (cf. Foulke, 1971, pp. 81-82). In this method, small (20-msec) tape segments are periodically deleted at regular intervals, with the remaining segments then abutted in time. When played back at normal speed, the result is speech reproduced in less than normal time, but without the distortion in pitch or quality which would, for example, accompany tape-recorder playback at faster than normal speed. The degree of time-compression is controlled by the frequency with which the tape segments are deleted.

Compressions were to 80%, 70%, and 60% of original playing time, corresponding to speaking rates of 231, 264, and 308 words/min. Together with the original recordings at normal speech rate, a total of 120 passages at four separate rates were thus available for the experiment.

Apparatus and Procedures

Subjects heard the speech passages binaurally on a TEAC Model 3300 tape recorder through Sharp MK II earphones at an average sound pressure level of 65 dB (re: .0002 dynes/cm²). The subjects' task was to listen to each passage as it was presented and to repeat aloud, as accurately as possible, as much of the passage as they could. It was explained to them that they were not to "shadow" the speech as they heard it (i.e., attempting to speak and listen simultaneously). This would, in any event, be virtually impossible during the high-speed-rate conditions. Instead, they were instructed to stop the tape recorder at intervals of their choosing to allow themselves time to report what they had heard. They were not to begin their reports until they had stopped the recorder.

The subjects were told to operate the pause control on the recorder to stop the input for intermediate reports whenever they wished, and as often as they wished, during the course of a passage, but to try to stop the machine no more frequently than necessary to insure good recall accuracy in their intermediate reports. The subjects were told only that their responses were being recorded for later scoring for accuracy. No mention was made of the fact that our primary interest was in *where* they chose to stop the recorder for intermediate reports. A second machine was used to record the subjects' spoken responses and the positions within the passages where the stimulus recorder was stopped.

Each subject heard a total of 20 test passages, five passages at each of the four speech rates employed (normal, 80%, 70%, and 60% compression). The particular passages heard at each rate were varied between subjects such that, by the end of the experiment, each passage was heard an equal number of times at each rate. No passage was heard more than once by the same subject.

The five passages at each rate were blocked and were preceded by two additional practice passages at that rate to familiarize the

subjects with the listening task. Twenty-eight university undergraduates served as subjects. All reported normal hearing, and all spoke American English as their first language. None reported previous experience with listening to time-compressed speech.

RESULTS

The subjects' responses were scored on four dependent measures: recall accuracy, number of words between successive responses, elapsed time between successive responses, and a linguistic analysis of observer segmentation intervals.

Recall Accuracy

Recall accuracy was based on the middle 100 words of each passage, with the first and last 10 words of each passage excluded from scoring to eliminate the influence of end effects. Credit was given for all words reported correctly, regardless of order. No partial credit was given for incomplete words. Failure to give a grammatical ending when it was present (e.g., reporting *quickly* as *quick*) lost credit for the entire word.

Table 1 shows mean recall scores and standard deviations for each of the speech-rate conditions. Analysis of variance following an arc sin transformation showed the overall reduction in recall accuracy from normal speech rate to 60% compression to be significant [$F(3,81) = 77.90, p < .001$]. On the other hand, the absolute magnitude of the reduction was small, representing only a 16.8% drop from normal speech rate to 60% compression (a 40% increase in speech rate), and even this fastest speech rate yielded 74.5% accuracy with our fairly strict scoring procedures.

Number of Words and Elapsed Time

The average number of words per recall segment was fairly uniform across speech rates, with mean segment lengths of 8.8, 7.9, 8.5, and 8.6 words obtained at normal, 80%, 70%, and 60% compression [$\chi^2(27) = 1.63, n.s.$]. While the absence of a systematic speech rate appears to favor a speech-content rather than speech-duration explanation of segmentation intervals, the variability of segment sizes from one segment to another was too large to be predictive,

Table 1
Mean Percentage of Words Recalled and Standard Deviations for Passages Heard at Normal Speech Rate and at Three Degrees of Compression When Subjects Were Allowed to Interrupt the Speech Input for Intermediate Reports

Compression Ratio (Percent Normal Time)	Recall Accuracy (Percent Correct)	
	Mean	SD
Normal	89.6	9.51
80	85.3	6.27
70	82.7	7.56
60	74.5	9.12

Table 2
Mean Duration (in Seconds) and Standard Deviations of Segments Spontaneously Selected for Intermediate Recall as a Function of Degree of Time Compression of the Speech Passages

Compression Ratio (Percent Normal Time)	Segment Duration	
	Mean	SD
Normal	3.0	.55
80	2.3	.38
70	1.9	.37
60	1.7	.35

ranging from as little as 1 to over 20 words between successive reports.

Consistent with the negation of segment duration as a primary variable in segmentation strategy, the measurement of mean time durations of recall segments showed a clear inverse effect of speech rate. Table 2 shows these mean durations and standard deviations for each of the speech rates tested. The overall reduction in segment duration with increasing speech rates was significant [$F(3,81) = 224.75, p < .001$], with the mean percentage reduction in segment duration almost exactly proportional to the percentage increase in speech rate. The variability in individual segment durations, however, was such that segment duration, like the number of words per segment, had little predictive quality in a priori determination of where individual passages would be segmented for report.

Linguistic Analysis of Input Segmentation

Since neither time duration nor simple number of words per segment seemed to define input segmentation across all speech rates, segmentation intervals were analyzed in terms of a phrase structure, linguistic constituent analysis.

Figure 1 shows, for each speech rate, the total frequency with which the speech passages were stopped for report at clearly defined syntactic boundaries and, within this, the relative frequency with which the passages were interrupted at the end of a sentence or clause (S boundary), between a subject noun phrase and verb phrase (NP-VP), between phrases or clauses within a verb phrase (within VP), or between a noun phrase and prepositional phrase under the same NP node (NP-PP).

Although the question of a hierarchy of boundary types is a complex one within linguistic theory (Rotenberg, 1978), syntactic structure of the recall segment was generally a good predictor of segmentation, accounting for between 68% and 73% of the segmentation data, depending on speech rate. There was some effect of speech rate on the frequency with which recall segments coincided with syntactic boundaries [$F(3,57) = 2.85, p < .05$], but pair-by-pair comparisons showed only the extreme differences between normal rate and 60% compression to be sig-

Figure 1. Total frequency with which speech passages were interrupted for intermediate recall at clearly defined syntactic boundaries, and relative contribution of each of four types of syntactic boundary to total segmentation strategy. Data are shown separately for each speech-rate condition.

nificant, and this was due almost entirely to a small reduction in the frequency of segmentation at S boundaries at the fastest speech rate.

The relative contribution of type of syntactic boundary to the total segmentation data can be slightly misleading, since some types of syntactic boundary occur more frequently than others in speech, such that the apparent importance of relatively minor, but frequent, boundaries will be inflated. When the data in Figure 1 are adjusted for the number of times a particular boundary type occurred in the passages, averaged across all speech rates, subjects segmented the materials at 64% of all S boundaries when they occurred, 26% of all NP-VP boundaries, 14% of all within-VP boundaries, and 9% of all NP-PP boundaries.

DISCUSSION

The generally good intelligibility at all speech rates employed can be accounted for by the subjects' freedom to segment the passages at points of their choosing, allowing time for input processing and report. This would be expected in view of the argument that the debilitating effects of compressed speech are due as much to depriving listeners of ordinarily available processing time, as to degradation of the speech signal itself (Aaronson, Markowitz, & Shapiro, 1971; Overmann, 1971).

One must be cautious using the terms "perception" or "memory" in connection with speech data such as ours. For some, "perception" implies immediate processing at the phonemic level, while others include perceptual organization at a syntactic, or sentence, level. For our present purposes, we are reluctant to employ an arbitrary distinction between "perception" and "memory," favoring a more con-

tinuous "levels" of processing approach without clear input-output boundaries (cf. Craik & Lockhart, 1972).

It is certainly the case, however, that there must be some upper limits on the size of processing units, bounded by the capacity and duration of the memory store in which they are presumably held for analysis. Analysis at the syllabic level, for example, has been identified with an auditory memory image lasting perhaps 100-250 msec, but certainly under 1 sec (Crowder & Morton, 1969; Huggins, 1974; Massaro, 1972). Analysis on the clausal level, on the other hand, would seem to imply the characteristics and presumed duration of storage more typically associated with short-term memory (Atkinson & Shiffrin, 1968; Broadbent, 1971; Shiffrin, 1976).

The significant reduction in duration of input report segments with increasing time compression implies that time duration is not the determining factor in segmentation at normal speech rates. The average time duration of segments ranged from 3 sec at normal speech rate to 1.73 sec at 60% compression. That is, as the average duration of speech elements was reduced by 40%, so the duration per segment was reduced by a nearly equivalent 42%. This proportional reduction in segment duration held for all compression rates with mean segment durations ranging from 2% to 5% shorter than would be predicted by compression ratio alone. Time compression may thus have served to increase processing load to a small degree over and above its simple effect on duration of the linguistic elements.

While time and/or capacity must clearly exist as limiting factors to a theoretical maximum segment size which could be held for analysis, speech content, as defined by syntactic structure, is a better predictor of subjects' segmentation intervals than either elapsed time or simple number of words per segment. This latter finding is robust, with the listeners' relative use of the boundaries remaining virtually unaffected by increasing speech rate. Indeed, adjusted for the average frequency with which the particular boundary types in our speech passages, segmentation frequencies were directly related to the linguistic importance of these boundaries regardless of speech rate. Evidence for within-major-clause segmentation is thus firmly established.

As clear as these particular results are, it would be wrong at this stage of the research to prematurely close on a fixed upper limit on the size of the analytic segment. That is, any data showing segmentation using major clause constituent units does not directly specify the upper-bound limits on unit size: one clause long, two clauses long, one sentence plus one clause, etc. The reason this is so is based on the possibility that conceptual content and syntactic complexity of a message are likely to interact with upper-bound size of an analytic segment. Put most simply, the limits

obtained may be different for conceptually complex vs. conceptually simple prose, an issue best resolved by further empirical tests. Specifically, our goal is to argue for the unit of measurement to be used as further research attempts to define the boundary limitations.

There is a further, and equally important, qualification to these otherwise straightforward results. Although segmentation intervals coincided with formal linguistic constituents in the clear majority of cases, even at normal speech rate some 27% of all segments could not be defined as ending at a recognized syntactic boundary. Some cases might have been due to subjects accidentally stopping the recorder when not intended or to stopping the input for report immediately after an unusual or unfamiliar word. For example, 22% of the subjects stopped for report after the word "isotropic," even though it did not terminate a syntactic boundary.

A third possible source of segmentation at unexpected places could also be implied from our data. Subjects frequently stopped the input for report at points which were not the ends of sentences or clause boundaries, but which might have been, given the preceding context. For example, several subjects segmented after hearing "There are many types of forests," only to discover when they restarted the tape that this had not been the end of a sentence, but that it went on to include: "... ranging from those of the humid tropics to fairly open woodland." Our listeners seemed not to be operating on a word-by-word basis, or on entire sentences or clauses passively held prior to any analysis. Rather, ongoing speech processing seemed to represent a continual attempt to impose structure on the heard utterances, with a preliminary analysis of each element in the speech stream being used to potentially confirm, reject, or modify these initial hypotheses as to structure. As implied by Chapin, Smith, and Abrahamson (1972), our subjects in these cases attempted to close boundaries at the earliest possible point: If a given word sequence could be a sentence or major clause, listeners may assume that it is one, until later context makes it obvious that it was not. The present paradigm would tend to highlight such events ordinarily unnoticed in natural speech perception.

Of particular relevance to these observations are several recent refinements of temporal on-line processing notions offered by the "knowledge driven" system of Marslen-Wilson (Marslen-Wilson, 1975; Marslen-Wilson & Welsh, 1978; Tyler & Marslen-Wilson, 1977), and the so-called augmented transition network (ATN) model of speech processing described by, for example, Kaplan (1972, 1973), Rumelhart (1977), and Wanner and Maratsos (Note 1). These notions offer the beginnings of the understanding of a formal system in which processing could proceed in an interactive on-line way, with

listeners doing as much syntactic and semantic analysis as they can on-line, as soon as they can.

Specifically, one could contemplate a language processor forming initial hypotheses of syntactic structure based on speech heard to a certain point, rejecting, sustaining, or modifying this hypothesis as additional information is received. One consequence of such ongoing, continuous processing would be a constant recording and consolidation of the input, such that capacity limits to a short-term memory store would rarely be reached.

This notion of ongoing interactive processing could weaken the earlier claims for processing units, in favor of identifying periodic crucial points of analysis. The difference is between the claim that clause boundaries are critical consolidation points, milestones, in analysis vs. the earlier notion that what lies between these points represents an integral unit held long enough for one-shot, top-down processing. The difference is a subtle, but important, one. It is important because it might force us to consider reconceptualizing upper-bound limits on unit size in terms of critical intervals for analysis. The distinction is a subtle one because the span between such intervals would also be expected to have upper-bound, linguistically based limits.

We can perhaps view the listener as a hypothesis maker, one concerned, as Shank (1972) has argued, as much with relationships among concepts (nominals, actions, and their modifiers), as among syntactic constituents, *per se*. In such a system, the primary function of syntactic processing would be to determine, for example, the main noun and verb in a sentence. Consider the following example, in which the slashes indicate a common segmentation strategy: "Those who performed manual labor / and worked with tools / were looked upon as lower in capacity than those who did not work." Several listeners segmented after *labor* (S boundary) and after *tools* (NP-VP boundary), points at which complete relationships have been expressed. Since relationships among concepts are typically, but not invariably, completed at syntactic boundaries, those boundaries most predictive of segmentation may, on further research, be those associated with such completion.

Further research using this and other procedures will be necessary to determine the full range of factors which may operate in subjects' segmentation hypotheses. It is likely, for example, that prosodic features such as stress, pausing, and intonation contour may play a key role in segmentation (Darwin, 1975; Wingfield, 1975a, 1975b; Wingfield & Klein, 1971). On the structural level, certain verbs may cue segmentation immediately after a subsequent noun phrase, while others may result in segmentation immediately after a subsequent prepositional phrase. Such a finding would point not only to a fundamental interface between syntax and semantics, but also might

lead to an emphasis on language performance strategies largely overlooked in the past two decades of psycholinguistic research.

REFERENCE NOTE

1. Wanner, E., & Maratsos, M. M. *On understanding relative clauses*. Unpublished manuscript, Department of Psychology and Social Relations, Harvard University, 1974.

REFERENCES

- AARONSON, D. Performance theories for sentence coding: Some qualitative observations. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, 2, 42-55.
- AARONSON, D., MARKOWITZ, N., & SHAPIRO, H. Perception and immediate recall of normal and "compressed" auditory sequences. *Perception & Psychophysics*, 1971, 9, 338-344.
- ATKINSON, R. C., & SHIFFRIN, R. N. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2). New York: Academic Press, 1968.
- BERTELSON, P. Listening from left to right vs. right to left. *Perception*, 1972, 1, 161-165.
- BEVER, T. G., LACKNER, J. R., & KIRK, R. The underlying structures of sentences are the primary units of speech processing. *Perception & Psychophysics*, 1969, 5, 225-231.
- BROADBENT, D. E. *Decision and stress*. New York: Academic Press, 1971.
- CHAPIN, P. G., SMITH, T. S., & ABRAHAMSON, A. A. Two factors in perceptual segmentation of speech. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 164-173.
- COLE, R. A., & SCOTT, B. Toward a theory of speech perception. *Psychological Review*, 1974, 81, 348-374.
- CRAIK, R. I. M., & LOCKHART, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 671-684.
- CROWDER, R. G., & MORTON, J. Precategorical acoustic storage. *Perception & Psychophysics*, 1969, 5, 365-373.
- DARWIN, C. J. On the dynamic use of prosody in speech perception. In A. Cohen & S. G. Nooteboom (Eds.), *Structure and process in speech perception*. Berlin: Springer-Verlag, 1975.
- FOULKE, E. The perception of time-compressed speech. In D. Horton & J. Jenkins (Eds.), *The perception of language*. Columbus, Ohio: Merrill, 1971.
- FODOR, J. A., & BEVER, T. G. The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior*, 1965, 4, 414-420.
- FODOR, J. A., BEVER, T. G., & GARRETT, M. F. *The psychology of language*. New York: McGraw-Hill, 1974.
- FOSS, D., & SWINNEY, D. A. On the psychological reality of the phoneme: Perception, identification and consciousness. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 246-257.
- GARRETT, M. F., BEVER, T. G., & FODOR, J. A. The active use of grammar in speech perception. *Perception & Psychophysics*, 1966, 1, 30-32.
- HEALY, A. F., & CUTTING, J. E. Units of speech perception: Phoneme and syllable. *Journal of Verbal Learning and Verbal Behavior*, 1976, 15, 73-83.
- HUGGINS, A. W. F. On perceptual integration of temporal phenomena in speech. *Journal of the Acoustical Society of America*, 1974, 51, 1279-1290.
- JARVELLA, R. J. Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 409-416.
- JARVELLA, R. J., & HERMAN, S. J. Clause structure of sentences and speech processing. *Perception & Psychophysics*, 1972, 11, 381-384.
- KAPLAN, R. M. Augmented transition networks as psychological models of sentence comprehension. *Artificial Intelligence*, 1972, 3, 77-100.
- KAPLAN, R. M. A general syntactic processor. In R. Rustin (Ed.), *Natural language processing*. New York: Algorithmics Press, 1973.
- LADEFOGED, P. *Three areas of experimental phonetics*. London: Oxford University Press, 1967.
- LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P., & STUDDERT-KENNEDY, M. Perception of the speech code. *Psychological Review*, 1967, 74, 431-461.
- MARSLER-WILSON, W. D. Sentence perception as an interactive parallel process. *Science*, 1975, 189, 226-227.
- MARSLER-WILSON, W. D., & WELSH, A. Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 1978, 10, 29-63.
- MASSARO, D. W. Preperceptual images, processing time, and perceptual units in auditory perception. *Psychological Review*, 1972, 79, 124-145.
- MCCNEILL, D., & LINDIG, K. The perceptual reality of phonemes, syllables, words and sentences. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 419-430.
- NORMAN, D. A., & BOBROW, D. G. On data-limited and resource-limited processes. *Cognitive Psychology*, 1975, 7, 44-64.
- OVERMANN, R. A. Processing time as a variable in the comprehension of time-compressed speech. In E. Foulke (Ed.), *Proceedings of the Second Louisville Conference on Rate and/or Frequency-Controlled Speech*. Louisville, Ky: University of Louisville Press, 1971.
- REBER, A. S. What clicks may tell us about speech perception. *Journal of Psycholinguistic Research*, 1973, 2, 286-287.
- REBER, A. S., & ANDERSON, J. R. The perception of clicks in linguistic and nonlinguistic messages. *Perception & Psychophysics*, 1970, 8, 81-89.
- REITMAN, J. S. Without surreptitious rehearsal information in short-term memory decays. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 365-377.
- ROTENBERG, J. *The syntax of phonology*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, 1978.
- RUMELHART, D. E. *Human information processing*. New York: Wiley, 1977.
- SHANK, R. C. Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, 1972, 3, 552-631.
- SHIFFRIN, R. M. Capacity limitations in information processing, attention, and memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 4) *Attention and memory*. Hillsdale, N.J: Erlbaum, 1976.
- TYLER, L. K., & MARSLER-WILSON, W. D. The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, 1977, 16, 683-692.
- WINGFIELD, A. Acoustic redundancy and the perception of time-compressed speech. *Journal of Speech and Hearing Research*, 1975, 18, 96-104. (a)
- WINGFIELD, A. The intonation-syntax interaction: Prosodic features in perceptual processing of sentences. In A. Cohen & S. G. Nooteboom (Eds.), *Structure and process in speech perception*. Berlin: Springer-Verlag, 1975. (b)
- WINGFIELD, A., & BYRNES, D. L. The decay of information in short-term memory. *Science*, 1972, 176, 690-692.
- WINGFIELD, A., & KLEIN, J. F. Syntactic structure and acoustic pattern in speech perception. *Perception & Psychophysics*, 1971, 9, 23-25.

(Received for publication January 23, 1979;
revision accepted March 31, 1980.)