

# The extraction of structure during reading: Evidence from reading prosody

ASHER KORAIAT

*University of Haifa, Haifa, Israel*

SETH N. GREENBERG

*Union College, Schenectady, New York*

and

HAMUTAL KREINER

*University of Haifa, Haifa, Israel*

According to the structural approach to reading, the extraction of structure precedes the analysis of meaning and paves the way for it. In this study, reading prosody was used to examine this proposition. Specifically, we tested the hypothesis that reading prosody produced on line reflects the early extraction of structure. In Experiment 1, readers were successful in assigning natural prosody to unfamiliar text immediately upon its reading. Experiment 2 showed that the prosodic patterns applied are tuned to the structure of the sentence and are largely indifferent to the content of the sentence or to its semantic coherence. The results join with other findings in speech production and comprehension in supporting the precedence of structure to meaning in reading.

A great number of studies of reading concern the processes that enable comprehension to occur on line. Text comprehension requires the coding of structure, as well as the coding of meaning (see Bock, 1990; Garrett, 1980). According to the structural precedence hypothesis, the processing of structure precedes the analysis of meaning and paves the way for it. Early in text processing, readers attempt to derive a structural frame for the text. This frame is assumed to guide the processing of content-laden units and to facilitate their assimilation into a coherent semantic representation. Most of the evidence in support of this hypothesis comes from work on letter detection during silent reading of text (for a review, see Koriat & Greenberg, 1994). The letter detection task, however, has been criticized as disrupting natural reading and, therefore, as being of limited validity for the study of on-line text processing (Rayner & Pollatsek, 1989). In this study, we examine yet another reading behavior—prosody—that occurs on line and can shed some light on the early analysis of structure. Prosody refers to the intonation, rhythm, and stress applied

when reading a sentence aloud. We propose that reading prosody is the output of processes that are concerned with the on-line analysis of text; therefore, it can provide insight into early structural processing.

This proposal receives some support from an observation we made when studying structural processing in reading. Consider the Jabberwocky poem by Lewis Carroll (1900): “’Twas brillig, and the slithy toves did gyre and gimble in the wabe. . . .” We observed that when asked to read this poem aloud, readers applied a fluent and coherent prosodic pattern that was quite consistent across different readers. Note that although the content units in this poem are nonsense, the function units preserve the morphosyntactic structure characteristic of English, and the prosody applied by readers seems to reflect that structure. Thus, we propose that reading prosody provides a window to the kind of structural processing that is assumed to take place on line prior to complete semantic analysis. In the experiments to be reported, we first tested the hypothesis that natural reading prosody is produced on line. Then, we proceed to show that the prosody applied on line during reading is tuned primarily to the structure of the sentence and is extracted largely independently of meaning.

---

This research was supported by a grant from the Israel Foundations Trustees and by the Max Wertheimer Minerva Center, University of Haifa. We are grateful to Hana Struminger for programming the experiments, Rachel Shitzer-Reichert for conducting them, and Limor Sheffer for helping in the analyses of the data. Correspondence concerning this article should be addressed to A. Koriat, Department of Psychology, University of Haifa, Haifa, Israel (e-mail: akoriat@research.haifa.ac.il).

—Accepted by previous editorial team

## **The Hypothesis of Structural Precedence: Evidence from Letter Detection Patterns**

Our letter detection work derived its impetus from the extensive research of Healy and her colleagues on the *missing-letter effect* (MLE). They observed that when readers search for a target letter in connected text, they miss

that letter more often when it appears in very common words, such as *the*, *and*, and *for* (for a review, see Healy, 1994). According to Healy's unitization account, familiar words are processed in terms of unitized, whole-word representations, thus preempting access to their constituent letters. In support of this account, several studies have indicated that fewer detection errors are made when the familiar shape of a common word is disrupted—for example, by misspelling or by case alternation (Drewnowski & Healy, 1977; Healy & Drewnowski, 1983). Recent modifications of the unitization view assume that the rate of letter detection errors is determined by processing time (Moravcsik & Healy 1995, 1998).

The MLE attracted our interest because the most common words in English, those that yield an inordinately high rate of letter omissions, are function words. Many students of reading argue that function words are particularly helpful as diagnostic clues to sentence structure (e.g., Just & Carpenter, 1987; Kimball, 1973). Others, in contrast, maintain that these words are largely redundant and are simply skipped over during reading (see Haber & Schindler, 1981). These contrasting views seem to reflect a discrepancy between those who stress structural organization and those who emphasize semantic content in reading.

In an attempt to resolve this discrepancy, we proposed a structural model of reading, according to which, early in text processing, readers rely on a variety of cues, mostly morphosyntactic, to establish a structural frame for the phrase or the sentence. This frame then guides semantic analysis and the integration of individual elements into an overall meaning representation. Functors play a crucial role in the analysis of structure, often signaling the construction of a new phrase (see Kimball, 1973), and are therefore monitored early in text processing. However, once a sentential or phrasal framework has been specified, the conveyers of that frame become less prominent, in comparison with the semantically rich content units. Thus, the MLE reflects a figure-ground organization of the sentence, in which the structural skeleton of a sentence helps to push forward the semantically rich units.

Indeed, a large number of experiments with English, Hebrew, French, and German (e.g., Koriat & Greenberg, 1991; Koriat, Greenberg, & Goldshmid, 1991; Moravcsik & Healy, 1995, 1998; Müsseler, Koriat, & Nißlein, 2000; Saint-Aubin & Poirier, 1997) provide evidence for the contribution of syntactic factors to letter detection. They indicate that the MLE for common words stems primarily from their syntactic function within the sentence. In Hebrew, for example, some function morphemes can be expressed as a one-letter prefix appended to a content word, thereby creating orthographic sequences that are not particularly frequent. Such single-letter function prefixes were more difficult to detect than the initial letters of content words (Koriat et al., 1991). Also, nonwords in both English and Hebrew produced more detection errors when placed in function slots in text than when placed in content slots (Koriat & Greenberg, 1991). Similarly, the initial letter of Hebrew nonwords was more difficult to detect

when context biased its interpretation in favor of a function prefix than when context led readers to interpret it as part of the stem of a content word. In French, Saint-Aubin and Poirier found more detection errors for the word *or* when it was used as a function word (meaning *however* or *whereas*) than when it was used as a content word (meaning *gold*). Greenberg and Koriat (1991) also observed that some function words in English reveal or conceal their letters, depending on their structural role within a phrase (e.g., *on* in *on switch* vs. *on my way*). Similarly, Moravcsik and Healy (1995, 1998) observed fewer errors for a function word when it was used as an adjective or adverb (e.g., “the *in* clothes,” “*the* harder you study”) than when it was used as a preposition. Findings obtained with German indicate that the structure of the sentence as a whole also affects the magnitude of the MLE, suggesting that the MLE is generated after phrase structure has been tentatively specified (Müsseler et al., 2000). Taken together, these findings support the idea that letter detection patterns reflect processes dedicated to unraveling the structural organization of text.

### Reading Prosody as a Window to Structural Processing

The hypothesis of structural precedence in reading accords well with the emphasis on structural processing in theories of speech production. According to frame-and-slot models of speech production (see Dell, 1986; Garrett, 1975), the internal representations of utterances to be spoken are constructed by inserting linguistic items into slots in independently created structures. Bock (1990) has stressed the independence between structure and meaning. Levelt (1989) further proposed that the prosody of a spoken utterance is derived from prelexical structural representation, and Ferreira (1993) showed that prosodic timing patterns roughly correspond to the sentence's syntactic structure (although she also stressed that syntactic and prosodic patterns are not completely isomorphic; see also Gee & Grosjean, 1983; Nespor & Vogel, 1987; Selkirk, 1984).

Much of the research and theorizing on prosody (including Levelt's, 1989, and Ferreira's, 1993, models) has concerned speech rather than reading. However, we propose that in reading, as in speech production, structure sets the stage for content and that the prosody applied on line when text is read captures the structural frame established for the phrase or the sentence. In fact, Goldman-Eisler (1972) argued that reader's prosody is even closer to the “ideal delivery” of grammatical structure than is the speaker's prosody. This was based on the finding that whereas in spontaneous speech less than a third of breathing pauses occur at clause boundaries, in reading aloud all occur at these locations.

Let us illustrate our thesis with the following two sentences:

1. The windy horse was singing when the books ate the house and saw all the justice.
2. The little girl was sleeping when the burglars entered the house and took all the jewelry.

In an exploratory study, we found that when asked to read aloud the first sentence, proficient readers consistently applied a fluent and coherent prosody that was very similar to the one they applied to the second sentence, which has the same syntactic structure. These observations suggest three hypotheses: First, natural coherent prosody is applied on line when a sentence is read. Second, the prosodic patterns applied roughly reflect the structure of the sentence. Third and most important, the correspondence between structure and prosody is preserved even in the absence of meaning, consistent with the idea that syntactic structure is extracted independently of semantic processing and possibly prior to it. These hypotheses were examined in the experiments to be reported.

Experiment 1 tested the proposition that natural reading prosody is produced on line. It examined the hypothesis that the prosody applied in reading a sentence immediately when it is presented is practically as natural as that applied after readers familiarized themselves with that sentence. Experiment 2 used reading prosody to examine the proposition that the structural analysis of a sentence is carried out prior to integrative semantic analysis. Hence, prosodic patterns were expected to respond to manipulations that change or destroy structure but to be indifferent to manipulations that change or destroy meaning. Because of our concern with structural analysis, in Experiment 2 we focused on only one aspect of prosody—pause pattern, which has been assumed to reflect primarily syntactic structure. Clearly, other prosodic parameters are prominent in disclosing the effects of pragmatic and semantic information, but we will not consider them here.

In sum, the experiments reported below are intended to show that reading prosody reflects the structural organization of the sentence, that this organization can be extracted by a reader on line without advance preparation, and that it is generated with relative independence of semantic information. Results consistent with these propositions will support our view that reading prosody discloses an intermediate representation of a sentence, one that follows structure analysis but precedes more complete semantic analysis.

## EXPERIMENT 1

Experiment 1 had two parts. In the first part, participants were recorded as they read aloud sentences under three conditions. In the *unpracticed* condition, they read each sentence as soon as it appeared on the screen. In the *practiced* condition, their reading was recorded only after they had familiarized themselves with the sentence. In the *arbitrary* condition, which provided a baseline, they had to read the sentence according to punctuation marks inserted in arbitrary positions. In the second part, the recorded versions of all sentences were presented to judges who rated the naturalness of the reading prosody of each sentence. If the unpracticed prosody is rated as being as natural as

the practiced prosody, that would indicate that reading prosody reflects on-line processing and can be used to examine the structural precedence hypothesis.

## Method

### Reading Part

**Participants.** Twelve University of Haifa students (5 males and 7 females), whose native language was Hebrew, were paid for participating in this part of the experiment.

**Stimulus materials.** Twelve Hebrew sentences containing 15–16 words were composed. Three matched sentences represented each of four different syntactic structures. All four syntactic structures were complex. Sentences were always presented on one line in order to avoid a pause when moving to a new line. In the unpracticed and practiced reading conditions, the sentences were presented with no punctuation marks, except for a period at the end of the sentence. In the arbitrary condition, two commas were placed in arbitrary positions in the sentence, but never in the natural clause or phrase boundaries.

**Procedure and Apparatus.** The participants were told that the experiment concerned the processes engaged by a reader when reading text for broadcasting. They were asked to read aloud each sentence clearly and with the correct intonation, as if they were reading it “on-the-air,” because their “broadcast” reading would be recorded and then evaluated by a group of judges. Each participant read all 12 sentences—first, in the unpracticed condition; second, in the practiced condition; and finally, in the arbitrary condition. In the unpracticed condition, the participants were asked to read aloud each sentence as soon as it appeared on the computer screen. In the practiced condition, the same list of sentences was presented again, but the participants read each sentence aloud four times, and only the fourth reading was recorded. Finally, in the arbitrary condition, the participants were told to read the sentences four times but to take care to read them according to the punctuation marks. Only the fourth reading was recorded. The order of presentation of the sentences conformed to a Latin square based on three sequences of four steps, so that all four syntactic structures were represented in each sequence and, across all participants, each syntactic structure was equally represented in every step. The experiment was controlled by a Silicon Graphics Indigo workstation. Sentences were presented on a screen, and the readings were recorded by the computer’s audio system. A keypress by the participant initiated the presentation of the sentence on the screen, and a second keypress terminated the recording and caused the sentence to disappear from the screen.

### Judgment Part

**Participants.** Twelve new University of Haifa students (5 males and 7 females), whose native language was Hebrew, were paid for participating in this part of the experiment.

**Stimulus materials.** The sentences recorded in the reading part of the experiment served as the stimuli. Each of the 12 sentences was played three times to each judge, once in each of the three reading conditions (unpracticed, practiced, and arbitrary). Altogether, each judge heard 36 sentences in such a way that the three sentence recordings representing the three reading conditions of each sentence were read by the same reader and each set of 3 such sentences was read by a different reader. Otherwise, the assignment of sentences to judges and the order of presentation to each judge were random.

**Procedure and Apparatus.** The judges saw each sentence on the computer screen and read it aloud twice. Then, they heard one of the recordings of that sentence and had to rate how natural the reading prosody sounded on a scale from 1 (*very low*) to 10 (*very high*). As a referent, the judges were told that a synthesized computer message, like a monotonous phone-information message, represented an example of low prosodic naturalness, whereas reading with clear punctuation according to the content of the message was mentioned as an example of high prosodic naturalness. The judges were encouraged

to use the entire range of ratings. They could hear the sentence again by pressing the space bar before typing their rating.

### Results and Discussion

Figure 1 presents the means of the naturalness ratings for the four syntactic structures for each reading condition. As can be seen, the arbitrary condition yielded much lower naturalness ratings (2.72) than did the unpracticed and practiced conditions (7.38 and 7.95, respectively), which differed only slightly between themselves. This pattern was obtained for each of the four syntactic structures: A syntactic structure  $\times$  reading condition analysis of variance (ANOVA) yielded  $F(2,22) = 176.48, p < .0001$ , for reading condition and  $F < 1$  for both syntactic structure and the interaction. A post hoc analysis (Scheffé) indicated that the difference between the practiced and the unpracticed reading conditions was not significant, whereas both differed significantly from the arbitrary condition. This result was obtained for the overall means, as well as for the means of each and every syntactic structure.

The results were also consistent across the 12 readers. The naturalness ratings for each reading condition were averaged for each reader across all judges and sentences. A syntactic structure  $\times$  reading condition ANOVA on the readers' means yielded a significant effect for reading condition [ $F(2,22) = 289.09, p < .0001$ ], but not for syntactic structure or for the interaction ( $F < 1$  for both). A post hoc analysis (Scheffé) indicated again that the difference between the practiced and the unpracticed conditions was not significant but that both differed significantly from the arbitrary condition.

These results clearly indicate that natural reading prosody is produced on the first reading of an unfamiliar sentence. Thus, whereas practiced and unpracticed readings were both judged as significantly more appropriate than arbitrary readings, the improvement resulting from the familiarization with the sentence was not significant. This

pattern of results was consistently found across different judges, different readers, and different syntactic structures and hence appears to be of broad generality. In addition, although it may seem trivial, the finding that the arbitrary condition yielded very low ratings of naturalness suggests that listeners rate reading prosody as more "natural" or appropriate when it conforms to the structural organization of the sentence.

These findings suggest that reading prosody reflects on-line processing. What are the underlying mechanisms that enable readers to produce natural prosody on line during reading? In Experiment 2, we examine the proposition that structural cues, possibly analyzed prior to semantic integration, play a major role in this process.

### EXPERIMENT 2

In Experiment 2, the participants were recorded as they read sentences aloud, and the prosodic pause patterns applied were used to examine the structural precedence hypothesis. Using different versions of each sentence, we tested two hypotheses. The first was that prosodic pause patterns reflect sentential structure. Therefore, well-structured sentences that represent different syntactic structures should yield clearly differentiated prosodic patterns, but when structure is destroyed, pause patterns should become less differentiated. The second hypothesis was that pause patterns are derived on line, on the basis of structural cues, prior to complete semantic integration. Therefore, they should be indifferent to modifications that change or even destroy meaning while preserving sentence structure.

### Method

#### Participants

Six Hebrew-speaking female students participated in the experiment for course credit. (Females were chosen because of technical considerations; the voice analysis apparatus was calibrated to female voices.)

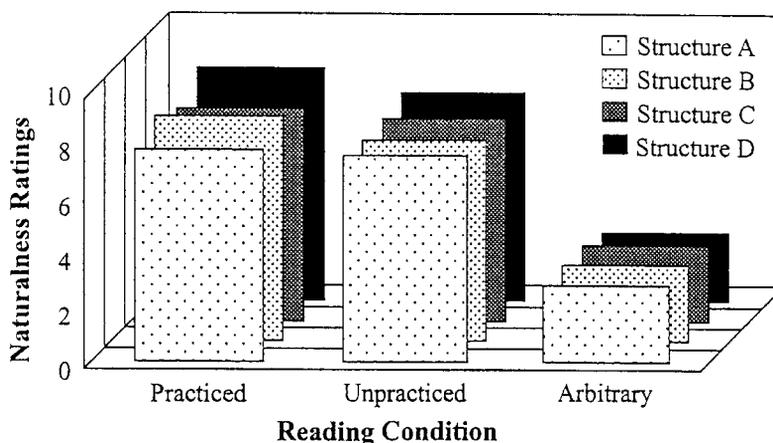


Figure 1. Mean ratings of naturalness for practiced, unpracticed, and arbitrary readings for sentences with four different syntactic structures (Experiment 1).

**Table 1**  
**Core Sentences Representing Syntactic Structures A and B**  
**and the Five Sentences Derived From the Core Sentence**  
**Representing Structure A**

Semantic Coherence	Structural Integrity	Sentence
Structure A		
Semantic	intact (core)	1. The fat cat <sup>1</sup> with the gray stripes <sup>2</sup> ran <sup>3</sup> quickly <sup>4</sup> to the little kitten <sup>5</sup> that lost <sup>6</sup> its way <sup>7</sup> in the noisy street.
	partial	2. Fat cat <sup>1</sup> gray stripe <sup>2</sup> run <sup>3</sup> quick <sup>4</sup> little kitten <sup>5</sup> loose <sup>6</sup> way <sup>7</sup> noise street.
	random	3. Way gray <sup>1</sup> run cat <sup>2</sup> kitten <sup>3</sup> loose <sup>4</sup> noise fat <sup>5</sup> street <sup>6</sup> stripe <sup>7</sup> quick little.
Nonsense	intact	4. The sad gate <sup>1</sup> with the small electricity <sup>2</sup> went <sup>3</sup> carefully <sup>4</sup> to the happy computer <sup>5</sup> that sang <sup>6</sup> the leafs <sup>7</sup> in the front book.
	partial	5. Sad gate <sup>1</sup> small electricity <sup>2</sup> go <sup>3</sup> careful <sup>4</sup> happy computer <sup>5</sup> sing <sup>6</sup> leaf <sup>7</sup> front book.
	random	6. Sing small <sup>1</sup> go gate <sup>2</sup> computer <sup>3</sup> leaf <sup>4</sup> front sad <sup>5</sup> book <sup>6</sup> electricity <sup>7</sup> happy careful.
Structure B	(core)	7. The old horse <sup>1</sup> that was injured <sup>2</sup> on the first race <sup>3</sup> last year <sup>4</sup> received <sup>5</sup> all the medical care <sup>6</sup> from the most experienced vet <sup>7</sup> in the racing field.

Note—The sentences were translated literally from Hebrew. The numbers inside the sentences indicate the locations of the measured pauses.

## Design

Four factors were manipulated in constructing the stimulus materials: syntactic structure, semantic content, structural integrity, and semantic coherence.

## Stimulus Materials

Examples of the sentences used in the experiment (translated as literally as possible from Hebrew) appear in Table 1.

Six normal Hebrew sentences were composed, three representing each of two different syntactic structures (e.g., Sentences 1 and 7 in Table 1). The three sentences representing each structure had the same number of words and contained the same function elements in the same ordinal positions. These sentences served as the core sentences, from which the other sentence types were derived. To destroy semantic coherence, every content word in the core sentence was replaced with another content word from the same lexical category so as to produce a nonsense sentence (e.g., Sentence 4 in Table 1, which was derived from Core Sentence 1). The destruction of structure was performed in two steps. First, partially structured sentences were created by eliminating all function elements from the core sentence (e.g., Sentence 2 in Table 1, derived from Core Sentence 1). Second, unstructured (or random) "sentences" were then derived by randomly reordering the content words of the partially structured sentences (e.g., Sentence 3 in Table 1, derived from Sentence 2). The same manipulations were also applied to the nonsense sentences (e.g., Sentence 5 in Table 1 was derived from Sentence 4, and Sentence 6 was derived from Sentence 5). Thus, from each of the six core sentences, we generated a set of 6 sentences representing all combinations of three levels of structural integrity (normal, partial, and random) and two levels of semantic coherence (semantic and nonsense). Sentences 1–6 in Table 1 represent a complete set. Altogether, there were six different sets, comprising 36 sentences.

In addition to the experimental sentences, 24 normal sentences (15–16 words each) were constructed as fillers, representing different syntactic structures. These served to balance the ratio of normal to distorted sentences and to increase the variability of syntactic structures.

## Procedure and Apparatus

The instructions and procedure were similar to those of the reading part in Experiment 1, except for the following. The participants were asked to read the sentences as soon as they were presented. They were told that some of the sentences were not regular but that they should nevertheless read them as if they were normal sentences, using the appropriate intonation. Each sentence appeared on the screen on one line. The experiment began with 3 practice sentences, one of which was a nonsense sentence. The order of presentation of the 60 experimental and filler sentences was determined randomly for each participant, except for the following restrictions. First, each block of 10 consecutive sentences included 6 experimental and 4 filler sentences. Second, the 6 experimental sentences in a block were drawn from each of the six different sets and represented all combinations of the semantic coherence and structural integrity conditions. Finally, 3 sentences in each block were derived from Structure A, and the other 3 from Structure B. The experimenter initiated the presentation of each sentence when the participant was ready and terminated it when the participant finished reading. There was a short break after each block of 10 sentences.

The experiment was conducted on the same workstation as Experiment 1, with a Motorola DSP56001 audio and digital signal processing system. This system analyzed the voice signal and presented it graphically both as a spectrogram and as an amplitude graph of input speech log normalized and weighted by a Hamming window. Both graphs were based on sampling voice at a rate of 8000 Hz in frequencies of between 100 and 4000 Hz.

## Measurements

The basic measure was the duration of pauses in prespecified sentence locations. In previous studies on pause patterns during reading, participants were deliberately asked to artificially slow down their reading pace in order to allow a measure of pauses between each pair of successive words (e.g., Grosjean, Grosjean, & Lane, 1979). This procedure, however, could not be applied here because of our focus on natural prosody derived on line. A pilot study indicated that

in normal reading, participants do not pause after every word but, rather, connect 2–3 words without a real pause. Hence, we decided to measure only the pauses occurring at phrase boundaries, which are generally longer and more distinct (see Cooper & Paccia-Cooper, 1980; Goldman-Eisler, 1972; Martin, 1970). As was mentioned earlier, the sentences were matched so that all fully structured core and nonsense sentences contained seven phrase boundaries. In the partly structured sentences, pauses were measured at the locations corresponding to the phrase boundaries in the core sentence. In the unstructured sentences, we measured the pauses that occurred in the serial positions corresponding to the phrase boundaries in the partly structured sentence from which they were derived (see Table 1, which indicates the seven locations at which pauses were measured). Thus, a vector of seven pause durations for each sentence will be referred to as the pause pattern of that sentence.

A graphic presentation was used for measuring the reading pauses. The beginning and end of each pause were located and marked by an independent judge, and the duration of the pause was then calculated by the computer. (Specifying the beginning and end of a pause on the basis of some algorithm raised problems. Previous studies also relied on judges for measuring pauses—e.g., Ferreira 1993; Goldman-Eisler, 1972; Grosjean et al., 1979.) In all, there were 36 pause patterns for every participant.

### Results and Discussion

Before addressing the two hypotheses, we will first examine intersubject consensus. Kendall's coefficient of concordance<sup>1</sup> was used to evaluate the similarity between the pause patterns produced by the 6 participants when reading

the same sentence. This index, to be referred to as  $W(\text{sub})$ , was calculated for each of the 36 experimental sentences. The  $W(\text{sub})$  values for the normal sentences ranged from .45 to .86 and averaged .65. They were significant for each of the six normal sentences. If we assume that the natural prosody for a sentence is the one on which different readers agree, the high intersubject agreement in pause patterns joins with the results of Experiment 1 to suggest that readers produce natural prosody on line when they read a new sentence.

### Prosodic Similarity Between Sentences Differing in Meaning

We turn now to examining the hypothesis that reading prosody is tuned to the structure of the sentence and is relatively indifferent to its content. This hypothesis is supported by the results presented in Figure 2, which suggest that pause patterns are relatively insensitive to manipulations that change or destroy the semantic content of the sentence while keeping its structure intact. The top panels in this figure present the pause patterns for sentences that differ in their content but share the same syntactic structure, either A or B, whereas the bottom panels present the same data for the respective nonsense sentences. Each line in the figure represents pause durations, averaged across the 6 participants, as a function of the serial position of the

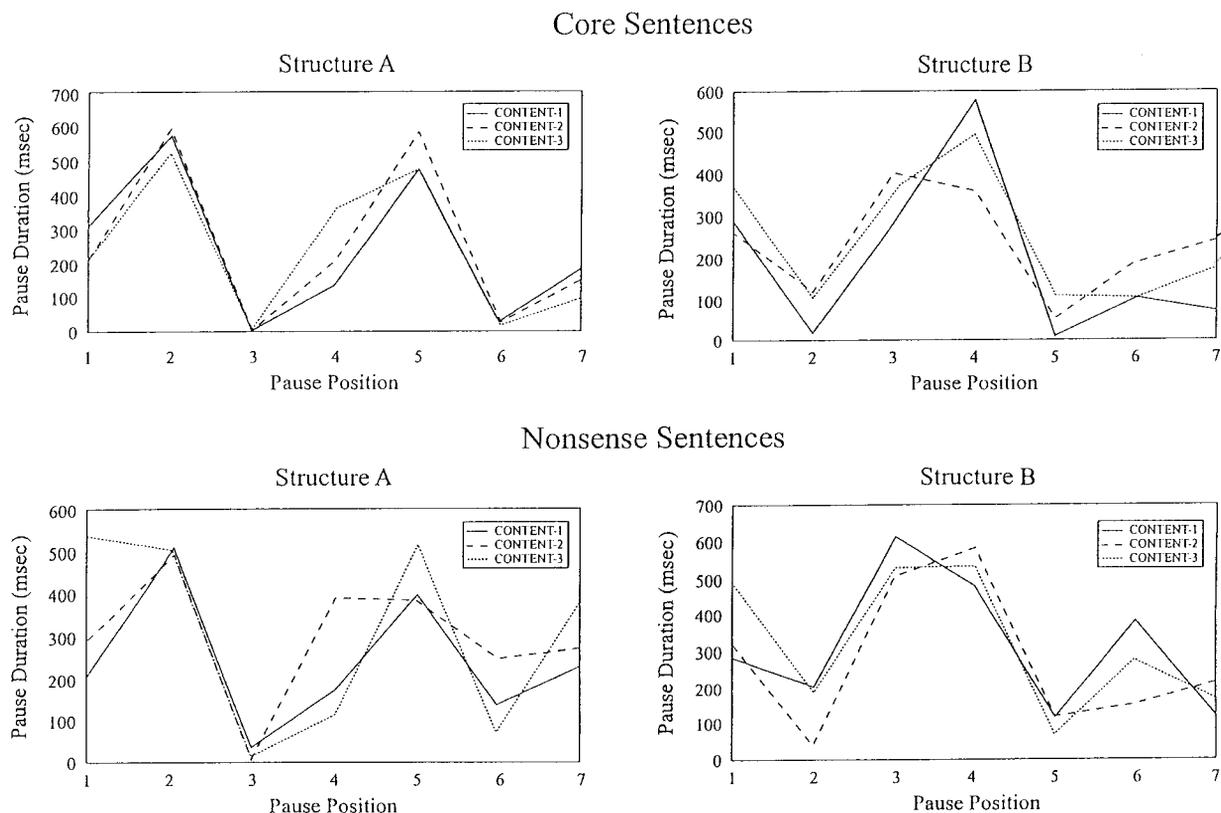


Figure 2. Pause patterns for fully structured core (meaningful) and nonsense sentences for Syntactic Structures A and B (Experiment 2).

**Table 2**  
**Range, Standard Deviation, and Mean of Kendall's  $W(\text{sent})$  Coefficients for Sentences**  
**Representing Different Levels of Structural Integrity and Semantic Coherence**

Structural Integrity	Semantic Sentences			Nonsense Sentences			All (Average $W$ )
	Range	$SD$	$M$	Range	$SD$	$M$	
Intact	.53-.93	.12	.77	.37-.88	.18	.66	.72
Partial	.17-.87	.23	.46	.24-.78	.16	.49	.47
Random	.09-.68	.14	.30	.12-.68	.18	.37	.34
All			.51			.51	.51

Note—The  $W(\text{sent})$  coefficient refers to the concordance between three sentences differing only in their content.

pause in the sentence. It can be seen that the three sentences sharing the same syntactic structure yielded very similar pause patterns, yet these patterns differed for sentences with different syntactic structures. Importantly, the similarity between the profiles of same-structure sentences was obtained whether the sentences were meaningful or nonsense—that is, devoid of coherent meaning (Figure 2, bottom panels).

To evaluate these trends statistically, Kendall's coefficient of concordance,  $W$ , was used again. This time it served as a measure of the similarity between the pause patterns applied to sentences that share the same syntactic structure and differ only in their content. Kendall's coefficient of concordance between the three sentences with the same syntactic structure, designated as  $W(\text{sent})$ , was calculated for each participant. As can be seen in Table 2 (first row), the 12 coefficients for normal sentences (6 participants  $\times$  2 syntactic structures) ranged from .53 to .93 and averaged .77. A similar analysis applied to the nonsense sentences yielded an average  $W$  of .66. Thus, pause patterns applied to sentences that share the same syntactic structure and differ only in their semantic content seem to be very similar. This similarity is relatively high even in sentences that are devoid of coherent semantic content.

The similarity of prosodic patterns seems to be due to the shared syntactic structure, and indeed, when syntactic structure is degraded, similarity decreases. Thus, the magnitude of Kendall's  $W(\text{sent})$  coefficients in Table 2 drops sharply as structural integrity is progressively impaired, whereas the effects of destroying semantic coherence are very small. Indeed, a structural integrity  $\times$  semantic coherence ANOVA on the 72 individual  $W(\text{sent})$  coefficients yielded  $F(2,71) = 23.09, p < .0001$ , for structural integrity,  $F < 1$  for semantic coherence, and  $F(2,71) = 5.2, p < .05$ , for the interaction. A post hoc analysis (REGWQ) indicated that  $W(\text{sent})$  coefficients were significantly higher for normal than for partly structured sentences and were significantly higher for the partly structured than for the unstructured sentences. To clarify the source of the interaction, separate one-way ANOVAs were carried out on each level of structural integrity. The results yielded  $F < 1$  for both the partly structured and the random sentences and  $F(1,23) = 6.76, p < .05$ , for the

structured sentences, suggesting that semantic information affects prosody only when structural integrity is intact. Structural integrity, on the other hand, exerted a significant effect for both levels of semantic coherence, yielding  $F(2,35) = 32.98, p < .0001$ , for semantic sentences, and  $F(2,35) = 8.38, p < .005$ , for nonsense sentences.

Taken together, these results suggest that pause patterns are relatively insensitive to semantic modifications: Sentences with different content yield very similar pause patterns if they have the same structure. Moreover, whereas the reduction or elimination of structural integrity reduces the degree of similarity, elimination of semantic coherence exerts only a small and mostly nonsignificant effect on pause patterns.

### Prosodic Differentiation Between Sentences Differing in Structure

The results reported in the previous section demonstrate that pause patterns are primarily sensitive to structural modifications. In this section, we focus specifically on the effects of structural modifications on pause patterns.

**Structural differentiation.** Figure 3 (top panel) presents mean pause patterns for the structured sentences representing Syntactic Structures A and B. These means were obtained by averaging the results across participants and across the three sentences representing the same structure. The results are presented separately for meaningful and nonsense sentences. It can be seen that the pause patterns associated with the two syntactic structures differ clearly from each other and that the difference between them is preserved regardless of semantic coherence. The correlations between the pause patterns support this impression: Whereas the correlations between the pause patterns associated with Syntactic Structures A and B are negative ( $-.67$  for meaningful sentences and  $-.77$  for nonsense sentences), those between the meaningful and the nonsense versions of the same syntactic structure are positive and high (.93 for Structure A and .94 for Structure B). Thus, sentences derived from the same syntactic structure produce very similar pause patterns, whereas sentences derived from different syntactic structures yield very distinct pause patterns.

The specific relationship between prosodic pattern and sentence structure has been discussed by many researchers

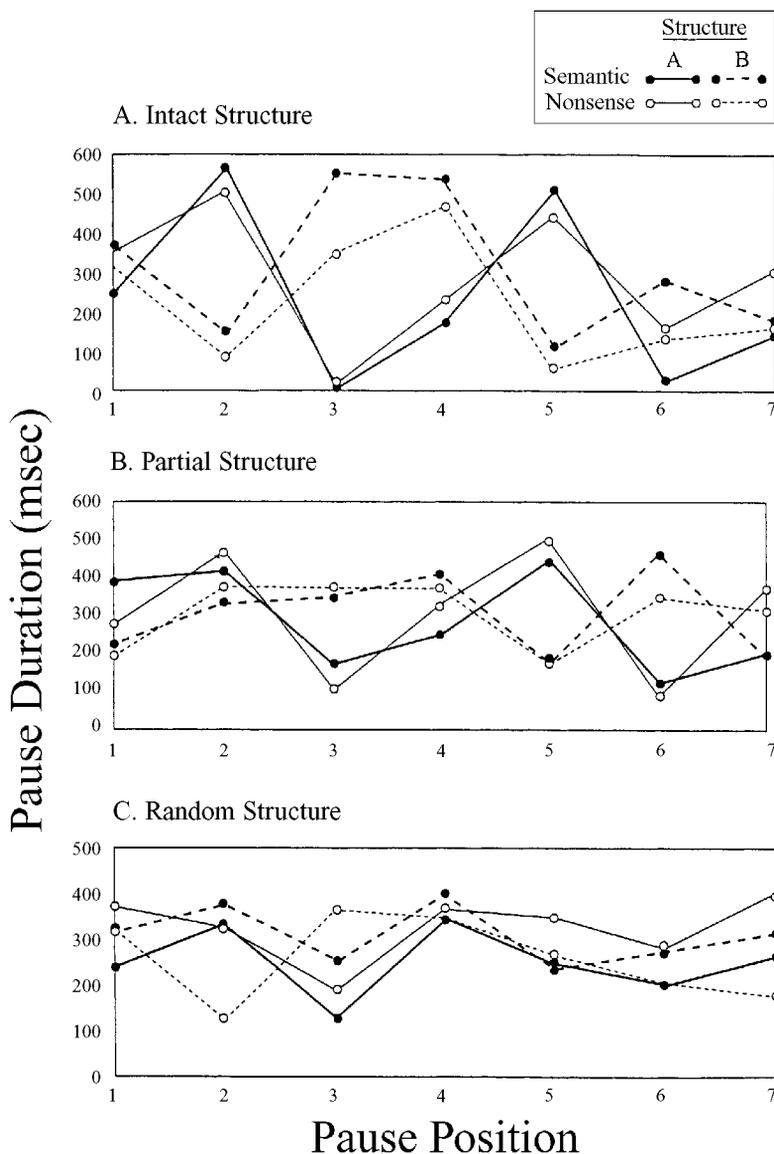


Figure 3. Pause patterns for Syntactic Structures A and B for semantic and nonsense sentences for the three conditions of structural integrity (Experiment 2).

in the context of spoken language (see Cutler, Dahan, & van Donselaar, 1997) and is beyond the scope of the present study. Nevertheless, we should note that in the normally structured sentences, the longer pauses generally occurred at the major phrase boundaries. For example, Pause 2 in Structure A occurred between the NP and the VP, whereas Pause 5 in Structure A and Pause 4 in Structure B occurred between the main clause and the relative clause. This observation is consistent with accumulating evidence from spoken language studies suggesting that prosodic and syntactic structures are correlated in the major phrase boundaries (Ferreira, 1993; Grosjean & Collins, 1979; Selkirk, 1980, 1984).

**Structural degradation.** If on line prosody reflects the output of early structural analysis, eliminating or reducing

structural information should impair readers' ability to produce distinct prosody. Indeed, the results presented in Figure 3 suggest that degradation of structural information results in less polarized and, therefore, less distinct pause patterns. The top panel in this figure presents the average pause patterns for normal core sentences representing Structures A and B. The middle and bottom panels present the same information for the partly structured and random sentences, respectively. Whereas structured sentences produced highly polarized and visibly distinct pause patterns (top panel), the removal of functors resulted in profiles that are less polarized (middle panel), consistent with the notion that function morphemes serve as cues for structure. When, in addition to the removal of functors, the content words were also scrambled, the random sentences thus

created (bottom panel) produced pause patterns that were less polarized still, suggesting that word order also serves as a cue for structure (see Müsseler et al., 2000; Roesler, Pechmann, Streb, Roeder, & Hennighausen, 1998). A simple index of polarization is the standard deviations of the pauses in the mean pause patterns. For the semantic sentences, these deviations averaged 188.5, 109.6, and 82.2 for the intact, partially structured, and random sentences, respectively. The respective values for the nonsense sentences were 173.9, 138.5, and 66.6.

In sum, the results indicate that the pause patterns produced on line clearly differentiate between sentences with different syntactic structures and are relatively insensitive to differences in content. Furthermore, the extent of differentiation of pause patterns is strongly impaired by manipulations that degrade structural integrity but is hardly affected by manipulations that destroy semantic coherence.

### GENERAL DISCUSSION

The structural view of reading has received most of its support from the letter detection task (Koriat & Greenberg, 1994). In the present investigation, additional evidence for this view was obtained by using the prosodic pattern applied by readers in reading unfamiliar text. Experiment 1 showed that natural reading prosody is produced on line. The prosody applied when reading a sentence on its first encounter was rated as equally natural as that applied to the same sentence after practice.

Experiment 2 obtained evidence that directly supports the structural precedence hypothesis: The prosodic pause pattern applied to a sentence was found to be sensitive to modifications that changed or destroyed the structure of the sentence but was largely indifferent to those that changed or destroyed its meaning. First, same-structure sentences elicited very similar pause patterns even when they differed in semantic content. Second, sentences with different syntactic structures elicited different prosodic patterns, with prosodic pauses largely corresponding to major phrase boundaries. Finally, the degree of polarization and differentiation of pause patterns was strongest for well-structured sentences but decreased as structural integrity was gradually destroyed. In contrast, destroying semantic coherence had a much more limited effect, so that pause patterns for the meaningful and the nonsense versions of the same syntactic structure were quite similar.

Taken together, these results are consistent with the structural precedence hypothesis, according to which, early in text processing, readers combine evidence from a variety of cues to construct a structural frame for the phrase or the sentence. This frame is established on line prior to complete semantic analysis and is used to guide the interpretation of individual units and their integration within an overall meaning schema. The results of Experiment 2 specifically highlight the role of such structural cues as functors and word order: The omission of functors impaired the tendency of on-line prosody to map the structure of the sentence, whereas the scrambling of the content words within the sentence destroyed this mapping entirely.

These results accord with the proposals that, early in text processing, functors serve to anchor local phrase structure (Just & Carpenter, 1987; Kimball, 1973) and word order is used to help specify the global organization of the sentence (see Tyler & Warren, 1987).

However, it appears that semantics may also make a small contribution to prosody: The effects of semantic coherence suggest an interplay between structural and semantic contributions, with semantic coherence affecting natural prosody only when structural integrity is intact, but not when it is impaired. It would seem that semantic content can be processed only within a structural frame and that when this frame is degraded, semantic analysis is also impaired. This pattern of results is consistent with the proposition that the early extraction of structure paves the way to semantic integration.

The structural precedence hypothesis in reading was inspired by frame-and-slot models proposed to account for speech production (e.g., Dell, 1986; Garrett, 1975). According to these models, the internal representation of an utterance is constructed by inserting linguistic items into slots defined by an independent structural frame created earlier. Supporting evidence for these models comes from the study of substitution errors, structural perseveration, and structural priming. This evidence was taken to suggest the existence of an intermediate stage in which syntactic forms run free of meaning (Bock, 1990). Koriat and Greenberg (e.g., 1994) proposed that a similar model, stressing early and independent structural analysis, might apply to text processing as well. The results reviewed earlier with the letter detection task and the present results with prosody suggest that, indeed, frame-and-slot models might be extended to reading.

It might be argued that reading is more similar to speech comprehension than to speech production. Indeed, the idea of structural precedence has been invoked in connection with speech comprehension as well (Forster & Ryder, 1971). Several models assume that, at an early perceptual stage, listeners monitor the speech stream for structural cues that can help them formulate an initial hypothesis about the structural organization of the sentence before analyzing it for meaning (Bever, 1970; Bever, Lackner, & Kirk, 1969; J. A. Fodor & Garrett, 1967; J. A. Fodor, Garrett, & Bever, 1968; Halle & Stevens, 1964). Prosodic information, such as stress patterns (Kelly, 1992), pause patterns (Gee & Grosjean, 1983), intonation, and pre-boundary lengthening (Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992) serves as a major source for structural cues. Results also suggest that prosodic information can be used on line to resolve potential structural ambiguities (Marslen-Wilson, Tyler, Warren, Grenier, & Lee, 1992; but see also Watt & Murray, 1996). Furthermore, ERP monitoring of the time course of prosodic processing suggests that intonational phrasing guides the initial analysis of sentence structure (Steinhauer, Alter, & Friederici, 1999). Thus, although the exact mapping between prosodic and syntactic structures has yet to be determined (see Ferreira, 1993; Grosjean et al., 1979; Nespor & Vogel, 1987; Selkirk, 1984), research clearly suggests

a close correspondence between them. Indeed, Kelly (1992) proposed that speakers and listeners share implicit knowledge about the correspondence between syntactic and prosodic organization, which allows prosody to be used for communicating the structure of the utterance.

On the whole, the experimental evidence can be seen to support the proposition that the processing of both spoken and written messages recapitulates the general architecture of speech production, where a rudimentary frame-and-slot organization of the sentence is established before content units are inserted (see Koriat & Greenberg, 1994). In reading, however, prosody is not available as a cue to structure. Rather, readers must generate the natural prosody of the sentence on the basis of various structural cues, and they seem to succeed in doing so on line. We submit that in reading, as in speech production, in parallel to the process of establishing phrase structure, a prosodic pattern emerges that roughly captures that structure. This proposal is consistent with Beach's (1991) view that syntactic representations and prototypic patterns of prosody are stored in memory in a way that affords mutual activation.

What is the function of reading prosody? Reading prosody can be seen to constitute an integral part of the phonological representation of text. Patterson and Coltheart (1987) argue that this representation might be a parasitic carryover from speech, which assists text comprehension by virtue of its compatibility with the phonological, speech-based codes used in short-term memory (Baddeley, Vallar, & Wilson, 1987; Shankweiler & Crain, 1986). Indeed, in speech comprehension, prosody is assumed to help the listener by transforming the serial auditory input into structured patterns roughly corresponding to the syntactic segmentation of the sentence (Cutler et al., 1997). Presumably, these structured patterns facilitate the maintenance of information in working memory while the utterance is being processed (see Leonard, 1973). Our speculation is that in reading, too, prosody plays an important role as a means of representing the output of the early structural analysis and in maintaining that output in working memory while semantic analysis is carried out (see also J. D. Fodor, 1998).

## REFERENCES

- BADDELEY, A., VALLAR, G., & WILSON, B. (1987). Sentence comprehension and phonological memory: Some neuropsychological evidence. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 509-529). Hillsdale, NJ: Erlbaum.
- BEACH, C. M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory & Language*, **30**, 644-663.
- BEVER, T. G. (1970). The cognitive basis for linguistic structure. In J. R. Hayes (Ed.), *Cognitive development of language* (pp. 279-362). New York: Wiley.
- BEVER, T. G., LACKNER, J. R., & KIRK, R. (1969). The underlying structures of sentences are the primary units of immediate speech processing. *Perception & Psychophysics*, **5**, 225-234.
- BOCK, K. (1990). Structure in language: Creating form in talk. *American Psychologist*, **45**, 1221-1236.
- CARROLL, L. (1900). *Alice's adventures in wonderland and through the looking glass*. New York: Grosset & Dunlap.
- COOPER, E. W., & PACCIA-COOPER, J. (1980). *Syntax and speech*. Boston: Harvard University Press.
- CUTLER, A., DAHAN, D., & VAN DONSELAAR, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language & Speech*, **40**, 141-201.
- DELL, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, **93**, 283-321.
- DREWNOWSKI, A., & HEALY, A. F. (1977). Detection errors on *the and*: Evidence for reading units larger than the word. *Memory & Cognition*, **5**, 636-647.
- FERREIRA, F. (1993). Creation of prosody during sentence production. *Psychological Review*, **100**, 233-253.
- FODOR, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research*, **27**, 285-319.
- FODOR, J. A., & GARRETT, M. (1967). Some syntactic determinants of sentential complexity. *Perception & Psychophysics*, **2**, 289-296.
- FODOR, J. A., GARRETT, M., & BEVER, T. G. (1968). Some syntactic determinants of sentential complexity, II. Verb structure. *Perception & Psychophysics*, **3**, 453-461.
- FORSTER, K. I., & RYDER, L. A. (1971). Perceiving the structure and meaning of sentences. *Journal of Verbal Learning & Verbal Behavior*, **10**, 285-296.
- GARRETT, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9, pp. 133-177). New York: Academic Press.
- GARRETT, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production: Vol. 1. Speech and talk* (pp. 177-220). Cambridge: Academic Press.
- GEE, J. P., & GROSJEAN, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, **15**, 411-458.
- GOLDMAN-ESLER, F. (1972). Pauses, clauses, sentences. *Language & Speech*, **15**, 103-113.
- GREENBERG, S. N., & KORIAT, A. (1991). The missing-letter effect for common function words depends on their linguistic function in the phrase. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 1051-1061.
- GROSJEAN, F., & COLLINS, M. (1979). Breathing, pausing and reading. *Phonetica*, **36**, 98-114.
- GROSJEAN, F., GROSJEAN, L., & LANE, H. (1979). The patterns of silence: Performance structures in sentence production. *Cognitive Psychology*, **11**, 58-81.
- HABER, R. N., & SCHINDLER, R. M. (1981). Error in proofreading: Evidence of syntactic control of letter processing? *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 573-579.
- HALLE, M., & STEVENS, K. S. (1964). Speech recognition: A model and a program for research. In J. A. Fodor & J. J. Katz (Eds.), *The structure of language: Readings in the philosophy of language*. Englewood Cliffs, NJ: Prentice-Hall.
- HEALY, A. F. (1994). Letter detection: A window to unitization and other cognitive processes in reading text. *Psychonomic Bulletin & Review*, **1**, 333-344.
- HEALY, A. F., & DREWNOWSKI, A. (1983). Investigating the boundaries of reading units: Letter detection in misspelled words. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 413-426.
- JUST, M. A., & CARPENTER, P. A. (1987). *The psychology of reading and language comprehension*. Boston: Allyn & Bacon.
- KELLY, M. H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review*, **99**, 349-364.
- KIMBALL, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, **2**, 15-47.
- KORIAT, A., & GREENBERG, S. N. (1991). Syntactic control of letter detection: Evidence from English and Hebrew nonwords. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 1035-1050.
- KORIAT, A., & GREENBERG, S. N. (1994). The extraction of phrase structure during reading: Evidence from letter detection errors. *Psychonomic Bulletin & Review*, **1**, 345-356.
- KORIAT, A., GREENBERG, S. N., & GOLDSHMID, Y. (1991). The missing-letter effect in Hebrew: Word frequency or word function? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 66-80.
- LEONARD, L. B. (1973). The roll of intonation in the recall of various linguistic stimuli. *Language & Speech*, **16**, 327-335.

- LEVELT, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- MARSLÉN-WILSON, W. D., TYLER, L. K., WARREN, P., GRENIER, P., & LEE, C. S. (1992). Prosodic effects in minimal attachment. *Quarterly Journal of Experimental Psychology*, **45A**, 73-87.
- MARTIN, J. G. (1970). On judging pauses in spontaneous speech. *Journal of Verbal Learning & Verbal Behavior*, **9**, 75-78.
- MORAVCSIK, J. E., & HEALY, A. F. (1995). Effect of meaning on letter detection. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 82-95.
- MORAVCSIK, J. E., & HEALY, A. F. (1998). Effect of syntactic role and syntactic prominence on letter detection. *Psychonomic Bulletin & Review*, **5**, 96-100.
- MÜSSELER, J., KORIAT, A., & NIBLEIN, M. (2000). Letter-detection patterns in German: A window to the early extraction of sentential structure during reading. *Memory & Cognition*, **28**, 993-1003.
- NESPOR, M., & VOGEL, I. (1987). *Prosodic phonology*. Dordrecht: Foris.
- PATTERSON, K., & COLTHEART, V. (1987). Phonological processes in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 421-447). Hillsdale, NJ: Erlbaum.
- RAYNER, K., & POLLATSEK, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice-Hall.
- ROESLER, F., PECHMANN, T., STREB, J., ROEDER, B., & HENNIGHAUSEN, E. (1998). Parsing of sentences in a language with varying word order: Word-by-word variations of processing demands are revealed by event-related brain potentials. *Journal of Memory & Language*, **38**, 150-176.
- SAINT-AUBIN, J., & POIRIER, M. (1997). The influence of word function in the missing-letter effect: Further evidence from French. *Memory & Cognition*, **25**, 666-676.
- SELKIRK, E. O. (1980). *On prosodic structure and its relation to syntactic structure*. Bloomington: Indiana University Linguistic Club.
- SELKIRK, E. O. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
- SHANKWEILER, D., & CRAIN, S. (1986). Language mechanisms and reading disorder: A modular approach. *Cognition*, **24**, 139-168.
- SIEGEL, S. (1965). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.
- STEINHAEUER, K., ALTER, K., & FRIEDERICI, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, **2**, 191-196.
- TYLER, L. K., & WARREN, P. (1987). Local and global structure in spoken language comprehension. *Journal of Memory & Language*, **26**, 638-657.
- WATT, S. M., & MURRAY, W. S. (1996). Prosodic form and parsing commitments. *Journal of Psycholinguistic Research*, **25**, 291-318.
- WIGHTMAN, C. W., SHATTUCK-HUFNAGEL, S., OSTENDORF, M., & PRICE, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, **91**, 1707-1717.

## NOTE

1. Kendall's coefficient of concordance,  $W$ , expresses the ratio between the maximum variance possible between different patterns and the variance actually found between these patterns.  $W$  is distributed from 1.0, for maximum concordance, to 0.0, for no concordance. This statistic is actually a linear transformation of the average correlation between patterns. Unlike other measures of correlation,  $W$  can be applied to more than two vectors (see Siegel, 1965, pp. 229-239).

(Manuscript received April 27, 2000;  
revision accepted for publication October 7, 2001.)