Journal of Experimental Psychology:
Learning, Memory, and Cognition
1996, Vol. 22, No. 2, 324–335

0278-7393/96/$3.00

Effects of Lexical Frequency and Syntactic Complexity in Spoken-Language Comprehension: Evidence From the Auditory Moving-Window Technique

Fernanda Ferreira, John M. Henderson, Michael D. Anes, Phillip A. Weeks, Jr., and David K. McFarlane
Michigan State University

In 2 experiments, a new technique called the *auditory moving window* was used to investigate aspects of spoken-language processing. Participants paced their way through spoken sentences divided into word or wordlike segments, and their processing time for each segment was recorded. The 1st experiment demonstrated that high-frequency words in spoken sentences require less time to process than do low-frequency words. The 2nd experiment demonstrated that words in syntactically demanding contexts (i.e., the disambiguating word of so-called *garden-path sentences*) are processed longer than the same words in syntactically simpler contexts. Helpful prosodic information appeared to facilitate reanalysis of garden-path structures but did not seem to prevent the misanalysis. The implications of these findings for issues in spoken-language comprehension are discussed. The authors conclude that the auditory moving-window technique provides a useful tool for addressing largely unexplored issues in spoken-language comprehension.

Most of the useful linguistic information that children receive during the several years preceding literacy is obtained through the auditory modality, and even as adults, the majority of our linguistic experience comes from spoken language. Given its predominance, one might expect that the study of how spoken language is understood would be far more advanced than the study of reading. However, this is clearly not so. Little is known about the processing of spoken language, particularly that of connected speech (i.e., phrases, sentences, and texts), in contrast to the large body of work on the processing of visual language. One reason for this state of affairs is that researchers have fewer adequate tasks for examining on-line language processing in the auditory domain. In this article, we present a new on-line task we have developed for studying spoken-language comprehension. We then describe the results of two experiments with this task. These experiments provide important new information about spoken-language comprehension, demonstrating that listeners are immediately sensitive to both lexical frequency and syntactic complexity during the auditory processing of coherent sentences.

To see the need for a new task like the one we have developed, we briefly consider the contrasting situation in the field of reading. The availability of the eye-movement monitoring technique has led to a large body of data concerning basic reading processes. For example, researchers have conducted experiments in which a high- or low-frequency word is placed in a semantically appropriate sentence frame and then the location and duration of eye fixations are measured. Such experiments have shown that readers fixate longer on rarer words (Henderson & Ferreira, 1990; Inhoff & Rayner, 1986; Just & Carpenter, 1980; Rayner & Duffy, 1986). This result is important, not only because it sheds light on the properties of the visual-language processing system (indicating that the lexical-processing module is frequency sensitive), but also because it demonstrates that the lexical-frequency effects observed when words are visually presented in isolation carry over to the more typical circumstances of connected text.

To take another example, this time from the domain of syntactic parsing, numerous studies have used a methodology in which participants read what are termed *garden-path sentences* (Altmann, Garnham, & Dennis, 1992; Britt, Perfetti, Garrod, & Rayner, 1992; Ferreira & Clifton, 1986; Ferreira & Henderson, 1990, 1993; Frazier & Rayner, 1982; Rayner, Carlson, & Frazier, 1983; Trueswell, Tanenhaus, & Kello, 1993). These sentences contain a temporary ambiguity (e.g., *The editor played the tape*) that is ultimately resolved toward a less preferred structure, as in *The editor played the tape agreed the story was important.* (We ignore for the moment whether the preference is due to lexical, syntactic, or contextual factors.) Again, the eye-movement monitoring technique has been invaluable to researchers, providing a moment-by-moment profile of processing load across the sentences. This research has shown that the main verb (e.g., *agreed*) of such sentences is the site of relatively long fixation times and regressive eye movements. These data are often taken to indicate that readers interpret the sequence of words *the editor played the tape* as a main clause and then encounter processing difficulty when they must reanalyze the string as a determiner, noun, and reduced relative clause (Ferreira & Clifton, 1986;

Fernanda Ferreira, John M. Henderson, Michael D. Anes, Phillip A. Weeks, Jr., and David K. McFarlane, Department of Psychology, Michigan State University.

This research was supported by an All-University Research Initiation Grant from Michigan State University and by Grant SBR-9319272 from the National Science Foundation.

Correspondence concerning this article should be addressed to Fernanda Ferreira, Department of Psychology, Michigan State University, East Lansing, Michigan 48824-1117. Electronic mail may be sent via Internet to fernanda.ferreira@ssc.msu.edu.

MacDonald, Just, & Carpenter, 1992; Rayner et al., 1983; Rayner, Garrod, & Perfetti, 1992).

Many researchers do not have access to an eye-movement monitoring system. Fortunately, the moving-window task (Just, Carpenter, & Woolley, 1982) is much less costly and appears to be reasonably sensitive to processing load across a visually presented sentence.[1] This task works as follows: A trial begins with every letter of a sentence concealed but indicated by a position marker. A participant begins by pushing a pacing button to reveal the sentence's first word. Once the participant has comprehended the word, he or she pushes the pacing button again. This button press simultaneously conceals that word and reveals the next one. The participant proceeds in this manner until the end of the sentence. Some subset of the sentences is typically followed by a question to ensure that participants read for comprehension. The moving-window task has been used extensively to explore syntactic processing (Ferreira & Clifton, 1986; Ferreira & Henderson, 1990; MacDonald, 1994; MacDonald et al., 1992; Taraban & McClelland, 1988; Trueswell & Tanenhaus, 1991). Although the task is not as sensitive or as unobtrusive as eye-movement monitoring, the overall pattern of data obtained with the moving-window technique is often similar to that obtained from eye movements (Ferreira & Henderson, 1990, 1991).

In the auditory domain, the tasks that have been available for studying spoken-language comprehension include monitoring tasks, cross-modal lexical decision, and word naming. Monitoring tasks require listeners to push a key on detection of a particular phoneme (e.g., Cairns & Hsu, 1980; Foss, 1969) or word (Marslen-Wilson & Tyler, 1980; Tyler & Warren, 1987) in a sentence. The logic is that as processing load changes across a sentence, resource allocation to the comprehension task changes as well. Thus, at a point of comprehension difficulty, many resources are devoted to comprehension, so fewer resources are available for the monitoring task (Foss, 1969). As a result, detection times are slow. A similar logic underlies tasks requiring participants to either name or make a lexical decision to an unrelated word during sentence comprehension (e.g., Clifton, Frazier, & Connine, 1984)—times should be long when the decision must be made during processing of a difficult word in a sentence. At least three assumptions underlie this logic: First, comprehension draws on a single pool of resources; second, the size of that pool does not change with changing processing demands; and third, the demands of the two tasks are additive so that performance on the two tasks is independent. Navon and Gopher (1979) and Friedman and Polson (1981) point out problems with each of these assumptions. (See Ferreira & Anes, 1994, for a summary of these arguments.)

A somewhat different auditory technique from the ones described thus far involves having participants name or make a lexical decision to a word related to a critical word within a spoken sentence. This task has been successfully used to explore how moved constituents in a sentence are related to their original syntactic positions (e.g., Nicol & Swinney, 1989; but see McKoon, Ratcliff, & Ward, 1994, for arguments that the findings may be artifactual). However, it is not clear that it is as well-suited to examining the influence of variables such as lexical frequency and syntactic complexity on ease of comprehension. The task measures the extent to which a concept is available at a particular point in processing; it is less clear that it measures how difficult the activation of that concept was. Thus, this task does not directly measure processing load. Yet another variation is the following: A sentence is auditorily presented to participants up to some critical word. The critical word is then visually presented on a computer monitor, and participants must name the word. For example, the participant might hear *The editor played the tape*, and then the word *agreed* would be presented on the monitor. Naming time is thought to reflect how easily the word can be integrated into the sentence (Boland, 1993; Marslen-Wilson, Tyler, Warren, & Grenier, 1992; Trueswell et al., 1993). A word that continues a sentence toward an unpreferred syntactic reading, then, could be expected to elicit longer naming times. This task has yielded some interesting data, but it has the drawback that it mixes visual and auditory processing and draws on both the comprehension and production systems. As a result, one cannot be confident that the task taps into spoken-sentence comprehension alone.

Finally, all the techniques for studying spoken language discussed to this point have two additional weaknesses. First, they distract the participants' attention away from the task that is ultimately of interest—namely, language comprehension. Second, none of the tasks can provide a profile of processing load across a sentence. In contrast, the eye-movement monitoring and moving-window techniques used for studying reading do not suffer from these weaknesses. In these visual paradigms, the participants' sole task is to read the sentences for comprehension, and reading times can be obtained for every word of the sentence. In addition, these visual tasks suffer from none of the other weaknesses that specifically affect the various auditory tasks described above.

On the other hand, it is also important to note that the auditory techniques that we have described have many strengths as well, and a great deal of important work has been published with them. In addition, the *auditory moving-window technique* we describe in this article has weaknesses, some of which we describe. We do not wish to argue, then, that all existing auditory techniques should be replaced with our own. Instead, we would argue that progress in the area of spoken-language comprehension will be achieved through the development of numerous paradigms and tasks, so that converging evidence can be provided for models and theories of spoken-language comprehension.

Therefore, we believe that researchers working in the field of language processing would find useful an auditory analogue of the eye-movement monitoring and moving-window techniques. This technique will help researchers address crucial questions concerning the on-line processing of spoken sentences. We focus on three questions in this article: First, are

[1] There is some confusion in the literature about the meaning of the term *moving window*. The term was coined originally by McConkie and Rayner (1975), who developed a technique in which a small window of text was made visible to the reader contingent on eye position. The term is also used to refer to the technique developed by Just, Carpenter, and Woolley (1982) that is described in the remainder of this paragraph.

lexical-frequency effects observed in spoken sentences? Second, do spoken sentences such as *The editor played the tape agreed the story was important* lead to garden-path effects, or does the prosody of the sentence cue the appropriate structure to the listener? Third, are frequency and garden-path effects observed immediately on encountering the relevant word in the sentence, or are the effects delayed? The two experiments we describe here begin to provide some answers to these questions.

We have labeled our new task the auditory moving window.[2] The task allows participants to listen to a sentence one word at a time (or in units of whatever size the researcher wishes to use) by pressing a pacing button to receive successive words. Times between button presses—*interresponse times* (*IRTs*)—are then recorded. We tested the task in two experiments. Experiment 1 was designed to examine lexical-frequency effects, and Experiment 2 was designed to examine syntactic garden-path effects. The results of Experiment 1 indicate that lexical-frequency effects are observed in spoken sentences. This finding is important, because the existence of frequency effects is still controversial in the literature on spoken-language comprehension. In Experiment 2, participants listened to garden-path sentences such as *The editor played the tape agreed the story was important* or to their structurally preferred active counterparts (*The editor played the tape and agreed the story was important*). Studies with both the eye-movement monitoring and moving-window techniques have demonstrated longer reading times on the second verb *agreed* for the garden-path sentences compared with the active versions. Our experiment with spoken sentences shows a remarkably similar pattern. This experiment thus demonstrates that syntactic-parsing preferences observed with visually presented materials carry over to spoken sentences, at least in the circumstances of our experiment.

The Auditory Moving-Window Technique

In general, setting up the auditory moving-window technique requires the following steps: Sentences are recorded, or synthesized, digitized, and stored as waveform files, or both. A marker (which we will refer to as a tag) is placed in the waveform at the locations defining the boundaries of presentation segments. For example, Sentence 1 might be tagged as follows: *The editor* ∧ *played* ∧ *the tape* ∧ *agreed* ∧ *the story* ∧ *was* ∧ *important* (∧ indicates a tag). Whenever possible, tags are placed in the waveform at areas of low signal amplitude, as indicated by auditory and visual inspection, to make segment-to-segment transitions smooth. However, in natural speech, word boundaries often do not coincide with areas of low signal amplitude. In these cases, we attempt to position the tags so as to maximize the intelligibility of the affected words. As data from a rating task that we report later indicated, occasionally this procedure results in words that cannot be easily identified without the word following. However, we constructed our stimuli so that critical words were always intelligible on their own.

The resulting waveform files are then made available to software for running reaction time experiments. (Notice that this technique does not require the use of any form of analogue audiotape.) When the participant pushes a button, the time of the button press is recorded, and the waveform up to the first tag is converted from digital to analogue format and played out on headphones or speakers. When the participant has understood the material, he or she pushes the button again. The button-press time is again recorded, and the material between the first and second tags is played out. Button presses occurring before the end of a segment cause truncation of the segment. This truncation discourages participants from pressing the button before they have actually heard and processed the speech segment. IRTs are then computed, which indicates the time required to play out and process each segment of the sentence.

Next, we describe the specific details of the technique. This description provides information regarding the apparatus used in the two experiments and also allows other researchers to duplicate our paradigm precisely, should they wish to do so.

Apparatus

To create stimuli for our experiments, we used Computerized Speech Laboratory (CSL version 4.0 from Kay Elemetrics)—a system for digitizing and modifying speech. This system includes software for manipulation and analysis of waveforms and a circuit board for the analogue to digital and digital to analogue conversion. The system was controlled by a 486-66 MS-DOS computer equipped with an additional digital input–output I/O board and button box for gathering IRTs.

Stimuli

The procedure for creating stimuli was as follows. A male speaker spoke each of the sentences to be used in a given experiment into the microphone. Each sentence was captured by CSL at a sampling rate of 10 kHz and was edited as required for the demands of the experiment. In addition, a 500 ms, 500 Hz sine wave tone was appended to the waveform of every sentence immediately following the offset of visible and auditory activity associated with the sentence-final word. During pilot testing of the auditory moving-window technique, we discovered that participants often could not tell when a spoken sentence had ended. It appears that, although there are prosodic variations associated with the end of a sentence, they are not as compelling or as unambiguous as a punctuation mark such as a period. The purpose of the tone, then, was to provide the listener with a clear signal indicating the end of the sentence so that the listener would not continue to push the button fruitlessly and would be prepared to receive the comprehension question.

Procedure

The names of the waveform files were organized into a list. The program controlling the experiment provided a unique

[2] Marschark (1979) and Pynte (1978) have described a similar technique.

random order of the files for each participant. When the program encountered a file name, it loaded the waveform file from the hard drive into memory and placed a message on the screen for the participant to push a pacing button when ready to begin the trial. Each successive button press then played the next presentation segment. The last segment ended with a tone that served as a warning to the participant that a yes–no comprehension question would be presented next over the speaker. The entire question was played as a single unit. The participant responded to the question by pushing a button on the button panel in front of him or her. The participant continued in this fashion until all the stimuli had been encountered. The program stored the data from each participant in a file, including all IRTs, question-answering times (defined from the end of the sentence to the onset of the participant's button press), and question-answering accuracy.

Data Analysis

As described above, the program automatically stored IRTs. For many experiments, this measure will properly reflect processing times for the segments of the sentences. However, IRTs include the duration of the segment as well as any extra time required by the participant to comprehend it. This is not a problem when the critical segments in the sentence are identical in duration across conditions. However, if the segments differ, as they do, for example, in Experiment 1 (because of the presence of a high- vs. a low-frequency word), a confound could arise if the word predicted to be processed more slowly also happens to have a longer duration. For example, it has been reported that frequent words have shorter durations (Wright, 1979; but see Geffen & Luszcz, 1983), so longer IRTs for less frequent words could be due either to their longer duration or to the greater amount of time necessary to comprehend them. Therefore, we introduced another dependent measure, which we term *difference time* (DT). This measure is the difference between the IRT and the segment's tag-to-tag duration, the latter of which we refer to simply as segment duration. Figure 1 shows the IRTs and DTs of a participant in Experiment 1 (as well as the segment durations).

Experiment 1

The first experiment had two purposes. One was to evaluate the viability of the auditory moving window. Participants listened to sentences that differed in the frequency of a critical word. These sentences had been visually presented to participants in an eye-tracking study by Henderson and Ferreira (1990), who found that gaze durations and total reading times were longer on the low-frequency words. Here, we asked whether a similar processing effect would be found with spoken sentences presented with the auditory moving window.

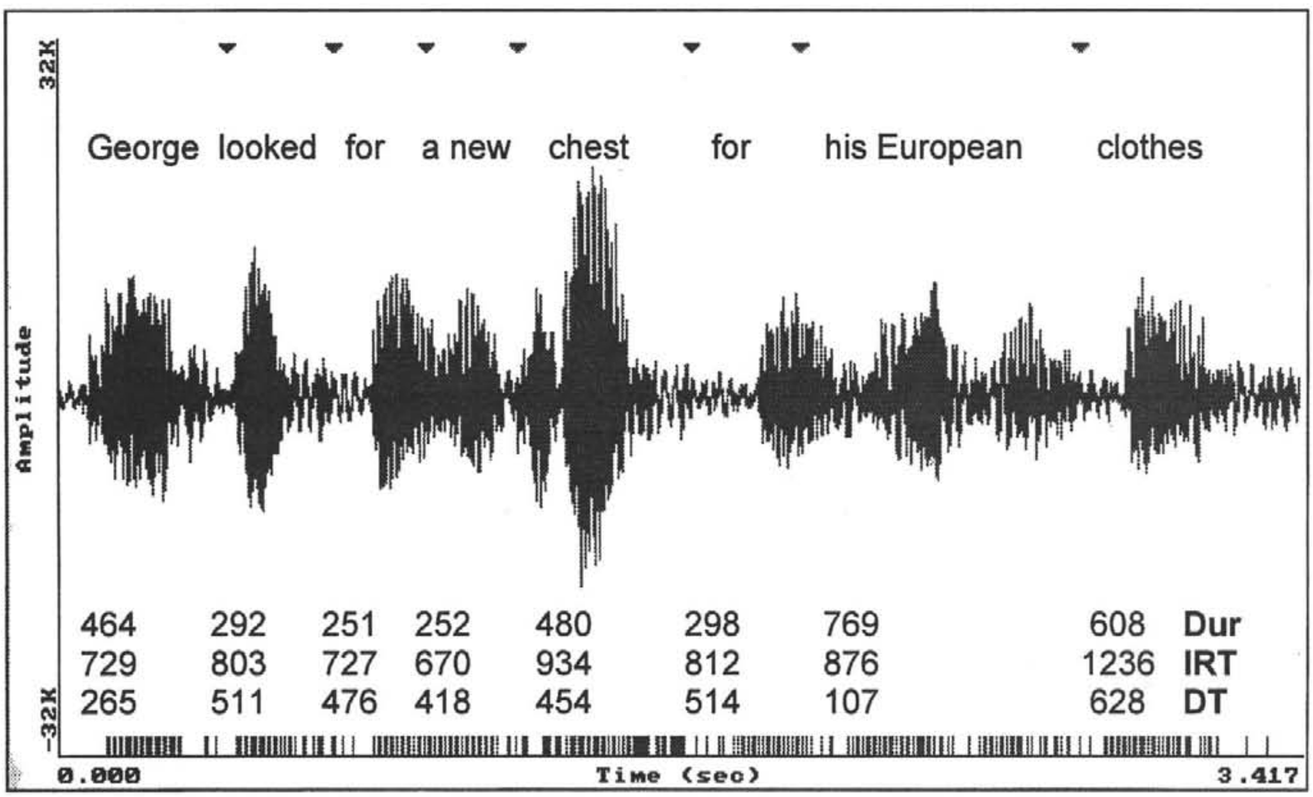

Figure 1. Waveform for a sample sentence used in Experiment 1, showing segment durations (Dur), interresponse times (IRT), and difference times (DT).

Would IRTs and especially DTs differ for the two types of words?

The second purpose of this experiment was to examine the substantive issue of the nature and extent of lexical-frequency effects in auditory-language processing. For words presented outside of any sentential context, frequency appears to influence participants' performance with auditorily presented stimuli in a number of tasks. For example, participants make lexical decisions more quickly for high-frequency words (Bradley & Forster, 1987; Eimas, Hornstein, & Payton, 1990), although the effect may be attenuated for multisyllabic words (Blosfeld's study, as cited in Bradley & Forster, 1987). Syntactic-category decisions are also made more quickly for high-frequency words (Eimas et al., 1990). Connine, Titone, and Wang (1993) presented listeners with words containing an ambiguous initial phoneme and asked participants to label the ambiguous phoneme. Participants tended to choose the phoneme that resulted in the stimulus forming a high-frequency rather than a low-frequency word. Goldinger, Luce, and Pisoni (1989) used a phonological-priming paradigm and found that the lower the frequency of a prime, the lower the accuracy of identifying a following target. Furthermore, two of the most influential models of auditory lexical processing, the cohort model (Marslen-Wilson, 1987) and the TRACE activation model (McClelland & Elman, 1986), assume that frequency affects the degree of activation of lexical candidates. Thus, according to these models, frequency influences the initial stages of lexical access. However, most of the experimental work appears to support a selection-based postaccess account (e.g., Connine et al., 1993; Goldinger et al., 1989).

The experiment reported here does not address the issue of whether frequency effects occur early or late in spoken-word recognition. Instead, the question we focus on is whether frequency influences the processing of words presented in a sentential context. The work that has been done on this question up to the present has used the phoneme-monitoring task because it has been considered the best technique available for studying auditory-sentence processing. Foss (1969) found that detection times for the first phoneme of a word were lengthened when the immediately preceding word was of low rather than high frequency, but Foss and Blank (1980) found that detection times on the word itself were not affected. Eimas and Nygaard (1992) examined the latter issue in a number of experiments and, consistent with Foss and Blank, found that frequency did not affect the time to detect the initial phoneme of a word. The only exception to this result was obtained in an experiment in which scrambled sentences were used. Thus, it appears that no study has shown effects of frequency on the word being processed in a spoken sentence. Instead, the frequency effect appears on the following word. This pattern is quite sensible if one assumes a largely bottom-up model of auditory word recognition (as argued by Eimas & Nygaard, 1992): Frequency could not affect the time to detect the initial phoneme of a word because the word has not yet been recognized. Nevertheless, it is striking that there is not yet a demonstration that the processing of an auditorily presented word in a sentential context is affected by its own frequency.

Method

Participants. The participants were 32 undergraduates from Michigan State University who participated in the experiment in exchange for partial credit in their introductory psychology courses. All participants were native speakers of American English, had normal hearing, and were not aware of the purposes of the experiment.

Materials. We used 50 pairs of sentences for this experiment. The two members of a pair differed only in the lexical frequency of a critical word approximately in the middle of the sentence. For example, one pair was *Mary bought a chest* (high frequency) or *trunk* (low frequency) *despite the high price.* The critical words were controlled for number of syllables. Of the 50 pairs of sentences, 32 were taken from Henderson and Ferreira's (1990) study. In that study, Henderson and Ferreira found that the high-frequency words received reliably shorter gaze durations than the low-frequency words did. We created 18 additional sentences by using the same critical words but creating new sentence frames for them. For example, we created the frame *George looked for a new chest (trunk) for his European clothes.* The sentences were organized into a list so that participants heard only one sentence frame in each of the two frequency conditions and did not hear a given critical word more than once during the experiment.

Experimental sentences were spoken by Michael D. Anes, who was judged to speak with an accent most similar to that of the participants who participated in the experiment. The sentences were digitized at a 10-kHz rate. The order in which the two versions of a single sentence pair were recorded was counterbalanced: For the first frame, the sentence with the high-frequency target was recorded first (e.g., *George looked for a new chest for his European clothes*), followed by the low-frequency target (e.g., *George looked for a new trunk for his European clothes*); for the second sentence pair, the frame with the low-frequency target was recorded first, followed by the high-frequency target, and so on for the remaining 48 sentence frames. For half of the experimental sentence pairs, the original recording of the frame containing the low-frequency target word was used for the low-frequency condition, and the high-frequency target word was spliced into this recording (replacing the low-frequency target word) to create the high-frequency condition. In the other half of the sentence pairs, the original recording of the frame containing the high-frequency target word was used for the high-frequency condition, and the low-frequency target word was spliced into this recording (replacing the high-frequency target word) to create the low-frequency condition. Thus, the sentence frames were exactly the same for a given sentence pair, with only the high- or low-frequency target word inserted. After splicing, tags were inserted in the waveforms so that each word of a sentence constituted its own segment, except for determiners and articles, which were always presented with the following word (typically an adjective or noun).

Yes–no comprehension questions were constructed and recorded for one quarter of the sentences. The unsegmented questions were presented, following the tone indicating the end of the sentence, and participants answered the questions by pushing one of two buttons on a button panel in front of them. Questions were virtually always answered correctly, and so they will not be discussed further. In addition to the experimental sentences, 32 filler sentences were recorded, digitized, and tagged in the same manner as the experimental sentences were. The filler sentences included a variety of syntactic structures and were of variable length.

Procedure. Participants were seated in a small room with the experimenter. They were told that they would be participating in an experiment in which they would listen to sentences and occasionally answer yes–no comprehension questions about them. They were told to pace their way through the sentences one segment at a time until they heard a tone signaling the end of the sentence. Participants were asked to read quickly but not to sacrifice comprehension, because their question-answering accuracy would be recorded. After the experimenter answered any questions from the participant, 12 practice trials were given to familiarize the participant with the procedure. The experimental session then began and lasted approximately 25 min. (Further details of the procedure can be found under the section The Auditory Moving-Window Technique above.)

Results

As described above, two dependent measures were used: IRTs and DTs. Data were analyzed with both subjects (F_1) and items (F_2) as random effects.

Results are shown in Table 1. With IRT as the dependent measure, high-frequency words took 27 ms less time to process than did low-frequency words (895 ms vs. 922 ms), $F_1(1, 31) = 9.43$, $MSE = 1{,}245$, $p < .05$; $F_2(1, 49) = 4.17$, $MSE = 3{,}801$, $p < .05$. The effect tended to spill over to the next word (824 ms and 843 ms for the word following the high- and low-frequency words, respectively) but was significant only by subjects, $F_1(1, 31) = 6.89$, $MSE = 807$, $p < .05$; $F_2(1, 49) = 1.2$, $MSE = 4{,}881$, *ns*.

As mentioned above, one possible concern about these results is that the IRT measure includes both the time necessary to present the segment as well as whatever time is needed to process it. Therefore, if high-frequency words have a shorter duration than low-frequency words do, it is possible that the differences in IRTs are due entirely to differences in the durations of the critical word. To examine this possibility, we measured the durations of the critical words used in this experiment. High-frequency words had an average duration of 538 ms, with a range of 333 to 791 ms and a standard deviation of 13.3. Low-frequency words had an average duration of 545 ms, with a range of 403 to 691 ms and a standard deviation of 10.1. The 7-ms mean difference was not reliable ($p > .20$). It should be noted that previous work has not shown a tendency for word durations to differ on the basis of their frequencies. Wright (1979) claimed that low-frequency words were 17% to 24% longer than high-frequency words were, but he did not separate word durations from the durations of any preceding pauses. Geffen and Luszcz (1983) separated the two and demonstrated that the difference found by Wright was almost entirely accounted for by pause duration.

Although our own analysis of the critical word durations did not show a reliable effect of frequency, we nevertheless conducted a second set of analyses with DT rather than IRT as our dependent measure. Recall that this measure is obtained by subtracting segment durations from IRTs and so offers one type of control for the length of the segment. In addition, the DT measure can take either negative or positive values: positive if the participant required extra time beyond the duration of the segment to comprehend that segment and negative if the participant comprehended the segment before its offset. (Recall also that presentation of the segment was truncated after the participant pushed the pacing button, so participants soon learned that they should not push the pace button unless they were confident they had comprehended the segment.) One might expect negative values if, for example, the word's uniqueness point—the point at which a word becomes unique from all others in the mental lexicon (Frauenfelder, Segui, & Dijkstra, 1990; Marslen-Wilson, 1987)—comes before its end or if the contextual constraint from the sentence together with some bottom-up input is sufficient to allow the participant to identify the word.

The results with the DT measure were consistent with the results obtained with IRTs: Participants needed 22 ms less time to comprehend a high-frequency word than to comprehend a low-frequency word (356 vs. 378 ms), $F_1(1, 31) = 6.50$, $MSE = 1{,}143$, $p < .05$; $F_2(1, 49) = 3.02$, $MSE = 9{,}488$, $p < .10$. The effect again appeared to spill over to the next word (417 vs. 435 ms for the word following the high- and low-frequency words, respectively). However, this spillover effect was significant only by subjects, $F_1\ (1, 31) = 6.25$, $MSE = 804$, $p < .05$; $F_2\ (1, 49) = 1.11$, $MSE = 5{,}610$, $p > .25$.

Finally, Marslen-Wilson (1987) has argued that any study of frequency effects with spoken words must take into account the words' uniqueness points. Therefore, we analyzed the high- and low-frequency words in this experiment to see whether their uniqueness points differed. Uniqueness points were identified with the procedure described by Frauenfelder et al. (1990). The high-frequency words contained 5.1 phonemes on average, and the uniqueness point came 4.5 phonemes into the word. The low-frequency words contained 5.3 phonemes, and the uniqueness point came 4.8 phonemes into the word. Neither the number of phonemes nor the uniqueness point differed for the two sets of stimuli ($p > .25$ for both). This finding is not surprising, for as Marslen-Wilson pointed out, most monosyllabic words do not become unique until the end of the word, and most of the critical words in this experiment were monosyllabic. Nevertheless, it is reassuring that the frequency effect obtained in this experiment does not appear to be attributable to differences in the uniqueness points of the high- and low-frequency words.

Table 1
Mean Processing Times (in Milliseconds) for Experiment 1

	Critical word		Next segment	
Measure	HF	LF	HF	LF
IRT	895	922	824	843
DT	356	378	417	435

Note. HF = high frequency; LF = low frequency; IRT = interresponse time; DT = difference time.

Discussion

This experiment makes two important points. First, it appears that frequency influences the time needed to process a word even in a spoken sentential context. Although the auditory moving window and the tasks used to study visual-language processing differ on many important dimensions, we compared the size of the frequency difference obtained in this experiment with that obtained by Henderson and Ferreira (1990), who used the same set of critical words. Henderson and Ferreira found gaze durations of 239 ms and 267 ms for the high- and low-frequency words, respectively—a difference of 28 ms. In this experiment, the difference with IRTs as the dependent measure was 27 ms, and the difference with DTs as the dependent measure was 22 ms. Thus, not only does frequency appear to affect the processing of spoken words in sentences, but the effect also seems to be of a similar size. Of course, this experiment cannot resolve the issue of whether frequency influences the initial activation of lexical candidates or only the selection of a candidate for integration into an ongoing semantic representation—our frequency effect is compatible with either possibility. What we have shown is that frequency has an influence at some point during processing of the currently heard word in spoken-language comprehension.

The second important point this experiment makes is that the auditory moving-window technique is sensitive to one variable that has been widely shown to influence processing time: lexical frequency. Indeed, as we noted above, the difference in processing time for the two types of words is remarkably similar in this experiment and in Henderson and Ferreira's (1990) experiment. Thus, we have some evidence that the auditory moving window could profitably be used as a technique for examining issues of word processing during spoken-language understanding. In the next experiment, we address whether the technique can be used to study syntactic processing and, more specifically, the comprehension of auditorily presented garden-path sentences.

Experiment 2

Many readers find visually presented sentences such as *The editor played the tape agreed the story was important* difficult to comprehend (Ferreira & Clifton, 1986; Just & Carpenter, 1992; MacDonald et al., 1992; Rayner et al., 1983, 1992). More specifically, studies in which participants' eye movements are monitored have tended to show that participants have little difficulty with the string *the editor played the tape* but that on receipt of *agreed,* participants make long fixations and execute regressive eye movements to earlier portions of the sentence (Ferreira & Clifton, 1986; Rayner et al., 1983, 1992). This pattern of data is taken to indicate that readers initially analyze *the editor played the tape* as a main clause and then must reanalyze the sentence to obtain the reduced relative reading of *played the tape* and to arrive at the structure in which *the editor played the tape* constitutes the subject of *agreed.* Recently, there has been much discussion about why such sentences are initially misanalyzed. Some researchers have argued that the misanalysis occurs because the human sentence-processing system initially relies on general-purpose parsing algorithms such as minimal attachment and late closure, according to which words should be incorporated into the ongoing phrase-structure tree with as minimal syntactic structure as possible and should be attached to the most recently postulated syntactic constituent (Frazier & Rayner, 1982). Others have argued that the pattern is due to the syntactic preferences of the lexical items used in such sentences. At the extreme, a sentence such as *The editor shown the tape agreed the story was important* would not be misanalyzed, because the lexical form *shown* is unambiguously past participle and is, therefore, unambiguously associated with the ultimately correct syntactic structure. This latter group of researchers argues that the need for syntactic reanalysis can be obviated in sentences composed of lexical items with the appropriate sorts of syntactic preferences (MacDonald, 1994; MacDonald, Pearlmutter, & Seidenberg, 1994; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993; Trueswell & Tanenhaus, 1991; Trueswell et al., 1993).

Experiment 2 does not decide between these two general theories of garden-path effects in sentence comprehension. Our goal is to assume the garden-path effect for visually presented versions of garden-path sentences and to explore whether processing load varies across the words of such sentences when they are auditorily presented in the same way as has been observed for visually presented sentences. To that end, we used two sets of materials for this study: The first set is made up of 16 sentences used by Ferreira and Clifton (1986). Ferreira and Clifton found reliable garden-path effects in their experiments with both the eye-movement monitoring and visual moving-window paradigms. The second set is made up of 15 sentences used by Rayner et al. (1992), which also showed reliable garden-path effects using eye-movement monitoring. Again, it is possible to remain agnostic about the source of these effects in the reading experiments and to merely use their existence as a baseline from which to explore whether comparable results can be found in the auditory modality.

A number of researchers have speculated that many putatively garden-pathing sentences would not cause the comprehender difficulty if they were presented auditorily (e.g., Beach, 1991). One important reason for this prediction is that spoken sentences contain rich cues to syntactic structure in the form of prosodic information. Cooper, in his pioneering work on the production of prosody (summarized in Cooper & Paccia-Cooper, 1980), showed that words are lengthened and that pauses often occur in the vicinity of a major syntactic boundary. Ferreira (1991, 1993) demonstrated that prosodic information is not a perfect predictor of syntactic structure but that the patterns are regular enough to be potentially useful to the comprehension system. However, it is not yet known what types of prosodic cues are produced by speakers, how reliably these cues are produced, or how the occurrence of such cues is influenced by the presence of other types of disambiguating information. In addition, it is possible that individual speakers vary in the extent to which they enrich their spoken sentences with prosodic cues and that sentences with structural ambiguities vary in the extent to which they can be disambiguated prosodically (Wales & Toner, 1979; Warren, 1985). Therefore, it is by no means self-evident that auditorily presented garden-path sentences do not occasion syntactic misanalysis.

In this experiment, we chose to examine the reduced relative–active main-clause ambiguity exemplified by *the editor played the tape agreed the story was important.* We chose this structure because with visual presentation it has produced robust garden-path effects. In addition, our informal observations and pilot testing indicated that these sentences were difficult to comprehend even when presented auditorily. Many instructors of psycholinguistics have had the experience of presenting the classic sentence *the horse raced past the barn fell*

(Bever, 1970) in a spoken lecture to students and finding that only a small minority were able to obtain the correct interpretation, regardless of what prosodic cues were offered. In contrast, contextual cues seem much more compelling (as shown by Ferreira & Clifton, 1986; Spivey-Knowlton et al., 1993; Trueswell & Tanenhaus, 1991). Clearly this observation does not provide evidence that the reduced relative–active main-clause ambiguity is indeed difficult to process in the auditory domain; it is given simply as an explanation of why we chose to examine this particular structural ambiguity rather than some other.

In summary, this experiment examined the differences in processing between reduced relative and active sentences presented auditorily. These sentences were produced in a natural fashion by one of the authors and so could be expected to have had the normal prosody associated with the two alternative structures. To examine whether the prosody of the sentence (as produced by this particular speaker) affected processing of the sentence, we introduced a second variable into the experiment: Sentences were presented either with no prosodic modifications or with what we term "mismatched prosody." In the mismatched prosody condition, the prosodic variations appropriate for the active structure were presented with the reduced relative condition and vice versa. The purpose of this manipulation was to see whether the sentences would be difficult to process with inappropriate prosody and whether any effects of mismatched prosody would be greater for the reduced relative structure than for the active structure—a finding that might be expected given the greater difficulty of the former structure.

Method

Participants. A group of 32 undergraduates from Michigan State University participated in the experiment in exchange for partial credit in their introductory psychology courses. All participants were native speakers of American English, had normal hearing, were not aware of the purposes of the experiment, and had not participated in Experiment 1.

Materials. The stimuli were 32 sets of sentences. The first 16 were Sentences 1–16 from Ferreira and Clifton (1986). The two versions of the sentences (active main clause vs. reduced relative) were taken directly from their stimuli, in which the active main-clause version of a sentence was simply the reduced relative version with a conjunction (typically *and* or *but*) inserted before the sentence's second verb. For example, the active sentence *the editor played the tape and agreed the story was big* is the same as the reduced relative version except for the word *and.*

The second set of stimuli was adapted from Rayner et al. (1992). In Rayner et al.'s study, the active and reduced relative versions differed substantially in their lexical content following the ambiguous portion of the sentences. For example, Rayner et al. used *the model perched on the dais and June took the picture* as the active version and *the model perched on the dais lost her balance* as the reduced relative version. Because we preferred to equate the two versions of the sentences to the greatest extent possible, we selected just one of these versions and then modified it to produce the two conditions of the experiment. For example, we selected the second version and inserted *but* to create its active counterpart (and replaced *dais* with *couch* because of the unfamiliarity of the former word to most participants), resulting in *the model perched on the couch but lost her balance.* No attempt was made to counterbalance the selection of Rayner et al.'s versions for use in this experiment. We simply selected the version we intuitively felt would result in the best stimuli for the two conditions. Finally, we created one more sentence to bring the number of stimuli in this experiment to 32 (viz., *the famous actor presented the award [and] started to cry*).

These sentences were recorded and digitized by Michael D. Anes, as in Experiment 1. For the first sentence, he recorded the reduced relative version first, and then he recorded the active version. For the second sentence, he switched the order in which the two versions were recorded and so on through the list of 32 sentences. It is important to note that the author who recorded the sentences was obviously not blind to the conditions of the experiment. Indeed, he was aware of the differences between the two versions and tried to realize those differences prosodically. We acknowledge the possibility that this speaker's intuitions about how to prosodically phrase the sentences may have been biased, inadequate, or idiosyncratic. Indeed, we maintain that research needs to be done to discover exactly how these sentences are normally produced by naive speakers as well as the extent of individual variation in the production of any prosodic cues.

The mismatched prosody conditions were created as follows: First, to create sentences with the reduced relative syntactic form but with active prosody, the active sentences were edited by removing the conjunction immediately before the second verb. For example, *and* was removed from the sentence *The editor played the tape and agreed the story was big.* Second, to create the sentence with the active syntactic form but with reduced relative prosody, the reduced relative sentences were edited by inserting immediately before the second verb the conjunction removed from the corresponding active sentences. For example, the *and* from the sentence above was inserted into the reduced relative version *The editor played the tape agreed the story was big.* The resulting sentences were free from any distortions or disfluencies that sometimes are associated with such editing procedures. Indeed, as shown below, the sentences with mismatched prosody were actually comprehended somewhat more easily than the sentences with matched prosody were.

The sentences were then tagged so that determiners were combined with the following word and all other words were presented alone. One example of a tagged sentence is *the model* $\wedge$ *perched* $\wedge$ *on* $\wedge$ *the couch* $\wedge$ *but* $\wedge$ *lost* $\wedge$ *her balance;* the reduced relative version was the same but without the word *but.* Tags in the waveform demarcated precisely the same presentation segments for the reduced relative sentences created by the deletion of the conjunctions from the original active sentences. Waveform tags for the active sentences that were created by inserting the conjunctions demarcated precisely the same presentation segments as the original reduced relative sentences had.

An additional 50 filler sentences were recorded, digitized, and tagged in the same manner as the experimental sentences were. These filler items were simple declarative sentences. As in Experiment 1, a tone was appended to all stimuli. Yes–no comprehension questions were created and recorded for one quarter of the sentences.

Procedure. The procedure was identical to that of Experiment 1.

Results

Results are shown in Table 2. The critical words in this experiment were the second verbs of the sentences in all four experimental conditions (e.g., *agreed* in *The editor played the*

Table 2
Mean Processing Times (in Milliseconds) for Experiment 2

Prosody	Critical word		Next segment	
	Act	RR	Act	RR
Interresponse times				
Matched	872	958	885	904
Mismatched	868	941	879	940
Difference times				
Matched	423	490	489	502
Mismatched	399	493	480	539

Note. Act = active sentence; RR = reduced-relative sentence.

tape [and] agreed the story was big).[3] With IRTs as the dependent measure, the critical word took 80 ms longer to process when it occurred in the reduced relative structure (950 ms) than when it occurred in the active structure (870 ms), $F_1(1, 31) = 26.76$, $MSE = 7{,}603$, $p < .01$; $F_2(1, 31) = 49.11$, $MSE = 4{,}148$, $p < .01$. There was neither a main effect of prosody nor an interaction of prosody with structure in the critical region (all Fs < 1). The main effect of structure spilled over to the next segment (typically an article plus noun combination): IRTs were 922 ms in the reduced relative condition and 882 ms in the active condition, $F_1(1, 31) = 7.80$, $MSE = 6{,}542, p < .01$; $F_2(1, 31) = 7.58$, $MSE = 6{,}747, p < .01$. This effect of structure on the segment following the critical segment was mediated by prosody, $F_1(1, 31) = 6.53$, $MSE = 2{,}235, p < .05$; $F_2(1, 31) = 4.58$, $MSE = 3{,}180, p < .05$, for the Structure × Prosody interaction. For the active sentences, there was no difference between the two prosody conditions ($p > .50$), but for the reduced relative condition, IRTs were faster in the matched prosody condition than in the mismatched condition ($p < .05$).

With DTs as the dependent measure, the critical word took 81 ms longer to process in the reduced relative structure (492 ms) than in the active structure (411 ms), $F_1(1, 31) = 24.99$, $MSE = 8{,}279, p < .01$; $F_2(1, 31) = 54.10$, $MSE = 3{,}981, p < .01$. No other effects were reliable on this segment (Fs ≤ 1). On the following segment, the reduced relative structure still took 36 ms longer to process than did the active structure (520 ms vs. 484 ms), $F_1(1, 31) = 5.27$, $MSE = 7{,}854, p < .05$; $F_2(1, 31) = 7.87$, $MSE = 6{,}730$, $p < .01$. There was also an interaction between the structure and prosody variables, $F_1(1, 31) = 10.45$, $MSE = 1{,}577, p < .01$; $F_2(1, 31) = 3.52$, $MSE = 4{,}320, p = .07$. The pattern was the same as the one observed for IRTs: For the active structures, the prosody variable had no effect ($p > .50$), but for the reduced relatives, the matched condition was easier than the mismatched condition ($p < .05$).

Discussion

In this experiment, we demonstrated that reduced relative sentences are more difficult to comprehend than their active counterparts are, even when presented auditorily. Furthermore, the effect appears on the critical, disambiguating word of the sentence—the same word on which garden-path effects have been localized in past research conducted with visually presented materials. Thus, it does not appear that this ambiguity can always be resolved with prosodic information. At the same time, the interaction of prosody and structure on the word following the critical word suggests that prosodic information can be used to aid in the reanalysis of the sentence.

These results do not rule out the possibility that prosody may be able to affect the initial processing of garden-path sentences under some circumstances. First, we examined only one structure, the reduced relative–active main-clause ambiguity. It is quite likely that other ambiguities are more susceptible to prosodic disambiguation. Indeed, according to constraint-based models of parsing (MacDonald et al., 1994; Trueswell et al., 1993), the extent to which any constraint, including prosody, can affect activation levels of alternative structures depends on the baseline preference for a particular structure. Perhaps the preference for the active structure in these sentences is so strong that it overpowers any possible influence of prosodic constraints.

Second, the way we implemented the auditory moving window in this particular experiment may have distorted the prosody of the sentences and so have weakened any potential prosodic effects. For instance, the segmentation of the sentences into wordlike chunks (where each segment would almost always be followed by a pause of some duration) could make it difficult for the normal prosody of the sentence to come through. However, it is important to note that an effect of prosody emerged in the experiment (albeit one word after it could have occurred), and so it does not appear that the task is entirely insensitive to prosody. Nevertheless, it would be useful to replicate these experiments with different segmentation for the sentences. For example, the entire ambiguous portion of each sentence could be presented as one segment (e.g., *the editor played the tape*) so that the prosodic features of that sequence would not be disrupted, and then effects on the processing time for the next word (the disambiguating word) could be examined. This implementation would make the auditory moving window more like the cross-modal naming and phoneme-monitoring techniques, in that the ambiguous portion of the sentence would be presented intact, and then the participant would perform a special task on the critical word.

Third, our stimuli were produced by just one speaker who was not blind to the experimental conditions or to the theoretical issues at stake in this work. Thus, it is possible that other speakers would provide prosodic cues for these sentences that would be useful to comprehenders in anticipating the ultimately correct syntactic form of the sentences. However, work we have conducted recently in our laboratory casts some doubt on this latter hypothesis. We asked eight undergraduates from Michigan State University who were drawn from introductory psychology courses to produce (read) a

[3] In some studies (e.g., Ferreira & Clifton, 1986, Experiment 1), the critical word in these reduced relative structures is identified as occurring earlier in the sentences. For example, in *The evidence examined by the lawyer turned out to be unreliable*, the *by*-phrase disambiguates the sentence before the second verb. However, the stimuli used here did not have the properties that would have led to this earlier disambiguation (e.g., an agentive *by* phrase following the first verb).

variety of reduced relative and active sentences, and the sentences occurred either in isolation or in an appropriately biased discourse context. We found little difference between the prosodic features of the reduced relative and active sentences, despite examining virtually every word of the sentences for differences in such acoustic correlates of prosody as word duration, pausing, and fundamental frequency variations. (Context had little effect on these measures either.) We suspect, then, that this structural ambiguity is not typically prosodically disambiguated by either sophisticated or naive speakers of English. Clearly, however, more work needs to be done on this largely unexplored issue of individual differences in the prosodic marking of sentences.

Naturalness of the Auditory Moving Window

How natural do participants find the task of listening to sentences in the auditory moving-window paradigm? Clearly, the task is somewhat unnatural—no participant had ever been asked to listen to sentences under these circumstances prior to arriving in our laboratory. But in and of itself, this is not a major criticism of any technique—few participants ever perform lexical decisions or read under conditions typically used in eye-movement laboratories unless specifically asked to do so by someone conducting a psychological experiment. Even a task as useful as the visual moving-window technique distorts normal reading in many ways. For example, participants are denied parafoveal preview and are unable to make regressive eye movements. The important question, we would argue, is whether the auditory moving-window task is sensitive to the sorts of effects that have been shown to influence ease of language processing in other experiments with more established techniques. The answer based on the results we described above appears to be yes.

Nevertheless, we asked participants during the debriefing session of each experiment to respond to four questions concerning their impression of the task and the stimuli. Participants responded to each question by circling a number between 1 and 5 on a scoring sheet. The first question was "How natural did the sentences in this experiment sound to you?" Participants were told that a response of 1 meant that the sentences were very unnatural and that a response of 5 meant that they were completely natural. In Experiment 1, the mean value given was 3.00, with a standard deviation of .95 and a range of 1 to 5. In Experiment 2, the mean was 3.16, with a standard deviation of .88 and a range of 1 to 5. The second question was "How comfortable were you with pressing the NEXT button to receive new words?" (1 = *very uncomfortable* and 5 = *completely comfortable*). In Experiment 1, the mean response was 3.53, with a standard deviation of 1.19 and a range of 1 to 5. In Experiment 2, the mean was 3.50, with a standard deviation of 1.11 and a range of 1 to 5. The third question was "How often were words presented that you felt you didn't understand?" (1 = *rarely* and 5 = *very often*). In Experiment 1, the mean response was 2.34, with a standard deviation of 1.10 and a range of 1 to 5. In Experiment 2, the mean was 1.97, with a standard deviation of .97 and a range of 1 to 4. Finally, participants were asked "If you misunderstood some words, how often did the words become clear by the end of the sentence?" (1 = *rarely* and 5 = *very often*). In Experiment 1, the mean response was 4.28, with a standard deviation of .77 and a range of 3 to 5. In Experiment 2, the mean was 4.22, with a standard deviation of .83 and a range of 2 to 5.

These results make it clear that the task was reasonably natural for participants to perform and that participants generally found the stimuli to be intelligible. The finding from the third question, that participants occasionally encountered a word they did not understand, is not surprising: Words such as prepositions always occurred as their own segments, and such words are often intelligible only after the word following is encountered (Selkirk, 1984). One difficulty with interpreting these results is that the numbers must be taken absolutely; that is, there are no comparable published data concerning the naturalness of the other major techniques used to study language processing, such as eye-movement recording, the visual moving window, or cross-modal priming. Therefore, it is impossible to tell at this point whether the auditory moving window is more or less natural than any of these other techniques. On the basis of the responses we obtained, however, we can conclude that participants found the auditory moving-window task more natural than unnatural in an absolute sense.

General Discussion

The goal of this study was twofold: First, we wished to assess whether the auditory moving window could be a useful task for exploring spoken-language comprehension, and to that end, we presented participants with sentences that differed either in the lexical frequency of a critical word or in the extent to which the sentences could be expected to induce a garden path. Second, we attempted to provide evidence relevant to some outstanding issues in language comprehension: Does frequency affect the ease with which a word can be processed in connected spoken sentences, and are garden-path effects entirely eliminated if sentences are spoken rather than written?

With respect to the first goal of our study, we found that the task could indeed be extremely useful to researchers wishing to study spoken-language comprehension. In Experiment 1, we obtained a reliable frequency effect, and furthermore, the increase in processing time associated with lower frequency occurred on the critical word; the effects were not delayed, although they did appear to spill over to the following word to some extent. These findings are consistent with work done with visual presentation of text (e.g., Rayner & Duffy, 1986). Thus, we are confident that the auditory moving-window technique can be used to study various aspects of lexical processing. In addition, in Experiment 2 we obtained reliable garden-path effects. Again, the effects were localized on the critical word in the sentence—the word that disambiguated the preceding sequence toward the more complex syntactic structure. The effect spilled over to the following word to an even greater extent than did the frequency effect, which is again consistent with data obtained from reading studies. Thus, we have provided evidence that the auditory moving-window technique is sensitive to the effects of syntactic complexity, and so we expect that the technique could be quite useful to researchers wishing to examine the effects of syntactic variables on the processing of spoken sentences.

With respect to the second goal of our study, we have found evidence that lexical frequency affects the time required to process a word in spoken connected sentences. As discussed earlier, until the present study, there had been no work showing that the processing of an auditorily presented word in a sentential context is affected by its own frequency. On the other hand, this evidence does not resolve the question of whether frequency (in spoken or written language) affects initial lexical-access procedures or affects only selection from among lexical candidates. More sophisticated manipulations will be necessary to address this issue of the locus of frequency effects.

We also obtained evidence that garden-path effects can be found even with spoken sentences. Thus, this study rules out the most extreme view of the relationship between prosody and syntactic parsing, which is that no syntactic ambiguities remain in the spoken versions of temporarily ambiguous sentences. However, the conclusion that no syntactic ambiguity can be disambiguated prosodically may be too strong given that there is some evidence for prosodic resolution of syntactic ambiguities (Beach, 1991; Marslen-Wilson et al., 1992; but see Watt & Murray, 1993). Furthermore, we examined only one structure, the reduced relative–active main-clause ambiguity, under conditions that were not optimally designed to test the extent to which prosody is used during parsing: We did not manipulate the prosodic characteristics of the sentences parametrically nor did we establish a priori the types of prosodic cues that such sentences could be expected to possess.

To begin to explore this issue in a more rigorous manner, we would need a theory of the relation between prosody and syntax. Ideally, such a theory would specify how syntactic and semantic information influence prosodic phrasing (e.g., Ferreira, 1993; Nespor & Vogel, 1987; Selkirk, 1984). In addition, an adequate theory would have to accommodate the variability in prosodic phrasing; that is, the existence of a many-to-one relation between syntactic and semantic structures, on the one hand, and phonological structures, on the other (Selkirk, 1984). Eventually, we would want to be able to predict the circumstances in which prosodic cues such as word lengthening, pausing, or pitch variations are invariably or at least reliably present versus the circumstances in which cues are produced only occasionally or with wide variations from speaker to speaker. It is not adequate simply to rely on one's intuition that a particular sentence that is ambiguous when presented visually will not be ambiguous when presented auditorily. As a result of such work, researchers could develop sophisticated experiments investigating the role of prosody in spoken language as well as in other aspects of language comprehension. We believe that the auditory moving-window technique provides one methodology with which to conduct these experiments and thus to obtain data relevant to deciding among competing theories of the architecture of the language-processing system.

References

Altmann, G., Garnham, A., & Dennis, Y. (1992). Avoiding the garden-path: Eye movements in context. *Journal of Memory and Language, 31,* 685–712.

Beach, C. M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language, 30,* 644–663.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279–362). New York: Wiley.

Boland, J. E. (1993). The role of verb argument structure in sentence processing: Distinguishing between syntactic and semantic effects. *Journal of Psycholinguistic Research, 22,* 133–152.

Bradley, D. C., & Forster, K. I. (1987). A reader's view of listening. In U. H. Frauenfelder & L. K. Tyler (Eds.), *Spoken word recognition* (pp. 103–134). Cambridge, MA: MIT Press.

Britt, M. A., Perfetti, C. A., Garrod, S., & Rayner, K. (1992). Parsing in discourse: Context effects and their limits. *Journal of Memory and Language, 31,* 293–314.

Cairns, H. S., & Hsu, J. R. (1980). Effects of prior context on lexical access during sentence comprehension: A replication and reinterpretation. *Journal of Psycholinguistic Research, 9,* 319–326.

Clifton, C., Frazier, L., & Connine, C. (1984). Lexical expectations in sentence comprehension. *Journal of Verbal Learning and Verbal Behavior, 23,* 696–708.

Connine, C. M., Titone, D., & Wang, J. (1993). Auditory word recognition: Extrinsic and intrinsic effects of word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 81–94.

Cooper, W. E., & Paccia-Cooper, J. M. (1980). *Syntax and speech.* Cambridge, MA: Harvard University Press.

Eimas, P. D., Hornstein, S. M., & Payton, P. (1990). Attention and the role of dual codes in phoneme monitoring. *Journal of Memory and Language, 29,* 160–180.

Eimas, P. D., & Nygaard, L. C. (1992). Contextual coherence and attention in phoneme monitoring. *Journal of Memory and Language, 31,* 375–395.

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language, 30,* 210–233.

Ferreira, F. (1993). The creation of prosody during sentence production. *Psychological Review, 100,* 233–253.

Ferreira, F., & Anes, M. (1994). Why study spoken language processing? In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 33–56). San Diego, CA: Academic Press.

Ferreira, F., & Clifton, C. E. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25,* 348–368.

Ferreira, F., & Henderson, J. M. (1990). The use of verb information in syntactic parsing: A comparison of evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 555–568.

Ferreira, F., & Henderson, J. M. (1991). How is verb information used during syntactic parsing? In G. B. Simpson (Ed.), *Understanding word and sentence* (pp. 305–330). Amsterdam: Elsevier.

Ferreira, F., & Henderson, J. M. (1993). Basic reading processes during syntactic analysis and reanalysis. *Canadian Journal of Experimental Psychology: Special Issue on Reading and Language Processing, 47,* 79–98.

Foss, D. J. (1969). Decision processes during sentence comprehension: Effects of lexical item difficulty and position upon decision times. *Journal of Verbal Learning and Verbal Behavior, 8,* 457–462.

Foss, D. J., & Blank, M. A. (1980). Identifying the speech codes. *Cognitive Psychology, 12,* 1–31.
Frauenfelder, U. H., Segui, J., & Dijkstra, T. (1990). Lexical effects in phonemic processing: Facilitatory or Inhibitory? *Journal of Experimental Psychology: Human Perception and Performance, 16,* 77–91.
Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14,* 178–210.
Friedman, A., & Polson, M. C. (1981). Hemispheres as independent resource systems: Limited capacity processing and cerebral specialization. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 1031–1058.
Geffen, G., & Luszcz, M. A. (1983). Are the spoken durations of rare words longer than those of common words? *Memory & Cognition, 11,* 13–15.
Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language, 28,* 501–518.
Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 417–429.
Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40,* 431–439.
Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329–354.
Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122–149.
Just, M. A., Carpenter, P. A., & Woolley, J. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111,* 228–238.
MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes, 9,* 157–201.
MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology, 24,* 56–98.
MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101,* 676–703.
Marschark, M. E. (1979). The syntax and semantics of comprehension. In G. D. Prideaux (Ed.), *Perspectives in experimental linguistics* (pp. 52–72). Amsterdam: John Benjamins B. V.
Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. In U. H. Frauenfelder & L. K. Tyler (Eds.), *Spoken word recognition* (pp. 71–103). Cambridge, MA: MIT Press.
Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition, 8,* 1–71.
Marslen-Wilson, W. D., Tyler, L. K., Warren, P., & Grenier, P. (1992). Prosodic effects in minimal attachment. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 45(A),* 73–87.
McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1–86.
McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics, 17,* 578–586.
McKoon, G., Ratcliff, R., & Ward, G. (1994). Testing theories of language processing: An empirical investigation of the on-line lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1219–1228.
Navon, D., & Gopher, D. (1979). On the economy of the human processing system. *Psychological Review, 86,* 214–255.
Nespor, M., & Vogel, I. (1987). *Prosodic Phonology.* Dordrecht, The Netherlands: Foris.
Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research, 18,* 5–19.
Pynte, J. (1978). The intra-clausal syntactic processing of ambiguous sentences. In W. J. M. Levelt & G. B. Flores d'Arcais (Eds.), *Studies in the perception of language* (pp. 109–127). New York: Wiley.
Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically ambiguous sentences. *Journal of Verbal Learning and Verbal Behavior, 22,* 358–374.
Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14,* 191–201.
Rayner, K., Garrod, S., & Perfetti, C. A. (1992). Discourse influences during parsing are delayed. *Cognition, 45,* 109–139.
Selkirk, E. O. (1984). *Phonology and syntax: The relation between sound and structure.* Cambridge, MA: MIT Press.
Spivey-Knowlton, M. J., Trueswell, J. C., & Tanenhaus, M. K. (1993). Context effects in parsing reduced relative clauses. *Canadian Journal of Psychology, 47,* 276–309.
Taraban, R., & McClelland, J. L. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. *Journal of Memory and Language, 27,* 597–632.
Trueswell, J. C., & Tanenhaus, M. K. (1991). Tense, temporal context, and syntactic ambiguity resolution. *Language & Cognitive Processes, 6,* 303–338.
Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 528–553.
Tyler, L. K., & Warren, P. (1987). Local and global structure in spoken language comprehension. *Journal of Memory and Language, 26,* 638–657.
Wales, R., & Toner, H. (1979). Intonation and ambiguity. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing* (pp. 135–158). Hillsdale, NJ: Erlbaum.
Warren, P. (1985). *The temporal organisation and perception of speech.* Unpublished doctoral dissertation, Cambridge University, Cambridge, England.
Watt, S. M., & Murray, W. S. (1993, March). *Complement sentence parsing in the auditory modality.* Poster presented at the Sixth Annual CUNY Human Sentence Processing Conference, Amherst, MA.
Wright, C. E. (1979). Duration differences between rare and common words and their implications for the interpretation of word frequency effects. *Memory & Cognition, 7,* 411–419.

Received November 9, 1994
Revision received May 26, 1995
Accepted June 2, 1995 ■