JOURNAL OF MEMORY AND LANGUAGE 30, 210–233 (1991)

Effects of Length and Syntactic Complexity on Initiation Times for Prepared Utterances

FERNANDA FERREIRA

University of Alberta, Edmonton, Canada

In three experiments, I examined initiation times for memorized utterances. Sentences varied in phonological word length, syntactic complexity, or semantic plausibility. The experiments demonstrated that number of phonological words and syntactic complexity (as measured by the number of nodes in a phrase structure tree) affected the time it took subjects to initiate the utterance. If a sentence had a syntactically complex subject and a syntactically complex object, speakers tended to pause at the subject–verb phrase boundary, and pause duration increased with upcoming complexity, just like initiation times. Semantic plausibility had no measurable effect. I argue that these results reflect the process of translating a semantic/syntactic representation of a sentence into a sound-based structure ultimately useable by the speech apparatus. For long and complex sentences, these resource-intensive processes cannot occur over the domain of the entire sentence, and so the production system must divide the sentence into two performance units, with the boundary between them occurring at a syntactically prominent location. © 1991 Academic Press, Inc.

This study brings together two lines of research in language production. The first is the work of Sternberg, Monsell, Knoll, and Wright (1978) demonstrating that the more words in an utterance, the longer it takes speakers to initiate it. The second line of research is work demonstrating the psychological reality of syntactic structures in sentence production (Bock, 1982, 1986, 1987; Bock & Kroch, 1989; Fromkin, 1971; Garrett, 1975, 1976, 1982; Kempen & Hoenkamp, 1987; Levelt, 1989). The way in which this work will be brought together is through the demonstration that the greater the syntactic complexity of an utterance, the longer it takes a speaker to initiate it.

Sternberg et al. (1978) proposed that a short-term memory buffer is used to hold the motor program for an utterance prior to its articulation. In order to produce the utterance, the speaker must retrieve it from this buffer. Sternberg et al. predicted that the more words in the utterance, the longer the retrieval time. This prediction was tested using word lists (varying in length from one to five words) such as *Wednesday–Thursday–Friday–Saturday*. Subjects were to begin producing a list as quickly as possible upon receipt of a visual cue. In a number of experiments, Sternberg et al. found that each word in the list added about 10 ms to the initiation time. It made no difference whether the items in the list had an inherent order, as in the example above, or were randomly ordered.

In further experiments, Sternberg et al. found that the unit of the initiation time effect is not the word, but rather the stress group. Lists such as *one and two and three* took no longer to initiate than lists such as *one two three*. Each list contains the same number of stressed syllables (three), but

The first and third experiments were supported by Grant OGP-37323 from the Natural Sciences and Engineering Research Council of Canada to the author. The second experiment was conducted at the University of Massachusetts, as partial fulfillment of the requirements for the Ph.D. degree, and was supported by NIH Grant HD-18708 to Charles Clifton and Lyn Frazier. I thank John Henderson and Charles Clifton, for comments on earlier versions of this article, and Lucille Lavender and Ramona Clifton, for assistance conducting the experiments. Correspondence concerning this article and reprint requests should be addressed to Fernanda Ferreira, Department of Psychology, University of Alberta, Edmonton, AB, Canada T6G 2E9 (e-mail: usernand@ualtamts.bitnet).

0749-596X/91 $3.00

differs in number of words. The finding that these two lists were equivalent indicates that it is the number of stressed syllables that affects initiation time. Sternberg et al. proposed that the stress group (a unit consisting of a single stressed syllable and any following unstressed syllables) is the unit of production. The more of these in the motor program for an utterance, the more units that must be retrieved from short-term memory and so the longer the initiation time. This result will be termed *the length effect*. Stress groups have also been referred to as phonological words (Selkirk, 1984), and I will use the latter term for the remainder of the article.

The length effect is remarkably robust (Levelt, 1989; Monsell & Sternberg, 1981). However, it has been shown with word lists and has not been demonstrated with sentences. In order to begin to incorporate such an effect into a model of sentence production, it must be shown that the number of phonological words affects the time to initiate syntactically and semantically structured utterances. One reason that the length effect might be different with sentences than with word lists is that the structure provided by syntactic and semantic groupings may cause some words to be clustered together. As a result, the length effect may be attenuated or even eliminated.

A second question that arises from the research done by Sternberg et al. is whether other characteristics of an utterance besides its length in phonological words could affect initiation time. The experiments presented here will examine the potential effects of syntactic complexity. Consider the following sentences:

a. The river near their city empties into the bay.
b. The river that stopped flooding empties into the bay. (1)

These sentences do not differ in length, regardless of whether length is measured in number of words, syllables, or phonological words. However, they do differ in syntactic complexity. The subject of (a) is syntactically less complex than the subject of (b), because the embedded prepositional phrase in the former requires fewer syntactic nodes than the relative clause in the latter (see panels B and C of Fig. 1). Sternberg et al.'s results would lead one to expect that initiation times for these two sentences would be equivalent. On the other hand, it is possible that the more syntactic nodes in the utterance about to be produced, the longer it will take a speaker to initiate it. On this view, (b) should take longer to initiate than (a).

There is good reason to expect that syntactic structure will affect initiation time for utterances. Research on motor production has shown that the patterns of inter-response times between elements of a motor sequence can be predicted if one assumes that the entire sequence is organized into a hierarchical tree that controls the execution of the response (Gordon & Meyer, 1987; Rosenbaum, 1985; Rosenbaum, Inhoff, & Gordon, 1984; Rosenbaum, Kenny, & Derr, 1983; Sternberg, Knoll, & Turock, 1990). For example, Rosenbaum and his colleagues have argued that simple finger movements are organized into a plan having some characteristics of a phrase structure tree. This tree is used in a two-part way to control production. When the motor production system receives a "go" signal, the tree is first edited to ensure that all elements of the sequence are known. Once the edit phase is completed, execution begins. Control returns to the top (root) of the tree, but now each element is executed as it is encountered. In both the editing and execution phases, the tree is decoded through a tree traversal process (Collard & Povel, 1982). The more steps required to decode an element (the more branches that must be traversed to reach that element), the longer the latency to produce it. Although it is certainly a large conceptual leap from the production of finger movements to the production of sentences, the research on motor planning at least suggests that a hierarchical tree may be used to control the production of complex motor sequences. The language production system may exploit the hierarchical structure found in sentences for the purpose of organizing and controlling the production of an utterance. It is important to bear in mind, however, that language is organized into hierarchical structures of different types—semantic, syntactic, and phonological. Any or all of these could be used to guide production.

The expectation that syntactic structure will affect utterance initiation time depends on the assumption that sentences cannot only be formally described as having a syntactic structure, but that sentence production normally incorporates a stage at which that structure is created. As Bock has pointed out (1982; Bock & Kroch, 1989), it is hard to imagine how a model of production could avoid this assumption, because speakers must produce sentences that conform to the syntactic rules of their language (Garrett, 1976, makes a similar point). For example, speakers of English must plan their utterances so that adjectives precede their nouns; speakers of French generally must produce sentences with postnominal adjectives. Because these syntactic decisions must be made, most models of production assume a syntactic stage at which the linear and precedence relations among words are established. In the speech error models proposed by Fromkin (1971) and Garrett (1975, 1982), a message level or semantic representation is organized into a surface structure. Garrett has argued that the processes that operate at this stage pay no attention to the semantic properties of the words they are organizing. This claim is based on the finding that speech errors (specifically, whole word exchanges) occur among elements that share syntactic characteristics, but are typically semantically dissimilar (although the elements are sometimes associatively or phonologically related; Dell, 1988; Dell & Reich, 1981).

Bock (1986, 1987; Bock & Kroch, 1989) has conducted a great deal of research on the computation of structure during sentence production. In one study, Bock (1986) demonstrated that the forms of two consecutive sentences have a tendency to be syntactically congruent. Subjects who just heard a passive sentence were more likely to describe an unrelated picture with a passive as well. From this result, Bock argued that the component of the sentence production system that creates structure is specialized and encapsulated from the semantic system. This study (along with a number of others Bock has conducted) demonstrates that a normal aspect of sentence production is the creation of the syntax of a sentence and that this structure cannot be reduced to its semantic content (Bock & Kroch, 1989).

Johnson (1966) conducted a study to examine the effects of syntactic variables on sentence initiation time. He compared sentences such as *The person who jumped over there is good* and *The person over there who jumped is good*. The syntactic structure of the first version is more right-branching than the structure of the second, and, so according to Yngve's (1960) complexity metric, it is less complex. Johnson predicted that because the less complex version was associated with fewer decoding operations, it could be initiated more quickly than its more complex counterpart. This prediction was supported. Johnson's sentences also differed in complexity as defined in the present article (i.e., a node count). In the first version, the prepositional phrase *over there* modifies a verb, and in the second version, the phrase modifies a noun, with the result that the second version has one more node in its phrase structure tree than the first (see Johnson, 1966, Fig. 1). The result Johnson obtained suggests that syntactic complexity can affect utterance initiation time, as measured by Yngve's metric and as measured by a node count. However, one problem with this experiment is that the more complex sentences were generally also less plausible, and either characteristic could account for the obtained initiation time differences (Fodor, Bever, and Garrett, 1974).

The purpose of the experiments presented here was to explore the following questions: First, what is the effect of the length of an utterance on initiation time when that utterance is a sentence rather than a word list? Will a greater number of words be associated with a longer initiation time, as was found in the original Sternberg et al. experiments? Second, beyond the number of words, does the number of syntactic nodes affect initiation time? The third question concerns the possible similarity between initiation times and pause durations. If a sentence is so complex that it must be divided into two production units, will the division between the units be the site of a pause, and will pause duration increase systematically in the same way that initiation times increase? The fourth question concerns the possible influence of semantic plausibility on the length effect. Will the plausibility of an utterance affect its initiation time? These questions were explored in three experiments, all employing the same paradigm. Speakers were presented with a sentence which they were asked to memorize. Upon receipt of a visually presented verbal cue, they were to begin saying the sentence as quickly as possible.

A final point that should be addressed before turning to the experiments is the link between the task used in this experiment and "real" language production. Clearly, speakers rarely produce sentences under the circumstances found in the present experiments. Nevertheless, my task may be useful for revealing the processes that do occur during spontaneous speech. Although the task I employed does not allow speakers to generate the utterance's semantic or syntactic content, it does require them to generate its phonological, phonetic, and motor representations and to execute the motor sequence. Thus, this task can give some insight into how a semantic/syntactic structure is eventually translated into a representation that can be produced by the speech apparatus. Further, the task has a nonexperimental analogue: Consider a situation in which two speakers are conversing. Speaker A has an important question to ask of Speaker B, but B has the floor. Under these circumstances, Speaker A might find him- or herself rehearsing the question until the "go" signal is given by Speaker B—that is, until B stops talking. Finally, as in many experiments, the circumstances may be somewhat artificial, but without the ability to hold some variables constant and manipulate others, it is impossible to determine which of the many different linguistic characteristics of words and sentences is influencing the processing measures.

EXPERIMENT 1

The first experiment had two purposes. The first was to assess whether the length effect obtained by Sternberg et al. (the finding that initiation time for an utterance increased with a greater number of phonological words in the utterance) would occur when the utterances were structured sentences rather than word lists. The second purpose was to determine whether the structural characteristics of an utterance would affect initiation time. Specifically, the experiment tested whether an increase in syntactic complexity (as defined by number of nodes in a phrase structure tree) would be associated with an increase in initiation time.

Method

Subjects. Subjects were 28 University of Alberta undergraduates who participated in the experiment in exchange for course credit. All subjects were native speakers of Canadian English and were not aware of the purposes of the experiment.

Materials. Twenty-four sentences appeared in one of four conditions, shown in Table 1. (A complete set of the materials appears in Appendix A.) In each case, the characteristics of the subject noun phrase

TABLE 1
EXAMPLE OF AN ITEM IN THE FOUR EXPERIMENTAL CONDITIONS, EXPERIMENT 1

Short condition
The river empties into the bay that borders the little town.

Long, low syntactic complexity condition
The large and raging river empties into the bay that borders the little town.

Long, medium syntactic complexity condition
The river near their city empties into the bay that borders the little town.

Long, high syntactic complexity condition
The river that stopped flooding empties into the bay that borders the little town.

differed. In the first condition, the subject contained only two words (an article and a noun) and was therefore labelled short. In the second, third, and fourth conditions, the subjects had more than two words and were labelled long. Across these three long conditions, the subjects had the same number of words, but differed in syntactic complexity. In the long, low syntactic complexity condition, the subject consisted of an article, one or more adjectives, and the head noun. In the long, medium complexity condition, the subject consisted of an article and a head noun, followed by an embedded prepositional phrase. In the long, high complexity condition, the subject consisted of an article, a head noun, and an embedded relative clause. Figure 1 illustrates that the increase in syntactic complexity for the three long conditions is attributable to an increase in the number of syntactic nodes in the structural representation of each phrase type.

Panel A of the figure depicts the tree for the long, low syntactic complexity condition. This tree contains eight syntactic nodes (not including the lexical items themselves, which add a constant of five to each of the three long conditions). Panel B, for the long, medium complexity condition, requires nine nodes. Panel C, for the long, high complexity condition, requires 12 nodes. These trees were created on the ba-

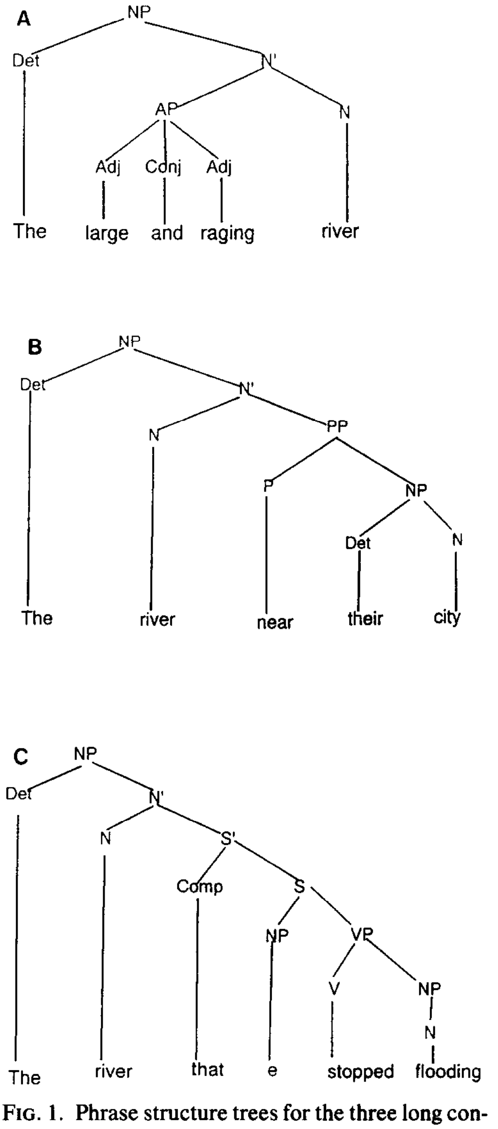

FIG. 1. Phrase structure trees for the three long conditions, Experiment 1. A illustrates the low syntactic complexity condition, B the medium complexity condition, and C the high complexity condition.

sis of two assumptions: First, that intermediate X-bar levels for phrases (Jackendoff, 1977) are not required unless they branch (Frazier, 1989); and second, that relative clauses such as the one shown in panel C are created through wh-movement, and such movement leaves behind an empty position or gap (shown in the figure as *e*) in the phrase structure (Fodor, 1989, summarizes the psychological evidence for the existence of these empty positions).

The short version of the sentences had fewer words, phonological words, content words, and syllables than the long versions. Across the three long subject conditions, sentence versions were equated on all dimensions except syntactic complexity. The three long versions had the same number of words and phonological words, content words, and syllables. In addition, the three long versions were equated on the overall frequency of the words used in the sentences.

Sentence plausibility was also assessed in a norming experiment. Twenty University of Alberta undergraduates, none of whom participated in the initiation time experiments, rated the sentences on their plausibility. Each subject was given a booklet containing the sentences. In each booklet, a sentence appeared in only one of its four conditions. As a result, there were four different booklets. Conditions were assigned to booklets using a Latin square procedure so that each subject saw all conditions of the experiment.

Subjects were asked to rate the sentences according to "how well the words of the sentence seemed to belong together or were associated with each other." Subjects rated the sentences on a scale of one to five, where a value of *one* meant that the words did not belong together at all, and *five* meant that the words belonged together perfectly. The instructions were clarified with examples. Subjects were told that a sentence such as *Martin got out a needle and some thread* might be rated quite highly; a sentence such as *Martin got out a hammer and some thread* would be rated somewhat lower, because hammers and thread are not typically associated in people's minds.

The results of the rating study are shown in Table 2. A one-way ANOVA performed on the data revealed an effect of condition, $F(3,57) = 7.81$, $MS_e = 22.2$, $p < .001$ with subjects as the random variable, and $F(3,23) = 6.04$, $MS_e = 34.5$, $p < .005$, by items. The value of 3.9 assigned to the short subject condition differed significantly from the next highest value of 3.5 assigned to the long, high complexity condition, $F(1,19) = 10.57$, $MS_e = 18.4$, $p < .005$ by subjects, and $F(1,23) = 8.68$, $MS_e = 26.9$, $p < .01$ by items. The three long conditions did not differ significantly from each other, $F(2,38) = 2.5$, $MS_e = 17.1$, $p > .05$ by subjects, and $F(2,46) = 1.35$, $MS_e = 38.5$, $p > .25$ by items. Note also that the plausibility ratings assigned by the subjects did not correlate with syntactic complexity, and so it is not possible to argue that an increase in syntactic complexity inevitably led to a decrease in the plausibility of the sentences. Thus, the results of this experiment are not vulnerable to the same criticism that has been raised against Johnson (1966).

The 24 experimental items were intermixed with 24 fillers. Each subject saw only one version of any one experimental item,

TABLE 2

PLAUSIBILITY RATINGS, INITIATION TIMES, AND MEMORIZATION TIMES, EXPERIMENT 1

Condition	Plausibility ratings[a]	Initiation times[b]	Memorization times[c]
Short	3.9	518 (.01)	7.8
Long, low complexity	3.4	550 (.02)	12.6
Long, medium complexity	3.2	633 (.04)	12.8
Long, high complexity	3.5	656 (.02)	12.8

[a] Values shown are plausibility ratings between one and five, where a sentence given a one is implausible, and a five, plausible.

[b] Values shown are in milliseconds; error rates are shown in parentheses.

[c] Values shown are in seconds.

but through a Latin square procedure, saw all conditions of the experiment. Item presentation was randomized for each subject.

Procedure. Subjects were seated in front a computer screen. A trial consisted of the following events. First, a message to begin the trial when ready appeared on the screen. After the subject pushed a button to indicate that he or she was ready, a sentence was presented. Subjects were told to take as long as necessary to memorize the sentence. They were encouraged to use whatever strategies they wished, including the use of verbal rehearsal. The experimenter stressed that it was critical to say the sentence exactly as it appeared on the screen. When the subjects felt confident they could say the sentence accurately, they pushed a button again, which caused the sentence to disappear. After a random delay varying from 500 to 1000 ms, a cue to begin speaking appeared on the screen. The random delay was used so that subjects would remain vigilant. The cue was the question, "What happened?" Subjects were told to begin saying the sentence as quickly as possible upon receipt of the cue. A Zenith 80286 microcomputer controlled the display of the items, and initiation times were gathered with a voice-activated relay interfaced with the computer.

The experimenter was situated in a different room and listened to the subjects' productions over headphones. If a subject made an error on any trial, the error was noted so that the trial could be eliminated from the data analysis.

Subjects were free to take as long as necessary between trials. Subjects were run individually, and each experimental session lasted approximately 20 min.

Results

Trials on which subjects made an error in saying the sentence were eliminated (approximately 2% of trials). Error rates across conditions are shown in Table 2. Errors included lexical substitutions, false starts, and stutterings which prevented the subject from completing the entire sentence. Means for the remaining responses were analyzed with both subjects (F_1) and items (F_2) as random effects.

Initiation times across the four conditions are shown in Table 2. The results were first analyzed as a one-way ANOVA, which revealed a significant effect of condition $F_1(3,81) = 4.69$, $MS_e = 25{,}848$, $p < .005$, $F_2(3,69) = 4.37$, $MS_e = 23{,}424$, $p < .01$.

To analyze this pattern, a number of planned comparisons were performed. The first crucial comparison is between the short subject condition and the average of the three long subject conditions. The mean for the short condition was 518 ms, and for the three long conditions overall, 613 ms. This difference was significant, $F_1(1,27) = 16.70$, $MS_e = 6342$, $p < .001$, $F_2(1,23) = 15.91$, $MS_e = 7181$, $p < .001$, and indicates that the more words in an utterance, the longer it takes to initiate its production.

Second, the long, low complexity condition was compared to the average of the medium and high complexity conditions (550 vs. 645 ms). This difference was reliable, $F_1(1,27) = 5.03$, $MS_e = 24{,}976$, $p < .05$; $F_2(1,23) = 11.85$, $MS_e = 9156$, $p < .005$, and indicates that subjects required less time to initiate production of the utterance of low syntactic complexity than one of medium or high complexity. The long, medium complexity condition and the long, high complexity condition did not differ reliably from each other (633 vs. 656 ms), both F's < 1.

The times to memorize the sentences (shown in Table 2) were also analyzed. An overall ANOVA revealed a significant effect of condition, $F_1(3,81) = 23.12$, $MS_e = 7{,}476{,}515$, $p < .001$, $F_2(3,69) = 17.75$, $MS_e = 6{,}321{,}761$, $p < .001$. The mean for the short subject condition was compared to the average of the three long subject conditions (7.8 vs. 12.8 s). This difference was significant, $F_1(1,27) = 58.88$, $MS_e =$

5,871,673, $p < .001$, $F_2(1,23) = 43.31$, $MS_e = 5,172,803$, $p < .001$. The means for the three long conditions did not differ significantly from each other, both F's < 1.

Discussion

The results of this experiment straightforwardly address two of the questions put forth in the Introduction. The first question was, what is the effect of the length of an utterance on initiation time when the utterance is a sentence rather than a word list? The answer is that the effect is the same: The greater the number of phonological words in an utterance, the longer it takes subjects to initiate it. Thus, the Sternberg et al. result is not limited to word lists, but occurs even in utterances that have internal structure. The second question was, does syntactic complexity affect initiation times, so that the greater the number of syntactic nodes in the utterance the speaker is about to produce, the longer the initiation time? A comparison of the low to the medium and high complexity conditions indicates that syntactic nodes act much like phonological words: the more of them, the longer the initiation time.

The length effect found in this experiment suggests further that the sentence production system attempts to produce a structurally defined unit—a unit that can contain varying numbers of words. To clarify this point, let us assume for the sake of argument that the unit in question corresponds to the subject noun phrase. The subject of the sentence in Table 1 consists of one phonological word in the short condition and three in each of the long conditions. The results of this experiment indicate that the sentence production system attempts to produce the subject as one unit; the more phonological words in that unit and the greater its syntactic complexity, the longer its initiation time. (The same argument holds if we assume that the unit is the whole sentence rather than the subject.) An alternative way in which the production system could work would be to buffer some constant number of words—say, three. If the system operated in this fashion, the length effect found in this experiment would not have occurred, because the system would simply have buffered three words in each of the four conditions. This experiment thus provides evidence that the production system attempts to produce units of a particular structural type, rather than arbitrary units of a particular size.

The second experiment allows for a further test of this claim. I varied the syntactic complexity of the subject and object noun phrases of subject–verb–object sentences and examined both initiation times and pauses before and after the sentence's main verb. The first experiment showed that syntactic nodes take up short-term memory capacity. By varying the syntactic complexity of the subject and object, it should be possible to create sentences that tax the production system and necessitate a pause. If the production system attempts to produce units of a constant structural type such as the subject, then pauses should tend to occur before the main verb, regardless of subject or object complexity.

EXPERIMENT 2

The second experiment was designed to address the third question of the study. If a sentence is so complex that it must be divided into two production units, will a pause occur at the boundary between them? An answer to this question would permit a test of the idea that the sentence production system places the break point within a sentence at a structurally defined location such as the subject–verb phrase boundary. In contrast, the pausing algorithm proposed by Gee and Grosjean (1983) predicts that the main pause in a sentence will shift location with changes in the complexity of pre- and postverbal material. The algorithm attempts to balance the size of the constituents on either side of a pause, and so it includes the verb with the following material if that material is simpler than the subject; the verb is included with the subject if the subject is simpler than the following material. These predictions are shown in Fig. 2.

The nodes represent prosodic units: A phi symbol (ϕ) represents a phonological phrase, and an *I* stands for an intonational phrase. The between-word location dominated by the highest *I* is associated with the highest probability of a pause. The algorithm predicts that with a complex subject and simple object, the most likely location for a pause is before the verb; with a simple subject and complex object, the most likely location is after the verb. The Gee and Grosjean algorithm is thus a specific instance of a model that is more concerned to balance the size of production units than to produce units that respect the syntactic structure of a sentence.

Further, if syntactic complexity affects pause duration just as it affects initiation times, it would indicate that pausing in speech can be caused by the difficulty of producing an upcoming segment (Butterworth, 1980; Cooper & Paccia-Cooper, 1980; Ford, 1982; Gee & Grosjean, 1983; Goldman-Eisler, 1967; Tannenbaum, Williams, & Hillier, 1965), and further, the experiment would localize the difficulty to the syntactic level. A result implicating syntax would run contrary to at least some theories of sentence production. For example, Levelt (1989) argues that because syntactic operations operate automatically, "the complexity of these syntactic operations will not be reflected in measures of mental load, such as reaction times and hesitation pauses" (p. 259). The finding that the syntactic complexity of an upcoming stretch of speech affects pause duration would indicate that even if syntactic operations are automatic, the execution of an utterance under the guidance of a syntactic structure takes more capacity the more complex that structure.

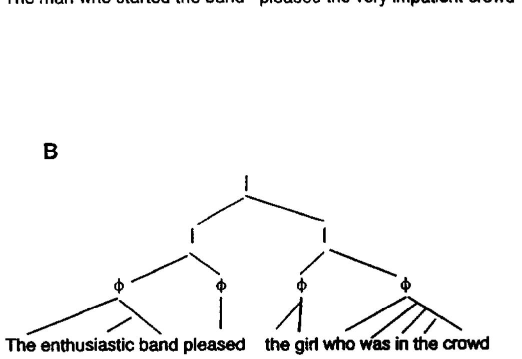

FIG. 2. Trees generated by the Gee and Grosjean (1983) algorithm. A shows the tree generated for the complex subject, simple object condition, and B the tree for the simple subject, complex object condition.

Method

Subjects. The subjects were 36 University of Massachusetts undergraduates who participated in the experiment in exchange for partial credit in their psychology courses. All subjects were native speakers of American English and were unaware of the purposes of the experiment.

Materials. Each item could appear in one of nine conditions illustrated in Table 3. (A complete set of the materials is given in Appendix B.) Subject and object syntactic complexity were independently varied. Each phrase could be either of low complexity (an article, one or more adjectives, and a noun), medium complexity (an article, a head noun, and an embedded prepositional phrase), or high complexity (an article, a head noun, and a relative clause). The complexity conditions are thus identical to those used in the first experiment.

The sentences were equated on length in the number of phonological words. Because of the number of conditions that had to be created for any one item and because the final word of the subject had to be identical across conditions for a given item and easily segmented in a waveform, it was not possible to equate items on the number of

TABLE 3
EXAMPLE OF AN ITEM IN THE NINE EXPERIMENTAL CONDITIONS, EXPERIMENT 2

Low complexity subject, low complexity object
The enthusiastic band pleased the very impatient crowd.

Medium complexity subject, low complexity object
The pianist in the band pleased the very impatient crowd.

High complexity subject, low complexity object
The man who started the band pleased the very impatient crowd.

Low complexity subject, medium complexity object
The enthusiastic band pleased the senator in the crowd.

Medium complexity subject, medium complexity object
The pianist in the band pleased the senator in the crowd.

High complexity subject, medium complexity object
The man who started the band pleased the very impatient crowd.

Low complexity subject, high complexity object
The enthusiastic band pleased the girl who was in the crowd.

Medium complexity subject, high complexity object
The pianist in the band pleased the girl who was in the crowd.

High complexity subject, high complexity object
The man who started the band pleased the girl who was in the crowd.

words as well. However, given Sternberg et al.'s (1978) finding that it is the phonological word or stress group, not the word, that is responsible for the length effect, it seemed appropriate to equate the sentences on the number of phonological words. Sentences were not equated across conditions on the frequency of use of the words in the phrases, or on plausibility. However, the results of the previous experiment suggest that syntactic complexity has an effect on initiation times beyond any contribution of frequency or plausibility. (The third experiment will also demonstrate that plausibility has little effect on initiation times.)

Twenty-seven items such as the one shown in Table 3 were constructed. The items were created so that for any one item, the same one syllable word was always used as the final word of the subject. Thus, the word before the verb, the verb, and the word after the verb (*the*) were always the same across the conditions. This constraint was imposed in order to permit accurate pause duration comparisons across conditions and from both the preverbal and postverbal locations. Each subject saw only one version of any one sentence, but through a Latin square procedure, saw all conditions of the experiment. The experimental sentences were intermixed with 36 filler items. Item presentation was randomized individually for each subject.

Procedure. The procedure was similar to that used in the first experiment. Subjects were told that their task was to memorize a target sentence, and then to produce it as quickly as possible upon receipt of a cue.

On each trial, the following events took place. First, a sentence designated as a context sentence appeared on a computer screen. Subjects were told that the sentence provided a context for the sentence they were supposed to say out loud and that they could take as long as necessary to read it. The purpose of this sentence was to aid subjects in their comprehension, and thus their memorization, of the often quite difficult target sentences. Subjects pushed a button on a button panel in front of them once they had understood the context sentence, and then it reappeared together with the target. (The context sentence reappeared to permit subjects to reread it if desired.) Subjects read and memorized the target sentence, as in the first experiment.

Once the subjects felt confident they could say the target sentence from memory, they pushed a button, causing the screen to go blank. A cue (a wh-question such as "who pleased whom") then appeared on the screen after a random delay varying from 500 to 1000 ms. The subject produced the target sentence as quickly as possible in response to the cue.

If subjects made an error in production, they repeated the trial from the beginning. This procedure was used to prevent subjects from avoiding some of the demanding sentences used in this experiment. Thus, all trials were included in the data analyses, because none were eliminated on the basis of subject error.

A microcomputer collected reading/memorization and initiation times automatically and controlled a tape recorder. When the subject pushed the button for the cue to produce the target sentence, the computer turned on the tape recorder, so that all productions of the target sentence were recorded on audio tape.

Measures. Measures taken were memorization and initiation times for the target sentences, probability of a pause before and after its main verb, and pause durations at the same locations. Pauses were measured using standard techniques outlined in Cooper and Paccia-Cooper (1980). Sentences were digitized at a 10 kHz rate and analyzed with a waveform editor. Pause duration before the verb was measured from the offset of visible activity associated with the final word of the subject to the onset of activity associated with the verb and could take any value greater than zero. Pause duration after the verb was measured similarly, from the offset of the verb to the onset of the first word of the object (the word *the*). The probability of pausing referred to the likelihood that any pause greater than zero occurred at the two critical locations.

Results

Initiation times. The initiation time results are shown in Fig. 3. Initiation times increased significantly as the syntactic complexity of the subject of the sentence increased, $F_1(2,70) = 3.72$, $MS_e = 18{,}189$, $p < .05$; $F_2(2,52) = 3.30$, $MS_e = 15{,}159$, $p < .05$. Initiation times for the sentences containing a low complexity subject were 480 ms; the medium complexity subject, 500 ms; and the high complexity subject, 529 ms. The complexity of the object had no effect, $F_1(2,70) = 1.10$, $MS_e = 18{,}455$, p

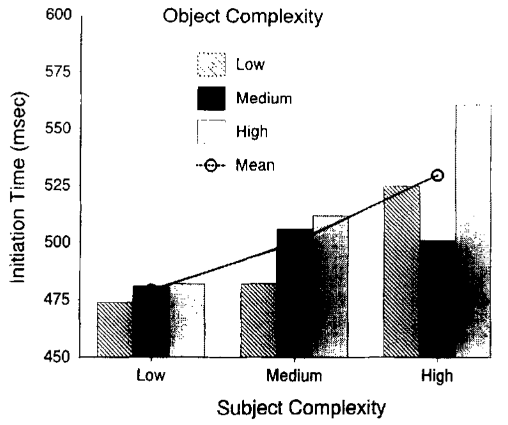

FIG. 3. Initiation times as a function of the syntactic complexity of the subject and object, Experiment 2.

$> .20$; $F_2 < 1$, and the two variables did not interact (both F's < 1).

Pause location. Pauses before and after the verb were analyzed as follows. For each subject, the likelihood of pausing before the verb and the likelihood of pausing after it across the nine conditions was assessed. Given that there were 27 sentences, a subject could pause three times before the verb and three times after it in any one condition. If more pauses occurred before than after the verb, the condition was counted as a "plus." For example, if a subject paused three out of three times before the verb and one out of three times after the verb, that condition was counted as a plus. If the opposite occurred, that condition was counted as a minus. If an equal number of pauses occurred before and after the verb, the condition was counted as zero. The result of this analysis is an indication of how many subjects paused more often before the verb than after, and how this tendency was manifested over the nine conditions.

The same procedure was then followed for each sentence. With 36 sentences and nine conditions, it was possible to pause four times before the verb and four times after it in any one condition. Pluses, minuses, and zeros were assigned as described above. The analysis over items indicates whether the majority of sentences tended to have preverbal pauses overall and where pauses tended to be located as a function of the experimental conditions.

Sign tests were performed on the data. Of the 36 subjects, two did not pause at all. Of the remaining 34 subjects, 32 paused more often before the verb than after, $z = 4.9$, $p < .001$. In addition, across all nine conditions, subjects were more likely to pause before the verb than after (the smallest z value, obtained for the condition with a complex subject and a simple object, was $z = 2.25$, $p < .01$). The same pattern held over items. For all 27 sentences, pauses occurred more often before the verb than after, $z = 5.01$, and this was true for each of the nine conditions (smallest $z = 2.00$, $p < .01$, obtained also for the complex subject–simple object condition). Thus, for virtually every subject and for every sentence, pauses tended to occur before the verb rather than after. Further, this pattern was evident in every experimental condition.

Pause probability. The probability of a pause before and after the verb was separately assessed. A 3 × 3 ANOVA at each location was performed on the data, where the first variable refers to subject syntactic complexity (low, medium, or high) and the second to object complexity (low, medium, or high).

First, as the result of the analysis on pause location would indicate, a pause was more likely to occur before the verb than after (.26 vs. .04). For the location after the verb, there was a main effect of object complexity by subjects, $F_1(2,70) = 3.67$, $MS_e = .009$, $p < .05$, but not by items, $F_2(2,52) = 1.51$, $MS_e = .015$ $p > .20$. With a low complexity object, the probability of a pause was .06; with a medium object, .03; and with a high complexity object, .04. To the extent that this effect is reliable, it indicates that the more complex the object, the less likely a pause is to occur after the verb. Subject complexity had no effect, nor was there an interaction between subject and object complexity (all F's < 1).

For the location before the verb, there was a significant effect of object complexity, $F_1(2,70) = 3.89$, $MS_e = .054$ $p < .05$, $F_2(2,52) = 3.56$, $MS_e = .046$, $p < .05$. The more complex the object, the more likely a preverbal pause. Subject complexity had no effect, nor was there an interaction between the two complexity variables (all F's < 1). The pattern of results over all conditions is shown in Fig. 4.

Pause duration. Pause duration before and after the verb was separately assessed. A 3 × 3 ANOVA was performed on the data, where the first variable was subject complexity and the second variable object complexity.

Pause durations were longer before the verb than after (110 vs. 15 ms). Further, for the location after the verb, pause durations were not affected by subject complexity, both F's < 1, nor were they affected by object complexity, $F_1(2,70) = 1.41$, $MS_e = 3708$, $p > .20$, $F_2(2,52) = 1.45$, $MS_e = 6117$, $p > .20$, and there was no interaction between these two variables, both F's < 1. Thus, pause durations after the verb are largely unaffected by the complexity manipulations.

Pause durations before the main verb of the sentence as a function of the nine complexity conditions are shown in Fig. 5. The prediction from the initiation time results is

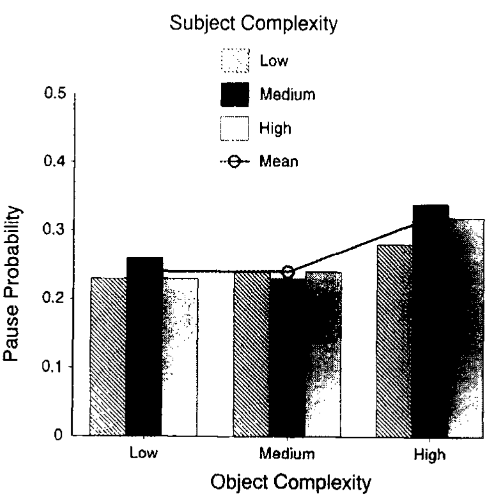

FIG. 4. Pause probabilities before the sentence's main verb as a function of the syntactic complexity of the subject and object, Experiment 2.

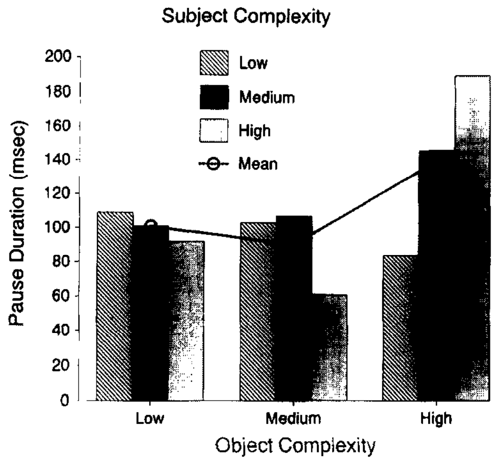

FIG. 5. Pause durations before the sentence's main verb as a function of the syntactic complexity of the subject and object, Experiment 2.

that pauses will increase in duration as a function of upcoming syntactic complexity—the complexity of the direct object of the verb. Consistent with this prediction, pause durations were affected significantly by the object, $F_1(2,70) = 3.99$, $MS_e = 18{,}347$, $p < .05$, $F_2(2,52) = 5.54$, $MS_e = 32{,}564$, $p < .01$.

Preverbal pause durations were not reliably affected by the preceding complexity of the subject noun phrase, $F_1 < 1$, $F_2(2,52) = 2.05$, $MS_e = 33{,}565$, $p > .10$. The mean for the pause following a low complexity subject was 98 ms; a medium complexity subject, 118 ms; and a high complexity subject, 115 ms.

The subject and object complexity variables interacted significantly, $F_1(4,140) = 2.68$, $MS_e = 21{,}721$, $p < .05$, $F_2(4,104) = 3.39$, $MS_e = 27{,}628$, $p < .05$, as can be seen in Fig. 5. If the subject of the sentence was of low or medium complexity, the complexity of the object had little effect on preverbal pause durations. But when the subject was most complex, pause durations were affected by the complexity of the object. Thus, object complexity seems to influence pause duration only when the earlier portion of the sentence is complex. Notice that this interaction was not obtained using probability of pausing as the dependent measure, perhaps because this measure is less sensitive than the duration measure.

Prepausal word duration. The finding that preverbal pause duration did not vary as a function of subject complexity (but did vary with upcoming complexity) is contrary to results reported by Cooper and Paccia-Cooper (1980). They found that the greater the syntactic complexity of a phrase, the longer the duration of the phrase-final word and the pause after it. Their results also show that in the environments in which a pause occurs, lengthening of the preceding word tends to occur as well; however, word duration tends to be a more stable measure. Therefore, in the present experiment, it is possible that word durations would show effects of preceding syntactic complexity not revealed by pauses.

The duration of the final word of the subject noun phrases was consequently analyzed and the data submitted to a 3 × 3 ANOVA. On the basis of Cooper's results, I would expect an effect of subject complexity, little effect of object complexity, and no interaction. In contrast, the complexity variables had no effect. With a low complexity subject, the duration of the phrase-final word was 299 ms; medium complexity subject, 304 ms; and high complexity subject, 309 ms, $F_1(2,70) = 1.30$, $MS_e = 1871$, $p > .25$; $F_2(2,52) = 1.39$, $MS_e = 1313$, $p > .25$. Object complexity had no effect, both F's < 1, nor was there an interaction, $F_1(4,140) = 1.08$, $MS_e = 1468$, $p > .30$, $F_2(4,104) = 1.05$, $MS_e = 1066$, $p > .35$. Thus, word durations were unaffected by the complexity manipulations, and preverbal pause duration varied only as a function of upcoming complexity.

Memorization times. Memorization times are shown in Table 4. Subjects' times to memorize the sentences increased with subject complexity, $F_1(2,70) = 16.36$, $MS_e = 10{,}433{,}916$, $p < .001$, $F_2(2,52) = 11.45$, $MS_e = 11{,}075{,}780$, $p < .001$, and object complexity, $F_1(2,70) = 5.46$, $MS_e = 15{,}815{,}300$, $p < .01$, $F_2(2,52) = 6.85$, $MS_e = 9{,}516{,}191$, $p < .01$. The two variables did

TABLE 4
MEMORIZATION TIMES (SECONDS), EXPERIMENT 2

Subject complexity	Object complexity		
	Low	Medium	High
Low	8008	8401	9650
Medium	10,213	10,666	11,270
High	9831	10,732	12,397

not interact, both F's < 1. These results indicate that subject and object complexity contributed independently to the memorization times, so that the more complex the constituent (either subject or object), the longer it took subjects to memorize the sentence.

Discussion

The results of this experiment replicate the syntactic complexity results found in Experiment 1. In the present experiment, initiation times increased as the syntactic complexity of the subject noun phrase increased. This result suggests again that the more syntactic nodes in the utterance a speaker is about to produce, the longer it takes subjects to initiate that utterance.

The finding that object complexity had no effect on initiation time qualifies the above statement in an important way. It is not the complexity of the entire utterance that matters; only the complexity of the subject affects initiation times. In sentences as long and/or complex as the ones used in this experiment, the subject–verb phrase boundary blocks any effects of object complexity. A sentence required a pause at the boundary when both the subject and object were syntactically complex.

This pattern is inconsistent with the predictions of the Gee and Grosjean (1983) algorithm. In the condition where the subject was syntactically complex and the object simple, the pause was predicted to occur before the verb; with a simple subject and complex object, the pause was predicted to occur after it. Instead, across all conditions, pauses were more likely and were of a longer duration before rather than after the verb. Thus, it appears that units of production respect major syntactic/semantic boundaries.[1]

A minor point is that the word and pause duration results obtained here are unexpected, given Cooper and Paccia-Cooper's (1980) theory of syntactic-to-phonetic coding. Their theory states that the more complex a phrase, the longer the duration of its final word and the following pause. Here, the durations of the subject-final word and the preverbal pause were unaffected by the complexity of the subject. (Recall that preverbal pauses were affected by object complexity.) One reason for the discrepancy may be that in Cooper and Paccia-Cooper's work, syntactic complexity and length were typically confounded—more complex phrases tended to have more phonological words in them. In Experiment 2, phonological word length remained constant while syntactic complexity varied. These null results may indicate that syntactic variations alone do not influence word duration (Ferreira, 1988; Ferreira & Clifton, 1987). Another implication is that phonetic pauses such as those studied by Cooper and Paccia-Cooper (1980) may be influenced by different variables than planning-based pauses such as the ones studied here (Ferreira, 1988).

The memorization time results from the present experiment contrast with the results obtained in Experiment 1. Here, memorization times were longer the more complex the subject or the object of the sentence, while in the first experiment, syntactic complexity had no effect on mem-

[1] In addition to the pre- and post-verbal locations, all other between-word locations were examined for pauses and revealed that pauses rarely occurred in any of these positions. Statistical treatment of the data is difficult because of the paucity of pauses and because the different between-word positions are not comparable across conditions or items. The data nevertheless strongly suggest that the most common place in a sentence for a pause to occur was the preverbal location.

orization times. A possible reason for the difference in results is that the sentences in Experiment 1 were better controlled, so that they varied only in syntactic complexity. The memorization times may have been different in Experiment 2 due to factors other than the sentence's syntactic complexity. Given that in the first experiment syntactic complexity did not lead to increased memorization times, it seems that at least some aspects of initiation time reflect factors other than the ease of learning and remembering the spoken sentences.

In summary, Experiment 2 provides answers to two of the questions given in the introduction. One question this experiment was designed to answer was whether syntactic complexity would affect initiation times. The answer appears to be that it does, just as in the first experiment. The more complex the constituent about to be produced, the longer it takes speakers to initiate the sentence. The second question this experiment was designed to answer was three-part: Could a sentence be so complex that it required a pause, and if so, where would the pause occur, and would pause duration vary in the same way as initiation times? The answer here is that long and/or complex sentences will tend to be spoken in two production units. The first unit will comprise just the subject of the sentence, and the time to initiate the sentence will vary with its complexity. The second unit will comprise the verb phrase, and the difficulty of initiating this unit will also increase systematically with its complexity.

Experiment 3

Thus far, three of the questions outlined in the introduction to the paper have been addressed. First, even in structured sentences, the more phonological words in an utterance, the longer it takes subjects to initiate it. Second, with length held constant, the more syntactic nodes in the utterance about to be produced, the longer it takes subjects to initiate it. Third, if an utterance is so complex that it must be produced as two units, a pause may occur at the syntactically most preferred location, and the duration of the pause will increase with upcoming syntactic complexity, just as initiation times increase.

The question that remains to be addressed is: What is the effect of the semantic plausibility of an utterance on initiation times? In the first experiment, an increase in syntactic complexity was not correlated with a decrease in the plausibility of a sentence, and yet initiation times varied with differing amounts of complexity. Thus, the experiment demonstrated that variations in semantic plausibility are not necessary to obtain variations in initiation times. The present experiment was conducted to examine whether variations in plausibility would be sufficient to affect initiation times. If plausibility is neither necessary nor sufficient to affect initiation times, then there is strong evidence that the processes that operate prior to the production of an utterance examine only its structural characteristics (including its length) and ignore its semantic and pragmatic content. To examine the influence of plausibility, sentences were constructed so as to vary on two dimensions: The subjects of the sentences were either short or long, and the sentences were either semantically plausible or relatively less plausible.

Method

Subjects. Twenty University of Alberta undergraduates participated in this experiment. All subjects were native speakers of Canadian English and were not aware of the purposes of the experiment.

Materials. A sentence appeared in one of four versions, illustrated below:

a. The plants were exposed to rain and frost.
b. The plants in the garden were exposed to rain and frost.
c. The cans were exposed to rain and frost.

d. The cans in the garden were exposed to rain and frost. (2)

(A complete set of the materials is shown in Appendix C.) The subject noun phrase was either short (an article and a noun) or long (an article, noun, and an embedded prepositional phrase). The sentences were either plausible (the *plants* conditions, a and b) or relatively implausible (the *cans* conditions, c and d).

At a given level of the length variable, the sentences were identical except for the head noun of the subject (*plants* vs. *cans* in the present example). A head noun for the subject was chosen so that it was either intuitively likely to occur in the context of the remainder of the sentence, or intuitively less likely. For each pair, the head nouns were matched on number of syllables and on frequency.

To obtain a more objective measure of plausibility, a norming study was conducted in the same way as the norming study described in Experiment 1. Twenty University of Alberta undergraduates who had not participated in any of the initiation time experiments or the previous norming study rated the sentences on a five-point scale, where a value of one meant that the sentence was implausible, and a value of five meant that the sentence was plausible. The same example sentences were given in the instructions as in Experiment 1.

Each subject was given a booklet in which one of the four versions of each of the 20 sentences appeared. Each item appeared in only one of its conditions in the booklet, but across items all conditions were represented. Thus, four different booklets were created, and five subjects responded to each of the four booklets.

The results for the plausibility ratings are shown in Table 5. Subjects rated the sentence versions designated "plausible" as significantly more plausible than the "implausible" versions, $F_1(1,19) = 45.46$, $MS_e = 51.67$, $p < .001$, $F_2(1,19) = 25.01$, $MS_e = 86.24$, $p < .001$. The plausible versions were given an average rating of 3.8, and the implausible versions were given an average rating of 2.7. In addition, the short sentences were rated as more plausible than the long sentences, $F_1(1,19) = 12.64$, $MS_e = 29.48$, $p < .005$, $F_2(1,19) = 7.23$, $MS_e = 51.93$, $p < .05$. The short sentences were given an average rating of 3.5, and the long sentences, 3.1.

TABLE 5

PLAUSIBILITY RATINGS, INITIATION TIMES AND ERRORS, AND MEMORIZATION TIMES, EXPERIMENT 3

	Plausible	Implausible	Mean
Plausibility ratings: 1 = implausible, 5 = plausible			
Short	4.0	2.9	3.5
Long	3.6	2.5	3.1
Mean	3.8	2.7	
Initiation times in milliseconds and proportion errors (in parentheses)			
Short	549 (.00)	583 (.01)	566 (.01)
Long	687 (.01)	645 (.02)	677 (.02)
Mean	618 (.01)	615 (.02)	
Memorization times in seconds			
Short	3.9	4.1	4.0
Long	6.3	6.6	6.4
Mean	5.1	5.4	

The 20 normed experimental sentences were used in the initiation time experiment and appeared together in a list with 24 filler sentences. Item presentation was randomized individually for each subject.

Procedure. The same procedure was employed here as in Experiment 1. The experimental session lasted approximately 15 min.

Results

Trials on which subjects made an error were eliminated from the analysis and constituted fewer than 2% of the trials. The results for initiation times, proportion errors, and memorization times are shown in Table 5.

Initiation times were shorter for the short versions (566 ms) than the long versions (667 ms). This result was significant by subjects, $F_1(1,19) = 4.63$, $MS_e = 44{,}242$, $p < .05$, and marginally significant by items, $F_2(1,19) = 2.90$, $MS_e = 78{,}755$, $p < .10$.

Initiation times were not affected by the plausibility manipulation, both F's < 1. The time to initiate the plausible version was 618 ms, and to initiate the implausible version, 615 ms. The length and plausibility factors did not interact, both F's < 1.

Memorization times were longer for the long sentences (6.4 s) than the short sentences (4.0 s), $F_1(1,19) = 65.03$, $MS_e = 1{,}959{,}882$, $p < .001$, $F_2(1,19) = 67.44$, $MS_e = 1{,}764{,}379$, $p < .001$ (as was found in Experiment 1). Memorization times were not affected by plausibility, $F_1 < 1$, $F_2(1,19) = 1.54$, $MS_e = 1{,}184{,}903$, $p > .20$. The length and plausibility factors did not interact, both F's < 1.

Discussion

The results of this experiment indicate that plausibility, at least at the levels manipulated in this experiment, does not affect utterance initiation times. Implausible sentences took no longer to initiate than plausible ones. One possible counterargument is to claim that the plausibility manipulation was not powerful enough to produce an effect. Although the implausible sentences were relatively less plausible than their counterparts, they were not strange or bizarre. A stronger manipulation, where one version of the sentences was highly plausible and the other bizarre, could affect initiation times. On the other hand, the sentences clearly differed significantly in plausibility as assessed by a norming task, and this degree of plausibility difference did not affect initiation times. The evidence available from this experiment implies that the processes operating just prior to producing an utterance are insensitive to its semantic plausibility.

General Discussion

This study was designed to bring together the work done by Sternberg et al. (1978) on the length effect and research on the production of syntactic structure. The study demonstrated that not only does the number of phonological words affect initiation times for sentences (a result which had not previously been demonstrated), but so does syntactic complexity: The more syntactic nodes in a sentence, the longer it takes subjects to initiate it. The research demonstrated further that this effect of syntactic structure cannot be attributed to semantic variables. The first and third experiments indicated that semantic plausibility is neither necessary nor sufficient to explain differences in initiation times. The sentences in the first experiment were equivalent in their plausibility but still differed in their initiation times. The sentences in the third experiment differed in plausibility, but this difference had no effect on initiation times.

The syntactic complexity effect cannot be attributed to propositional complexity either. Consider the sentences in Table 1. The short version consists of three propositions (viz., that a river empties into a bay, that the bay borders a town, and that the town is little). The long, low complexity version adds two more propositions to the short version (that the river is large and that the river is raging); the long, medium complexity version adds only one proposition to the short version (that the river is near a city); and the long, high complexity condition adds only one proposition to the short version (that the river stopped flooding). In general, the low complexity condition will add one proposition for every adjective to the short version; the medium complexity condition will add one proposition for the embedded prepositional phrase; and the high complexity condition will add one proposition for the relative clause (Kintsch, 1974). Thus, across the three complexity conditions, syntactic complexity increased while propositional complexity remained constant (or in the items with two adjectives in the low complexity condition, propositional complexity varied in the opposite direction from syntactic complexity). These experiments, then, provide evidence that the syntactic characteristics of an utterance systematically affect the processing that takes place prior to its articulation and that these characteristics are not confounded either with semantic plausibility or propositional complexity.

What are the theoretical implications of these findings for models of sentence production? We may begin answering this question by attempting to answer another, more specific, question: What takes place in the production system during the initiation or pause time? Here I will explore the possibility that what takes place is the translation of a semantic/syntactic structure into a sound-based, or phonological, representation.

For any sentence to be spoken, it must be translated into a representation that can control the speech apparatus. The subjects in the current experiments were faced with a sentence having a specified semantic and syntactic structure, but had to provide the phonological and phonetic representations themselves. My hypothesis is that this translation from a semantic/syntactic to a phonological/phonetic representation occurs during the initiation or pause interval. A representation containing syntactic constituents must be translated into one containing phonological constituents and eventually into a motor representation. This translation occurs during the initiation or pause interval. The more syntactic nodes that must be translated, the longer the translation takes and, so, the longer the initiation or pause time. The length effect may be a consequence of the same mechanism: The more words in a sentence, the more syntactic nodes it will contain (because every word must be syntactically integrated) and, so, the longer the initiation/pause time.[2] And finally, semantic plausibility has no effect on this translation process, because the plausibility of a sentence does not affect its structural complexity, and thus it does not make the translation of the sentence into a phonological representation any easier or any more difficult.

How is this translation brought about? Selkirk (1986) has proposed what she terms the "X-max algorithm." This algorithm scans a sentence's syntactic structure and converts it to a phonological representation. To explain its operation, it is necessary to clarify the term "X-max." According to the X-bar theory of phrase structure (Chomsky, 1970; Jackendoff, 1977), each lexical category (verbs, adjectives, prepositions, and nouns) projects a corresponding phrasal category. Projections can go up a couple of levels (represented as bars), culminating in what is called the maximal projection of a lexical category. The maximal projection of a verb is VP (verb phrase, equivalent to V-double-bar); of an adjective, AP (adjective phrase); of a preposition, PP (prepositional phrase); and of a noun, NP (noun phrase). A maximal projection always has the form XP; thus the term "X-max" refers to any maximal projection.

The algorithm searches for the right boundary of maximal projections in a sentence's surface structure. For example, consider the following sentence from Table 3:

$$[_{S}\,[_{NP1}\,\text{The}\,[_{A'}\,\text{enthusiastic}]_{A'}\,\text{band}]_{NP1}\;[_{VP}\,\text{pleased}\,[_{NP2}\,\text{the senator}\,[_{PP}\,\text{in}\,[_{NP3}\,\text{the crowd.}]_{NP3}\,]_{PP}\,]_{NP2}\,]_{VP}\,]_{S} \qquad (3)$$

The right boundaries of maximal projections (indicated with right brackets) can be found in two places: after *band,* and at the end of the sentence. (Selkirk stipulates that adjectives do not project all the way to AP, but only to A′ (A-single-bar), based on syntactic evidence provided by Stowell, 1981.) The algorithm takes all of the material up to the right edge of a maximal projection and converts it into a phonological phrase (abbreviated as *PPh*). In addition, the entire sequence is considered to be an utterance (abbreviated as *Utt*). Thus, the above sen-

[2] This account of the length effect does not preclude the one offered by Sternberg et al. (1978). It is quite possible that the processes that occur with structured sentences include the ones described by Sternberg et al., as well as the ones described in this article.

tence would have the following phonological representation:

$$(_{\text{Utt}}\ (_{\text{PPh}}\ \text{The enthusiastic band})_{\text{PPh}}\ (_{\text{PPh}}\ \text{pleased the senator in the crowd})_{\text{PPh}}\ )_{\text{Utt}} \qquad (4)$$

a single utterance dominating two phonological phrases.[3] Notice that this phonological theory, unlike Gee and Grosjean's (1983), is consistent with the evidence from Experiment 2 showing that pauses occur more often before the verb than after. A phonological phrase could not end after the verb, because that location does not line up with the right edge of a maximal projection.

A set of processes like the ones described in the X-max algorithm could be occurring in the sentence production system before the initiation of an utterance. The syntactic structure would be scanned left to right and translated into a phonological representation. It would be expected that the more structure to be converted (and the more words that are syntactically structured), the longer the conversion would take. In addition, there is no reason why plausibility should have an effect, because neither of the representations involved in the operation of the X-max algorithm contain information about real-world plausibility.

Thus far, this model can account for the length effect obtained with structured sentences, the effect of syntactic complexity, and the absence of any effect of semantic plausibility. What about the finding (most obviously seen in the second experiment) that a complex sentence often contains a pause in a syntactically prominent position? Recall that the translation of syntactic structure to phonological structure takes short-term memory capacity. If a sentence is long and/or structurally complex, the translation of the entire sentence will exceed this capacity. As a result, the sentence will be divided into two performance units.

One of the many remaining questions is whether this division occurs only in those sentences that contain a pause. Levelt (1989) has argued that the fluency of most speech makes it unlikely that planning can occur only when the system is not engaged in articulation. Instead, planning (including the translation of one representation into another) may occur during the execution of a previous utterance. Consider the following account. Assume that virtually every sentence used in Experiment 2 had to be spoken in two performance units. While subjects were executing the first unit, they were simultaneously preparing to produce the second—i.e., translating the syntactic constituents to phonological ones. For the easier sentences, these two processes could overlap and so the translation of the second was completed before articulation of the first unit was finished. But particularly if both the subject and object were syntactically complex, the translation of the second unit was not complete before execution of the first unit was completed. As a result, the speaker needed more time to prepare the second unit, and therefore paused.

In addition to pauses, there may be other acoustic cues to the existence of performance units. One intriguing possibility is that each performance unit is also an intonational phrase (Gee & Grosjean, 1983). One would then expect to find the typical fundamental frequency characteristics (described, for example, in Pierrehumbert, 1980) that are associated with the ends of such phrases. One strategy for future research would be to explore the way in which sentences of the type used in these experiments are intonationally phrased and whether different phrasings are associated with different processing states of the production system. We know little about the

[3] Selkirk's theory also covers the creation of phonological words, but I will not discuss them here, because they are not created from syntactic constituents. The same is true for intonational phrases: Selkirk assumes that they are created on the basis of a speaker's semantic and pragmatic interests (Selkirk, 1984) and not directly from syntactic structure. Of course, a general theory of how a sentence is converted into a phonological representation would have to account for the creation of all of these phonological constituents.

factors affecting a speaker's choice among different intonational options, but many researchers have argued that the decision is governed largely by stylistic and pragmatic concerns (e.g., Levelt, 1989; Selkirk, 1984). The current research leads to the speculation that the choice among different intonational possibilities is also affected by the processing state of the sentence production system.

In conclusion, the present study begins to address a fundamental question in sentence production: How is a sentence phonologically encoded? I have argued that the encoding takes place over a structurally defined unit, whose size is determined by the limits of short-term memory. The encoding involves translating a syntactic structure into a structure containing phonological phrases, using the X-max algorithm proposed by Selkirk (1986). The more structure to be translated, the longer the translation takes. The challenge now is to find further evidence in support of these claims and to discover the way in which other phonological constituents such as phonological words and intonational phrases are created.

APPENDIX A: EXPERIMENTAL SENTENCES USED IN EXPERIMENT 1

The first option within brackets is the short condition; the second option is the long, low syntactic complexity condition; the third option is the long, medium complexity condition; and the fourth option is the long, high complexity condition.

1. [The dog / The big and hairy dog / The dog in Dad's garden / The dog that bites children] went to the pond that's next to the museum.
2. [The customer / The tall and handsome customer / The customer at Sam's market / The customer who was shopping] asked for changed from his ten dollar bill.
3. [The rope / The small but heavy rope / The rope behind my shed / The rope that came apart] was in the garage we just finished painting.
4. [The student / The vain yet clever student / The student on Bonn's metro / The student who was reading] wanted a vacation in the Swiss Alps.
5. [The girls / The small and giggling girls / The girls in Jane's sandbox / The girls who were laughing] played with the toys they got for Christmas.
6. [The boxes / The small yet awkward boxes / The boxes on his workbench / The boxes that were empty] were in the bedroom next to the stairs.
7. [The movie / The short but tedious movie / The movie in our theatre / The movie that was premiering] had a sad ending that made everyone cry.
8. [The doorknob / The round and shiny doorknob / The doorknob on our entrance / The doorknob that is broken] sticks in the cold weather that we get in Edmonton.
9. [The coyotes / The strong and agile coyotes / The coyotes of those mountains / The coyotes that are hungry] hunt in the hills that are next to the lake.
10. [The player / The proud but gifted player / The player in the outfield / The player who missed practice] left the team that was threatening to trade him.
11. [The lawyer / The tough and angry lawyer / The lawyer in Steve's courtroom / The lawyer who was speaking] won the case that had been covered in the newspapers.
12. [The car / The large but foreign car / The car on York's highway / The car that was rented] belongs to the girl who we met at the dance club.
13. [The river / The large and raging river / The river near their city / The river that stopped flooding] empties into the bay that borders the little town.
14. [The doctor / The old and friendly doctor / The doctor at her clinic / The doctor who was quitting] fired the nurse who was stealing drugs.
15. [The torches / The dim but flaming torches / The torches on their building / The torches that were costly] burned in the city that won the Stanley Cup.
16. [The hotel / The old yet fancy hotel / The hotel on our ravine / The hotel that just opened] was near the train station everyone wants to restore.
17. [The martians / The green yet friendly martians / The martians from this spacecraft / The martians who had landed] needed some fuel that contained unusual compounds.
18. [The cake / The rich and tasty cake / The cake in Bill's oven / The cake that was tasty] was served at the dinner that preceded the mayor's speech.
19. [The birds / The large yet fragile birds / The birds in your bushes / The birds that are frightened] flew into the forest that the company wants to chop down.
20. [The troops / The young and fearless troops / The troops from their barracks / The troops that are training] marched through the woods that are full of bears.
21. [The plane / The small but mighty plane / The plane under thick clouds / The plane that was leaving] climbed higher into the sky that was clear and blue.
22. [The horse / The young but tired horse / The horse in Sue's pasture / The horse that was training] bit the young child who was wearing a pink shirt.
23. [The instructor / The friendly and kind instructor / The instructor from King's college / The instructor who was useless] taught the introductory course that's now a requirement.
24. [The thief / The smart but crafty thief / The thief on Pat's rooftop / The thief who was cornered] escaped from the police who were trying to catch him.

APPENDIX B: EXPERIMENTAL SENTENCES USED IN EXPERIMENT 2

Option (a) is the low syntactic complexity condition; (b) is the medium complexity condition; and (c) is the high complexity condition. The first noun phrase is the subject of the sentence, and the second is the verb's object.

1. frightened
(a) The friendliest cop / the enterprising girls
(b) The friend of the cop / the boyfriend of the girls
(c) The guy who's a cop / the boy who kissed the girls

2. chased
(a) The old and flea-ridden dog / the sanitary car
(b) The horse right next to the dog / the buyer of the car
(c) The horse that hated the dog / the drunk who stole the car

3. smacked
(a) The superstitious kid / the orange-colored ball
(b) The brother of the kid / the owner of the ball
(c) The man who saw the kid / the man who lost the ball

4. hates
(a) The incredibly kind nun / the terribly evil judge
(b) The enemy of the nun / the attitude of the judge
(c) The girl who looks like a nun / the guy acting as the judge

5. dropped
(a) The uninvited guest / the cinnamon pineapple pie
(b) The jerk beside the guest / the candy on top of the pie
(c) The jerk who brought the guest / the cake which was next to the pie

6. pleased
(a) The enthusiastic band / the very impatient crowd
(b) The pianist in the band / the senator in the crowd
(c) The man who started the band / the girl who was in the crowd

7. shot
(a) The California crooks / the terribly frightened nurse
(b) The leader of the crooks / the baby beside the nurse
(c) The men who look like crooks / the doctor who loved the nurse

8. challenged
(a) The simple minded test / the secretary's group
(b) The questions on the test / the smart-ass in the group
(c) The man who gave the test / the ones who joined the group

9. swam in
(a) The Massachusetts duck / the beautiful salt-water sea.
(b) The offspring of the duck / the lake at the edge of the sea
(c) The guy who owns the duck / the river that's next to the sea

10. drove
(a) The poverty-stricken old tramp / the Chevrolet four-wheel drive truck
(b) The son of the neighborhood tramp / the car underneath all the trucks
(c) The guy who avoided the tramp / the car that was next to the trucks

11. wore
(a) The seemingly crazy old Scot / a beautiful plaid-covered kilt
(b) The friend of the crazy old Scot / a shirt with a plaid-covered kilt
(c) The guy who believes he's a Scot / a skirt that I think was a kilt

12. saw
(a) The rowdiest class / the Boris Karloff film
(b) The girl in my class / the sequel to the film
(c) The girl who's in class / the man who liked the film

13. went to
(a) The genuine Norwegian prince / the totally fixed-up new gym
(b) The former guitarist for Prince / the bar at the back of the gym
(c) The guy who likes listening to Prince / the bar that's next-door to the gym

14. fell on
(a) The beautiful pendulum clock / the waxy linoleum floor
(b) The flowers right next to the clock / the carpet right under the floor
(c) The ashtray that's next to the clock / the carpet I'd laid on the floor

15. requested
(a) The Heritage Federal bank / a regular two-sided form
(b) The president of the new bank / a duplicate side of the form
(c) The student who used that bank / a letter that backs up the form

16. broke
(a) The arrogant coach / the Alabama law
(b) The wife of the coach / the letter of the law
(c) The man who's the coach / the rule that's now a law

17. haunts
(a) The creepiest ghost / the Mexican–Indian house

(b) The son of the ghost / the mansion right next to the house
(c) The guy who's a ghost / the mansion that borders the house

18. made
(a) The Lebanese–Italian chef / the Uncle Ben's Spanish-style rice
(b) The daughter of our finest chef / the mushrooms and meat for the rice
(c) The girl who wants to be a chef / the pasta that takes just like rice

19. sleeps on
(a) The ugly finicky cat / the old Hungarian rug
(b) The puppy next to the cat / the bed instead of the rug
(c) The puppy chased by the cat / the bed that's soft as a rug

20. got
(a) The laziest boss / the teletype programming job
(b) The friend of the boss / a high school degree for the job
(c) The guy just made boss / the files he needs for the job

21. stole
(a) The sentimental priest / the beautifully hand-painted cups
(b) The lawyer for the priest / the wine and the beer for the cups
(c) The boy who helps the priest / the box that contained all the cups

22. worried
(a) The violent strike / the peaceful little town
(b) The length of the strike / the people in the town
(c) The man who's on strike / the folks who live in town

23. avoided
(a) The huge overfed rat / the old Canadian cheese
(b) The mice under the rat / the slice of Parmesan cheese
(c) The mice chased by the rat / the trap containing the cheese

24. shocked
(a) The pornographic book / the sheltered little child
(b) The contents of the book / the parents of the child
(c) The author of the book / the folks who raised the child

25. cherished
(a) The Mississippi tribe / the unenlightened chief
(b) The members of the tribe / the killer of the chief
(c) The ones who formed the tribe / the man they'd made their chief

26. is made of
(a) The Continental bed / a controversial foam
(b) The mattress on the bed / a brand new kind of foam
(c) The frame that holds the bed / some stuff that's just like foam

27. broke
(a) The predatory bird / its incredibly long beak
(b) The duck right by the bird / its wing instead of its beak
(c) The duck that chased the bird / the glass by using its beak

APPENDIX C: EXPERIMENTAL SENTENCES USED IN EXPERIMENT 3

The first option within brackets is the plausible conditions, and the second option is the implausible conditions. The prepositional phrase in parentheses was included only in the long versions.

1. [The car / The girl] (in the garage) had just been sitting outside.
2. [The boy / The snake] (in the bed) was very ugly and scary to look at.
3. [The birds / The cats] (in the tree) scattered everywhere when the gun fired.
4 [The dish / The drum] (in the sink) was covered with grime and dirt.
5. [The crumbs / The socks] (under the table) made the room appear messy and dirty.
6. [The papers / The sandwich] (in the briefcase) got all wet because of the rain.
7. [The poster / The knitting] (on the wall) had just been mounted and hung.
8. [The thief / The nurse] (in the prison) had to serve a long sentence.
9. [The plants / The cans] (in the garden) were exposed to rain and frost.
10. [The headlights / The sunbeam] (on the car) shone with great intensity.
11. [The book / The rug] (on the shelf) was removed because it looked so ugly.
12. [The accountant / The engineer] (near the ledger) was prepared to go to work.
13. [The music / The roses] (on the stereo) amazed the impressionable young student.
14. [The snow / The hail] (on the driveway) was proof that the city had been hit by a storm.
15. [The customer / The principal] (at the market) was loudly demanding immediate service.
16. [The doorknob / The ivy] (on the entrance) didn't seem to serve any useful purpose.
17. [The swimmer / The biker] (near the lifejackets) almost fell out of the little boat.
18. [The wine / The salt] (from the decanter) had been aging for over twelve years.
19. [The soldier / The housewife] (near the helicopter) was returning from a mission.

20. [The waiter / The artist] (with the menu) spilled coffee all over the customer.

References

Bock, J. K. (1982). Towards a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, **89**, 1–47.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, **18**, 355–387.

Bock, J. K. (1987). An effect of the accessibility of word forms on sentence structures. *Journal of Memory and Language*, **26**, 119–137.

Bock, J. K., & Kroch, A. S. (1989). The isolability of syntactic processing. In G. N. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing*. Dordrecht: Kluwer Academic.

Butterworth, B. (1980). Evidence from pauses in speech. In B. Butterworth (Ed.), *Language production: Vol. 1. Speech and talk*. London: Academic Press.

Chomsky, N. (1970). Remarks on nominalization. In R. Jacobs & P. Rosenbaum (Eds.), *Readings in English transformational grammar*. Waltham, MA: Ginn.

Collard, R., & Povel, D. (1982). Theory of serial pattern production: Tree traversals. *Psychological Review*, **6**, 693–707.

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

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, **27**, 124–142.

Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, **20**, 611–629.

Ferreira, F. (1988). *Planning and timing in sentence production: The syntax-to-phonology conversion*. Unpublished dissertation, University of Massachusetts, Amherst, MA.

Ferreira, F., & Clifton, C. E. (1987, November). *Planning and prosodic structure in sentence production*. Paper presented at the Annual Meeting of the Psychonomic Society, Seattle, WA.

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language: An introduction to psycholinguistics and generative grammar*. New York: McGraw-Hill.

Fodor, J. D. (1989). Empty categories in sentence processing. *Language and Cognitive Processes*, **4**, 155–209.

Ford, M. (1982). Sentence planning units: Implications for the speaker's representation of meaningful relations underlying sentences. In J. Bresnan (Ed.), *The mental representation of grammatical relations*. Cambridge, MA: MIT Press.

Frazier, L. (1989). Against lexical generation of syntax. In W. D. Marlsen-Wilson (Ed.), *Lexical representation and process*. Cambridge, MA: MIT Press.

Fromkin, V. A. (1971). The nonanomalous nature of anomalous utterances. *Language*, **47**, 27–52.

Garrett, M. F. (1975). The analysis of sentence production. In G. Bower (Ed.), *Psychology of learning and motivation: Vol. 9*. New York: Academic Press.

Garrett, M. F. (1976). Syntactic processes in sentence production. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms*. Amsterdam: North-Holland.

Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions*. London: Academic Press.

Gee, J. P., & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, **15**, 411–458.

Goldman-Eisler, F. (1967). Sequential temporal patterns and cognitive processes in speech. *Language and Speech*, **10**, 122–132.

Gordon, P. C., & Meyer, D. E. (1987). Control of serial order in rapidly spoken syllable sequences. *Journal of Memory and Language*, **26**, 300–321.

Jackendoff, R. (1977). *X'-syntax: A study of phrase structure*. Cambridge, MA: MIT Press.

Johnson, N. F. (1966). On the relationship between sentence structure and the latency in generating the sentence. *Journal of Verbal Learning and Verbal Behavior*, **5**, 375–380.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, **11**, 201–258.

Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Erlbaum.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Monsell, S., & Sternberg, S. (1981). *Speech programming: A critical review, a new experimental approach, and a model of the timing of rapid utterances* (Part 1). Unpublished manuscript.

Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. Unpublished dissertation, MIT.

Rosenbaum, D. A. (1985). Motor programming. In H. Heuer, U. Kleinbeck, & K.-M. Schmidt (Eds.), *Motor behavior: Programming, control, and acquisition*. Berlin: Springer-Verlag.

Rosenbaum, D. A., Inhoff, A. W., & Gordon, A. M. (1984). Choosing between movement sequences: A hierarchical editor model. *Journal of Experimental Psychology: General*, **113**, 372–393.

ROSENBAUM, D. A., KENNY, S. B., & DERR, M. (1983). Hierarchical control of rapid movement sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 86–102.

SELKIRK, E. O. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

SELKIRK, E. O. (1986). Derived domains in sentence phonology. *Phonology Yearbook*, 3, 371–405.

STERNBERG, S., MONSELL, S., KNOLL, R. L., & WRIGHT, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning*. New York: Academic Press.

STERNBERG, S., KNOLL, R. L., & TUROCK, D. L. (1990). Hierarchical control in the execution of action sequences: Tests of two invariance properties. In M. Jeannerod (Ed.), *Attention and Performance XII: Motor representation and control*. Hillsdale, NJ: Erlbaum.

STOWELL, T. (1981). *Origins of phrase structure*. Unpublished doctor dissertation, MIT.

TANNENBAUM, P. H., WILLIAMS, F., & HILLIER, C. S. (1965). Word predictability in the environments of hesitations. *Journal of Verbal Learning and Verbal Behavior*, 4, 134–140.

YNGVE, V. H. (1960). A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society*, 104, 444–466.

(Received January 6, 1990)
(Revision received July 16, 1990)