

Quantifying contextual contributions to word-recognition processes

LORRAINE K. TYLER

*Max Planck Institute für Psycholinguistik, Nijmegen, The Netherlands
and MRC Applied Psychology Unit, Cambridge, England*

and

JEANINE WESSELS

Max Planck Institute für Psycholinguistik, Nijmegen, The Netherlands

The experiment reported here used the gating paradigm (Grosjean, 1980) to investigate the following issues: To test the validity of the claims made by the "cohort" theory (Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Welsh, 1978) for the interaction of sensory and contextual constraints during the process of recognizing spoken words, and to determine the relative contribution of two kinds of contextual constraint—syntactic and interpretative—in reducing the amount of sensory input needed for recognition. The results both provide good support for the cohort model, and show that although strong syntactic constraints on form-class only marginally reduce the amount of sensory input needed, a minimal interpretative context has a substantial facilitatory effect on word recognition.

In a normal conversational setting, people speak to each other at a rate of 2-3 words/sec. Listeners have to interpret the speech input at roughly the same speed as it is produced, if they are to avoid a backlog of uninterpreted input. How is this process of rapid interpretation achieved? To answer this question, we need, first of all, to specify the processes involved in recognizing individual words, for it is only on the basis of the syntactic and semantic information made available when a word is identified that the listener can construct a meaningful representation of the input.

In recent years, a number of experiments have elucidated some of the basic processes involved in recognizing spoken words (Cole & Jakimik, 1980; Grosjean, 1980; Marslen-Wilson & Tyler, 1975, 1980; Marslen-Wilson & Welsh, 1978). These studies have shown that listeners can identify words very rapidly—even before all of the word has been heard and when a large number of words are still compatible with the sensory input. Furthermore, these fast word-recognition decisions are strongly affected by the syntactic and interpretative context in which the word appears.¹ In order to correctly identify a word, listeners need to hear less of the sensory signal when the word is heard in an appropriate context than when the same word is heard in isolation (Grosjean, 1980; Marslen-Wilson & Tyler, 1980).

Special thanks are extended to William Marslen-Wilson for his comments on the manuscript. The authors' mailing address is: Max Planck Institute für Psycholinguistik, Berg en Dalseweg 79, Nijmegen, The Netherlands.

To account for these properties of spoken word recognition, Marslen-Wilson developed the "cohort model" (Marslen-Wilson, 1980; Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Welsh, 1978). According to this model, the first one or two phonemes of a word² serve to activate all of those words in the listener's mental lexicon which begin with that initial sensory sequence. When a word is heard in isolation, these word candidates continue to be assessed against the demands of the subsequent sensory input. When a mismatch occurs, a word candidate drops out of the pool. This process continues until only a single word candidate matches the sensory input and, at that point, the listener recognizes the word. When a word is heard in an utterance, however, the suitability of each word candidate is assessed against the syntactic and interpretative specifications of the context, as well as the incoming sensory input. In this case, recognition occurs when a single word candidate remains which matches both the context and the sensory input.

The advantage of this model, then, is that it makes precise a priori predictions about the exact point in a word, going from left to right, at which that word should be recognized—both in and out of context. These predictions have been verified for words in isolation using the lexical decision and phoneme monitoring tasks (Marslen-Wilson, 1980, 1983).

For words in context, the specific predictions of the cohort theory have not, before now, been tested. Only a general claim concerning contextual interactions has been verified. Namely, that listeners need to hear less sensory input when a word appears in a sentence context

than when the same word is heard in isolation (Grosjean, 1980; Marslen-Wilson & Tyler, 1980). In these experiments, however, the relationship between recognition point and activation of word candidates was not explicitly manipulated, and therefore the studies do not test the more precise claim of the cohort theory: that a word in context will be recognized when only one member of the initial cohort is consistent with both the sensory input and contextual constraints, even though at that point in the word there may be other members of the original cohort compatible with the sensory input. This is the claim that was tested in the experiment to be reported here.

However, it is impossible to test claims about the precise effect of contextual constraints without further specifying what these constraints actually consist of. A very obvious distinction to draw is between syntactic and interpretative sources of constraint. Unfortunately, the cohort theory, like all other theories of word recognition, has not explored in detail the effects of these different types of contextual constraint on word recognition. Although all theories acknowledge that such a distinction exists in the sense that each source of constraint is assumed to facilitate word recognition, most theories have not considered the further issue of whether there are quantitative or qualitative differences in the way that each source of constraint functions with respect to word recognition processes. Instead, the theories appear to assume that syntactic and interpretative information are functionally indistinguishable in terms of their effects on the recognition of spoken words. The major divergence between theories is in the various claims they make about the locus of contextual effects—that is, whether they operate at the preaccess stage (e.g., Morton, 1969), during access (e.g., Marslen-Wilson & Tyler, 1980), or at a postaccess stage (e.g., Forster, 1979).

Consistent with this general lack of theoretical interest in the effects of different kinds of contextual constraint is the paucity of experimental work concerned with this issue.³ There is no research which explicitly examines the relative facilitative effects of these different forms of constraint on word-recognition processes, in the sense of determining the extent to which the presence of each source of constraint reduces the amount of sensory input needed for recognition.

The research which does bear on this issue, although indirectly, falls into three categories. First, there are those studies which have been undertaken to determine whether there is any interaction at all between syntactic and semantic structural constraints and sensory input, without taking the further step of comparing the relative effects of each type of constraint. What these experiments show is that when the availability of either syntactic or semantic structure is manipulated, the absence of either increases the listener's difficulty in recognizing individual words (e.g., Marslen-Wilson & Tyler, 1975, 1980; Miller, Heise, & Lichten, 1951; Miller & Isard, 1963).

Other research on the effects of contextual constraints on word recognition processes has focused almost exclusively upon the effects of semantic context, and has tended to ignore syntactic constraints. These studies unanimously show that semantic constraints facilitate the speed with which a word can be identified, whether these constraints are generated on the basis of semantic associations (e.g., Blank & Foss, 1978; Kalikow, Stevens, & Elliott, 1977) or on the basis of sentential meaning (e.g., Cairns, Cowart, & Jablon, 1981; Cole & Jakimik, 1978; Grosjean, 1980; Marslen-Wilson & Tyler, 1975, 1980; Morton & Long, 1976; Underwood, 1977).

The only experimental investigation of the effect of syntactic form-class constraints on word recognition has focused on a very specific aspect of lexical processing—namely, the resolution of a single reading of an ambiguous word. The experiments in this domain have usually been conducted within the wider framework of a comparison of various types of contextual constraint on the resolution of ambiguous words. Taken together, these experiments can be interpreted as showing that semantic constraints based upon interlexical associations constrain the interpretation of a word soon after it has been heard, whereas syntactic form-class and pragmatic constraints take longer to exert their influence (Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Swinney, 1979; Tanenhaus, Leiman, & Seidenberg, 1979; Oden & Spira, 1978). However, these experiments are only marginally related to the present issues, since they focus on the effectiveness of various types of constraint in suppressing contextually inappropriate readings of ambiguous words, rather than determining the extent to which these constraints facilitate a word's recognition by reducing the amount of sensory input the listener needs to hear.

This latter issue clearly remains underdetermined by the data, and for this reason it was the focus of the experiment to be reported here. The aim of the study was to determine the relative extent to which syntactic and interpretative constraints each reduce the amount of sensory input the listener needs to hear in order to identify a word and thereby speed up the process of integration of the word into the interpretation of the utterance. The cohort model provided the theoretical framework within which this issue was examined.

To answer this question, we used the gating task, developed by Grosjean (1980), in which subjects hear successive presentations of fragments of a target word. At each presentation, the size of the fragment is increased by a constant amount. If the first fragment consists, for example, of the first 50 msec of the word, then the second will consist of the first 100 msec, and so on, until the whole word has been presented. After each fragment, subjects write down the word they think is being presented, together with a rating of how confident they are about their choice. This procedure provides an estimate of the amount of sensory input listeners need to hear in order to correctly identify a word.

Furthermore, following Grosjean, we assume that the gating task reflects the normal processes involved in on-line comprehension of speech. That is, the task tells us what is the maximum information that the subject can extract from a given amount of sensory input, and we assume that this accurately reflects what the listener, when he hears a word normally, can extract from the incoming word at different points in time. Thus, the processing profile that we can build up, increment by increment, in the gating task corresponds to the real-time profile of normal spoken word recognition. One source of evidence for this claim is the very close correspondence between the recognition time estimated from standard reaction-time tasks and those estimated from the gating task (Grosjean, 1980).

On the basis of data from his gating study, Grosjean (1980) has argued that word-recognition processes are somewhat more complex than had been previously assumed. That is, most models assume that once a word-recognition element reaches some criterial value, the word in question will be recognized. However, Grosjean's gating data suggested that the processes involved in identifying a word require at least two phases of analysis. During the first phase, the listener isolates a particular word candidate but may still feel unsure about this choice. During the second phase, he or she continues to monitor the sensory input until some criterial level of confidence is reached, and it is at the end of this second phase that the listener can be said to have recognized the word. Given this distinction between isolation of a word candidate and its recognition, the data from this experiment will be analyzed in terms of both isolation and recognition points.

To determine the role of syntactic and semantic context in reducing the amount of sensory input required for identification, the strength of syntactic and semantic constraints on target words were varied. We contrasted "minimal" semantic⁴ with no semantic context, and strong syntactic constraints on form class with weak syntactic constraints. Using the gating paradigm, we were able to determine the amount of sensory input required for identification as a function of the strength of each type of constraint independently.

The cohort model makes the strong prediction that a word will be recognized when it is the only remaining member of the initial cohort which matches both the sensory input and the demands of the syntactic and semantic context, even though at that point there may still be a number of other word candidates which match the sensory input alone. Since the gating paradigm enables us to locate the point at which a word is isolated and recognized, the present experiment allows us to evaluate these claims with respect to each phase of the word-recognition process to determine which produces the best fit with the theoretical predictions.

With respect to the differential effects of syntactic and semantic constraints on word-recognition processes,

the cohort model predicts that syntactic form-class constraints can have only a small effect on reducing the size of the cohort. There are two reasons for this. First, such constraints have, in principle, a limited ability to narrow in on a unique word candidate. The most they can do is to reject those word candidates which do not match the syntactic specifications of the context. In most cases, however, the syntactic structure of an utterance rarely places very narrow constraints on the form-class of possible continuations. Even when these restrictions do limit possible continuations to a single form-class category, as in the present experiment, it is almost never the case that the cohort contains only a single member of that category. Second, many words are ambiguous with respect to their syntactic category (many words can function as nouns, verbs, and adjectives, for example), and therefore cannot be eliminated from the cohort purely on syntactic grounds.

For these reasons, the cohort theory predicts that both strong and weak form-class constraints will lead to only a small reduction in the amount of sensory input needed for recognition.

Semantic constraints, on the other hand, can have a large effect in reducing the size of the initial cohort to a single member. The size of the effect will be determined by the strength of semantic constraints and the extent to which word candidates are semantically inappropriate. However, even the weak semantic constraints used in the present experiment should produce some advantage in terms of eliminating semantically inappropriate candidates and therefore reduce the amount of sensory input needed for recognition.

METHOD

Subjects

Sixty paid subjects participated in the experiment. They were all native speakers of Dutch and the experiment was carried out in The Netherlands.

Materials

The materials consisted of 25 target words, all of which were infinitive verbs of two, three, or four syllables in length. Each target word was chosen so that the size of its initial cohort, based on the first two phonemes, was over 60 members, with at least 20 of these being infinitive verbs. The strength of syntactic and semantic constraints was varied in the following way. For each target word, a set of four sentence pairs was constructed, with the target occurring in the second sentence of each pair. Two of the four pairs consisted of sentences that were semantically anomalous but syntactically normal. These provided a no-semantic context condition. The other two sets of sentence pairs consisted of syntactically and semantically normal material. In these conditions the material preceding the target provided a minimal interpretative context for the target word, which could be contrasted with the no-semantic context condition.

The strength of syntactic constraints was manipulated by varying local constraints on form-class. In the weak syntactic constraint condition, we chose structures in the Dutch language which placed minimal constraints on the form-class of the target. In general, the only syntactic restriction here was that certain inflected forms of verbs were prohibited. In contrast, the strong

syntactic constraint condition was the most syntactically constraining—in terms of narrowing down the set of possible form-class continuations—that the Dutch language allowed. In this condition, the target was preceded by the word *te*, which can only be followed by one of two forms—either the infinitive verb form or one of a small number of adjectives (e.g., *te groot* = too big).⁵

These materials were then pretested, using an auditory cloze procedure, to ensure that the interpretative context in the minimal semantic context condition did indeed impose only minimal constraints on the target words.

In this pretest, listeners heard the first 100 msec of each target word in the four context conditions. We chose this size acoustic segment for the following reason. We wanted a strong test of the nonpredictability of target words. If subjects had not been provided with any sensory input corresponding to the target, and had produced a wide range of responses, we would have judged the predictability of the target to be low. But this would have been only a weak test of predictability. If, on the other hand, subjects hear the first phoneme of the target and still do not produce the word, then the target can be confidently considered not to be predictable.

So that each listener would hear a target word only once, four versions of the materials were constructed, with conditions pseudorandomly distributed within a version. Eight subjects were tested on each version. The listener's task was to say the word he or she considered to be an appropriate continuation for the sentence pair. The following scoring procedure was used: 1 = identity with the target word; 2 = synonym of target word; 3 = related to target word; 4 = contextually appropriate but unrelated to target; and 5 = contextually inappropriate and unrelated to target.

Two independent judges scored the subjects' responses according to the above set of criteria. Any disagreements that occurred were discussed and resolved.

The mean ratings for the two semantically normal conditions were 3.95 and 3.96, indicating that although subjects' responses were contextually appropriate, they rarely produced the target word itself. For the two semantically anomalous conditions, the ratings were 4.88 and 4.97, showing that subjects' responses were unrelated to the intended target.

Covarying syntactic and semantic constraints resulted in four experimental conditions: (1) minimal semantic constraint + strong syntactic constraint; (2) minimal semantic constraint + weak syntactic constraint; (3) no semantic constraint + strong semantic constraint; and (4) no semantic constraint + weak syntactic constraint.

A fifth experimental condition—a no-context condition—in which the target was presented alone, was also included. This provided a baseline measure of the amount of sensory input required for recognition when no contextual constraints were available. An example stimulus set (in Dutch),⁶ showing these four context conditions, is given below.

Target word: *profiteren*

- (a) Minimal semantic/strong syntax: *De afspraak met de tandarts gaat niet door. Jan probeert te . . .*
- (b) Minimal semantic/weak syntax: *De afspraak met de tandarts gaat niet door. Jan kan . . .*
- (c) No semantic/strong syntax: *De adem met de leugen schuift pas door. Het terras tracht te . . .*
- (d) No semantic/weak syntax: *De adem met de leugen schuift pas door. Het terras wil . . .*

A female native speaker of Dutch, who was unaware of the purpose of the study, recorded the materials in each of the four context conditions. At the same time, the target words spoken in a neutral carrier phrase ("The following word is . . .") were also recorded. Targets were then excised from the neutral carrier phrases and digitized at a sampling rate of 20 kHz. The duration of the word was displayed on a screen and 50-msec

segments were marked off, starting from the onset of each word, by means of a cursor. Each target was then output from the computer in a sequence of segments, each of which increased by 50 msec in duration so that the first segment of a word consisted of the first 50 msec, the second consisted of the first 100 msec, and so on, until the entire word had been output. The total number of segments for each word depended upon the total duration of the word. For the 25 targets used in this experiment, the total number of segments ranged from 11 to 19.

These segments, taken from the recordings of the target words in the neutral carrier phrase, were then inserted into the recordings of each of the four context conditions. These sequences of segments were also used as stimuli in the no-context condition. In the context conditions each presentation of a segment was preceded by the context material, but in the no-context condition, the sequence of segments was presented in isolation.

Excising the targets from the neutral carrier phrases and inserting them into each of the experimental conditions meant that the same acoustic tokens appeared in each condition. Since each target word occurred in all five experimental conditions and we wanted a target word to be heard in only a single condition by each subject, five versions of the materials had to be constructed. Each version contained five instances of each of the five experimental conditions, with the five no-context items blocked at the beginning of a version and the items in the other four conditions pseudorandomly distributed across each version. Twenty-five filler items were interspersed between the test materials to obscure their regularities, and 10 practice items preceded the testing sequence. Twelve subjects were tested on each of the five versions. The final list for each version consisted of 514 items—263 fillers and 251 test items. Total presentation time was 3½ h, with each subject being tested in two separate sessions of 1½ h. The two sessions were held at the same time on successive days.

Procedure

The subjects were tested in groups of four. They were instructed to listen to the material as carefully as possible and, after hearing each fragment, to write down the word they thought was being presented. Then they were to indicate how confident they were about their choice by selecting a point on a scale of 1-10, with 10 being absolutely confident and 1 being completely unsure.

RESULTS AND DISCUSSION

In view of the distinction Grosjean (1980) has drawn between isolation and recognition points, the data were analyzed in terms of each of these phases of analysis.

Relative Effects of Syntactic and Semantic Constraints

Isolation points. We will discuss first the effects of syntactic and semantic constraints on word isolation processes, since it is only after determining the amount of sensory input required for isolation in the various context conditions that we can evaluate the predictions made by the cohort theory.

Subjects' responses were examined to locate the segment at which they had correctly identified the word and had not subsequently changed their minds. This was done for all targets in all five conditions. These "isolation" points were a measure of the amount of sensory input each listener needed to identify the target words. The mean isolation points for each condition are displayed in Table 1. This table shows that semantic constraints substantially reduced the amount of sensory

Table 1
Mean Isolation Points (in Milliseconds) for Each Condition

No Context	Minimal Semantic/ Strong Syntax	No Semantic/ Strong Syntax	Minimal Semantic/ Weak Syntax	No Semantic/ Weak Syntax
409	298	393	322	433

input required in order to isolate the correct word candidate, whereas syntactic constraints exerted a much smaller effect.

The difference in the amount of sensory information required in the various experimental conditions is shown graphically in Figure 1. This figure presents the cumulative distributions of the percentage of correctly recognized words as a function of the number of segments needed for isolation in each condition. As the figure shows, the presence of a minimal semantic context results in words being isolated earlier than in either the no-semantic context or the no-context conditions, with 50% of the words being isolated by the fifth segment—that is, after the listener has heard, on average, 250 msec of the word. In contrast, in the no-semantic context conditions, or when targets appeared without any context at all, isolation occurs considerably later, with 50% of targets being isolated after listeners have heard between 350 and 400 msec of the sensory signal (between 7 and 8 segments).

The mean isolation points for each item in each of the five conditions were entered into an analysis of variance, with items crossed by conditions.⁷ The effect of conditions was highly significant [$F(4,92) = 14.791, p < .001$]. A set of Newman-Keuls post hoc comparisons was subsequently performed, using the error term derived from the ANOVA. These comparisons revealed, first, that the amount of sensory input required for correct recognition in the semantic constraint conditions was significantly less than in the no-context condition. Words were recognized 87 msec earlier in the minimal-semantic-constraint/weak-syntax condition than in the isolation condition ($p < .01$) and 111 msec earlier in the minimal-semantic-constraint/strong-syntax condition ($p < .01$). The difference of 24 msec between the two semantic context conditions, however, was not significant ($p > .05$). Semantic constraints exerted a relatively constant influence, irrespective of whether syntactic constraints were strong or weak.

The same pattern appears in the comparison between

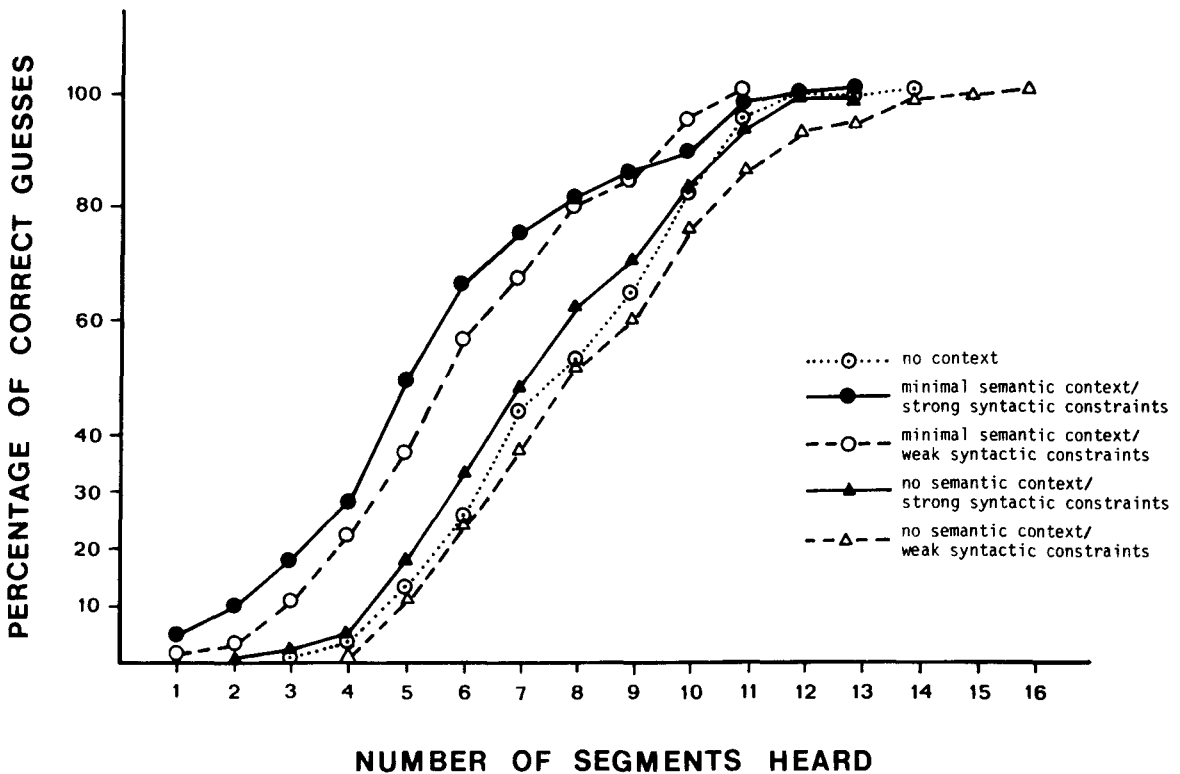


Figure 1. Cumulative distributions of the percentage of correctly identified words as a function of the number of segments needed for isolation in each of the five conditions.

minimal and no semantic context in the weak- and strong-syntax conditions. In the strong-syntax conditions, the difference between no semantic context and minimal semantic context of 96 msec was significant ($p < .01$), as was the difference of 100 msec in the weak-syntax condition. This also shows that minimal semantic constraints exerted a strong facilitatory effect on isolation points that was relatively unaffected by the strength of the accompanying syntactic constraints.

In contrast to the effects of semantic constraints, syntactic constraints did not significantly reduce the amount of sensory input required, relative to the no-context condition. The difference of 15 msec between the no-semantic-context/strong-syntactic-constraint and the no-context conditions was not significant ($p > .05$); the difference of 24 msec between the no-semantic-context/weak-syntactic-constraint and the no-context conditions was also not significant ($p > .05$). The difference between the two no-semantic-constraint conditions was 39 msec, but this also failed to reach significance ($p > .05$).

However, given the trend in the data towards strong syntactic constraints exerting a small facilitatory effect in both the no-semantic-constraint and the minimal-semantic-constraint conditions, we carried out an additional analysis of the data to determine whether there was a main effect of syntax. For this analysis, the isolation points for the no-context condition provided the baseline condition against which the effects of context could be evaluated. Therefore, the mean isolation point (collapsing across subjects) for each item in each of the four context conditions was subtracted from each item's mean isolation point in the no-context condition. These differences were a measure of the degree to which each type of context reduced the amount of sensory input necessary for isolation relative to the amount required when the word was heard without any prior context.

An analysis of variance was performed on these differences, with syntax and semantics as fixed effects. The analysis showed a main effect of both syntactic and semantic constraints. There was, however, a large difference in the size of the effect due to each source of constraint. Semantic constraints reduced the amount of sensory input by 103 msec [$F(1,23) = 24.85, p < .001$], whereas syntactic constraints reduced it by only 31 msec [$F(1,23) = 4.60, p = .043$]. These effects were constant across conditions, as shown by the lack of any interaction between syntax and semantics ($F < 1$). Strong syntactic constraints exerted their effect independently of the presence or absence of a semantic context. Similarly, semantic constraints exerted a relatively con-

stant influence, irrespective of whether syntactic constraints were strong or weak.

Taking these two analyses together, then, we find that syntactic constraints exert a small, but significant, effect. The reason that this effect did not reach significance in the first analysis was presumably because of the conservative nature of the Newman-Keuls statistic. Semantic context, however, clearly provides a much greater degree of facilitation. It reduces the amount of sensory input required for recognition three times more than does the presence of a strong syntactic context.

Recognition points. Following Grosjean (1980), the recognition point was defined as that segment at which subjects identified the target and were confident of their choice. We chose a confidence rating of 80% as our cutoff value and took as the recognition point that segment at which subjects correctly produced the target word with a confidence rating of 80%, and did not subsequently change their minds. Table 2 shows the mean recognition points for the five conditions. Although the absolute amount of sensory input needed in each condition is greater than that required for isolating a word, the relationship between each condition remains the same as in the prior analysis. Not surprisingly, then, the results of the ANOVAs performed on these recognition data were very similar to those carried out on the isolation data. In a first ANOVA on the mean recognition points for each item in each of the five conditions, the effect of experimental conditions was highly significant [$F(4,92) = 16.902, p < .001$] and the results of the Newman-Keuls post hoc comparisons paralleled those carried out on the isolation data. A second ANOVA, carried out on the differences between each context condition and the no-context condition, produced a small, but significant, main effect of syntax [$F(1,23) = 4.55, p = .044$] and a much larger effect of semantics [$F(1,23) = 27.31, p < .001$].

To statistically establish that there were, indeed, no significant differences between the isolation and recognition points in terms of the effects of syntactic and semantic constraints, we ran an ANOVA on the two sets of data. Although there was an overall difference between them, with recognition points being consistently later than isolation points [$F(1,46) = 16.34, p < .001$], there was no difference in the pattern of contextual effects for the two sets of data ($F < 1$). It is interesting to note that we find the same degree of contextual facilitation from the earliest point at which listeners can recognize the word and yet are relatively uncertain about their choice, and from the point at which they recognize it and are sure of their decision. It is clear

Table 2
Mean Recognition Points (in Milliseconds) for Each Condition

No Context	Minimal Semantic/ Strong Syntax	No Semantic/ Strong Syntax	Minimal Semantic/ Weak Syntax	No Semantic/ Weak Syntax
469	381	465	401	490

from these results, then, that the two distinct stages of word identification proposed by Grosjean (1980) have no consequences for the interaction between contextual constraints and sensory input during the process of identifying a word. Where differences between isolation and recognition points do emerge will become apparent in a later section, in which we discuss the cohort theory.⁸

Identification of *Te*

These analyses, then, both with and without confidence ratings, show that the strongest syntactic constraints on form-class that the Dutch language allows exert only a minimal facilitatory effect on word isolation and recognition processes. Could this have been due to listeners' difficulty in correctly perceiving the syntactically constraining information that preceded the target? The strong syntactic constraint was provided by a single-function word, *te*, which preceded the target word. Given the sentence structures used in the experiment, *te* could have been misperceived by listeners as either *de* or as an inflection on the prior modal or auxiliary. If so, then the strong syntax condition would not, in fact, have been strongly constraining.

We therefore conducted a posttest to determine the accuracy with which *te* was perceived. This was done by presenting 16 further subjects with the second sentence of each sentence pair up to, but not including, the target word, and asking them to write down exactly what they had heard.⁹

On the basis of this posttest, we found that there were 17 items for which 75% of the listeners correctly perceived the syntactically constraining *te*. Four separate ANOVAs were performed on these 17 items, repeating those performed on the full set of items.¹⁰ The results of these analyses, and the set of Newman-Keuls comparisons between the experimental conditions, were similar to the earlier ANOVAs on the total set of items, in that there was a large effect of semantics and a small effect of syntax. The size of the syntactic effect was, on average, 11 msec larger for the 17 items than it had been for the full set of 24 items. Similarly, the size of the effect of semantics was, on average, 18 msec larger.

The remaining items, in which *te* had not been correctly identified in the posttest, were also subjected to the same set of four ANOVAs.¹¹ These analyses showed no significant effect of syntactic constraint ($F < 1$) and, at the same time, a reduced, although significant, effect of semantic constraints. A possible explanation for this is that uncertainty over the syntactic category of a word might increase the difficulty of determining the semantic role the word plays in the utterance.

The results of the posttest, then, eliminate the possibility that the small size of the effect of syntactic constraints was due to subjects' inability to correctly perceive the syntactically constraining information. Considering just those items in which *te* could be unambiguously perceived, the effect of syntactic constraints was only marginally larger than in the main analyses reported earlier.

Contextual Interactions and Predictions of the Cohort Theory

According to the cohort theory, word candidates are initially activated on the basis of the sensory input. These candidates are continually evaluated against the demands of the context and of the subsequent sensory input. A word is recognized when only one member of the initial cohort matches both the context and the sensory input, and word candidates can be dropped from consideration because they fail to conform to either syntactic or interpretative contextual constraints.

This theory, then, makes a priori predictions about the exact point in a word at which it should be recognized. However, because of problems involved in determining the range over which coarticulation effects extend, it is extremely difficult in practice to establish the theoretical recognition point. This makes it problematical to establish the exact point at which the critical acoustic-phonetic information first becomes available to the listener. Given these difficulties, we decided to test the claims of the cohort theory in a different way. That is, rather than calculating in advance the theoretical recognition point for each word, we took the point at which listeners actually identified each word and determined whether there were any other word candidates remaining at the point that satisfied both sensory and contextual criteria. According to the cohort theory, there should be only one word candidate possible at the point at which a word is identified, and that should be the target word itself.

This prediction was tested with respect to both the isolation and recognition points for each word. We begin with the analysis on the isolation points. First of all, we assessed the number of word candidates compatible with the sensory input alone (ignoring contextual appropriateness) at each isolation point in the five experimental conditions. This was done by counting the number of dictionary entries compatible with the sensory input at each isolation point. Given the uncertain theoretical status in the mental lexicon of inflectional and derivational variations, this count was done in two ways.¹² The first count consisted only of base morphemes, and the second count consisted of all possible inflectional and derivational word forms. In view of the general lack of consensus about the definition of "base morpheme," we followed Quirk and Greenbaum (1979) and defined base morphemes as those morphemes to which the rules of word formation apply. When the same base morpheme had two distinguishable meanings, two entries were included in the count.

Our next step involved determining which of these word candidates were, in fact, ruled out as being either syntactically or semantically inappropriate. This was done for the base morpheme count and the combined inflectional and derivational account separately. Four judges independently evaluated each word candidate in terms of its syntactic and semantic compatibility with the prior context. All four judges agreed unani-

mously on decisions of syntactic appropriateness. However, the issue of whether or not a word candidate was semantically appropriate for the context was more difficult. Since we preferred to err on the side of caution, we classed a word candidate as inappropriate only if all four judges agreed. For those cases in which judges' disagreements could not be resolved, the word candidate was included in the group of semantically appropriate word candidates. Since the results of the base morpheme analysis provided the best fit with the predictions made by the cohort theory, they will be considered in some detail. First of all, however, we will briefly describe the results for the combined inflected and derived word counts.

The combined inflected and derived word counts for the isolation points are given in Table 3, and those for the recognition points, in Table 4. The top panel of each table gives the counts when the sensory input alone is taken into account. The bottom panel gives the counts when contextual constraints also are taken into account. The tables clearly show that, in all conditions, the counts are large when based upon the sensory input alone, and are considerably reduced when candidates that are contextually inappropriate are eliminated. Although the data in the bottom panel of Table 4 provide the best fit with the predictions made by the cohort theory, the counts for the no-context and the no-semantic/weak-syntax conditions, in particular, are still a relatively poor fit when compared with the results for the base morpheme counts.

The base morpheme counts for the isolation points are given in Table 5. As the top panel of the table shows, there are a large number of base morphemes which are compatible with the acoustic-phonetic input in the minimal-semantic-context conditions. These numbers are considerably smaller in the no-context and

the no-semantic-context conditions because the isolation points occur later in the word. When these word candidates are assessed against the demands of the context, their number is considerably reduced. This is especially clear for the minimal semantic context conditions. The mean number of word candidates remaining at the isolation points, averaging across conditions, was 1.6, including the target base. This value is close to the cohort theory's predicted value of 1.0, and represents a considerable reduction over the mean of 5.1 obtained in the acoustic-phonetic count.

A similar analysis was performed on the recognition data. Here the mean number of word candidates that matched the acoustic-phonetic input was somewhat smaller at each recognition point (Table 6) than at the isolation points, reflecting the fact that some word candidates that had been available at the earlier isolation points were now no longer compatible with the sensory input and had dropped out of the cohort. The mean number of base morphemes compatible with the sensory input alone ranged from 2.9 to 1.1, with a mean of 1.8. When we evaluate the contextual appropriateness of these remaining word candidates, we find that a large proportion of them drop out as being either syntactically and/or semantically inappropriate. Moreover, we also obtain a better fit with the predictions of the cohort theory than we do with the corresponding isolation point analysis. In fact, the results of the base morpheme count are almost exactly those predicted by the cohort theory. That is, the theory predicts that at the recognition point there should be exactly one member of the initial cohort that is still compatible with both contextual and sensory constraints, and we find that the estimated number for each condition is only fractionally larger than this.

Moreover, the goodness of fit we find here between

Table 3
Mean Number of Inflected and Derived Word Forms Remaining at Isolation Points in Each Experimental Condition

No Context	Minimal Semantic/ Strong Syntax	No Semantic/ Strong Syntax	Minimal Semantic/ Weak Syntax	No Semantic/ Weak Syntax
Based on Sensory Input Alone				
11.7	54.5	12.4	37.5	9
Based on Sensory Input and Contextual Constraints				
11.7	1.6	1.9	2.8	5.4

Table 4
Mean Number of Inflected and Derived Word Forms Remaining at Recognition Points in Each Experimental Condition

No Context	Minimal Semantic/ Strong Syntax	No Semantic/ Strong Syntax	Minimal Semantic/ Weak Syntax	No Semantic/ Weak Syntax
Based on Sensory Input Alone				
5.7	16.6	6	16.9	5.9
Based on Sensory Input and Contextual Constraints				
5.7	1.4	1.2	2.1	3.6

Table 5
Mean Number of Base Morphemes Remaining at Isolation Point in Each Experimental Condition

No Context	Minimal Semantic/ Strong Syntax	No Semantic/ Strong Syntax	Minimal Semantic/ Weak Syntax	No Semantic/ Weak Syntax
Based on Sensory Input Alone				
2.1	10.1	2.1	6.8	1.5
Based on Sensory Input and Contextual Constraints				
2.1	1.5	1.5	1.8	1.5

Table 6
Mean Number of Base Morphemes Remaining at Recognition Point in Each Experimental Condition

No Context	Minimal Semantic/ Strong Syntax	No Semantic/ Strong Syntax	Minimal Semantic/ Weak Syntax	No Semantic/ Weak Syntax
Based on Sensory Input Alone				
1.2	2.8	1.1	2.9	1.1
Based on Sensory Input and Contextual Constraints				
1.2	1.2	1.0	1.4	1.1

the estimated and predicted counts cannot be dismissed on the grounds that it merely reflects frequency effects—that is, that the target was more likely to be produced because it was of a higher frequency in the language than those other competing word candidates which were not related to the target base. A breakdown of the incorrect candidates produced by subjects revealed that half of them were of higher frequency than the target and half were of lower frequency.

Comparing the size of the set of remaining word candidates at the recognition and isolation points, we see that the recognition point provides a better measure of the point in a word at which the intersection of contextual constraints and sensory input have narrowed down the initial set of word candidates to a single member. Note that the average word length for the targets used in the experiment was 750 msec and that average recognition points in the various experimental conditions ranged from 381 to 490 msec. This means that only 50%-65% of the word had been heard when listeners recognized it with 80% confidence in their choice.

Finally, it is interesting to note that it is here in the cohort analysis that we see the effect of taking confidence into account, and not in the earlier analysis, in which the effects of syntactic and interpretative constraints on word-recognition processes were compared.

CONCLUSIONS

In the present experiment, we found a large contrast in the size of the facilitatory effects of syntactic and interpretative constraints on word-recognition processes. Measured precisely, and in a highly favorable situation, syntactic form-class constraints gave listeners little

advantage in terms of facilitating the earliness with which they recognized a word. Note that this weak effect of syntax was observed in situations in which syntactic form-class constraints were the strongest the language afforded. In general, syntactic constraints are much weaker.¹³

Can this weak effect of syntax be attributed to the fact that we used target words for which there were an unusually large set of word candidates compatible with the first two phonemes? Could such artificially large set sizes have functioned to minimize the potential effects of syntactic constraints, and if we had used smaller sets would we have found a larger discriminatory effect of syntax? There are two sources of evidence against this argument.

First, in our experiment there were, on average, 60 word candidates compatible with the first two phonemes of each target word. This is not an unusually large number, given that in American English (the only language for which such a count is available) the median set size for the first two phonemes of a word is 87 members (Marslen-Wilson, 1983).

Second, the argument that the effectiveness of syntactic constraints is a function of the size of the cohort is contradicted by the present data. Comparing the minimal-semantic/weak-syntax condition with the no-semantic/weak-syntax condition, the recognition point in the former occurs considerably earlier in the word (401 msec) than it does in the latter (490 msec). This means that the size of the set of word candidates compatible with the sensory input in the minimal-semantic condition is larger than in the no-semantic-context condition. Nevertheless, in both cases we find the same reduction in the amount of sensory input required for recognition as a result of strong syntactic constraints. Therefore, the lack of a strong effect of

syntactic constraints in the present experiment cannot simply be due to the size of the set of word candidates.

The reason why we find such a small effect of syntactic form-class constraints is because such constraints can, at best, only reduce the size of the initial cohort to those members which belong to a single form-class category. They can rarely narrow down the set to a single member. Even in those rare cases in which, in principle, syntax could distinguish a unique candidate on the basis of form class—as in the case of systematic ambiguities (e.g., watch)—it does not seem to do so immediately. As Seidenberg et al. (1982) have shown, the effects of syntactic form-class disambiguation can be detected only about 200 msec after a word has been identified.

If syntax has such a limited function in reducing the amount of sensory input needed for the identification of a word, what role does it play in the processing system? With respect to word recognition, we propose that syntax imposes constraints on permissible form-class categories and on word endings, such as inflections and derivations. One rather speculative proposal is that although syntactic constraints do not significantly facilitate identification of the stem of a word form, they do function to reduce the amount of sensory input needed for recognition of its suffixes. Where syntax does play a major and essential role is in language understanding in its wider sense. That is, without syntactic information, a meaningful interpretation of an utterance cannot normally be developed. Perhaps one moral of the present study is that an information source that is important in one aspect of the system does not need to be important in every aspect.

In contrast to the restricted role of syntactic form-class constraints in word recognition, even a weak semantic context produced a large facilitatory effect. This is because a semantic context provides constraints which, when they intersect with constraints provided by the sensory input, facilitate the rapid reduction of the initial cohort to a very small set, which, in the limit, consists of a single member.

The large facilitatory effect of semantic context suggests that the semantic properties of word candidates are available to listeners early in the processing of a word. If this had not been the case, there would have been no basis for contextual facilitation to have had its effect. This early availability of the semantic properties of words is exactly what is predicted by the cohort theory. The model claims that when word candidates are initially activated on the basis of the bottom-up input, their syntactic and semantic properties are also accessed and form the basis for the continuing evaluation of each word candidate against the demands of the syntactic and interpretative context. What the experiment also shows is that the semantic basis for the early evaluation of the semantic suitability of each word candidate cannot be explained simply in terms of interlexical semantic associations. The context sentences used in the experiment specifically did not contain

words that were semantic associates of the target word. Therefore, the interpretative representation against which word candidates are assessed primarily involves sentential rather than interlexical meaning.

These results also bear on the claim made, for example, by Grosjean (1980) and Morton (1969) among others, that semantic context functions to preselect a subset of word candidates even before listeners have heard any of the sensory signal. In the present experiment, semantic context did not make a specific target word predictable in the absence of any sensory input (as measured by the cloze pretest), but did have a large facilitatory effect after some acoustic input had been heard. These data, then, are at least consistent with a view against preselection of word candidates and in favor of word recognition being the result of the intersection of sensory and contextual constraints, given an initial set specified solely from the bottom up (Marslen-Wilson & Tyler, 1980).

Also, the present analyses of the recognition point data clearly provide good support for the predictions made by the cohort theory concerning the point at which a word will be recognized when it is heard in and out of context. While it is not the case that the present results are incompatible with other theories, none of these theories makes similarly precise a priori predictions about the selection processes involved, with respect to the availability of the acoustic input over time.

The differences in the extent to which the isolation and recognition points fitted the cohort theory's predictions require explanation. We propose that the discrepancy here reflects an artifact generated by the gating paradigm. The task encourages the listener to guess the word on the basis of as little sensory input as possible. Consequently, listeners produce the target word even when they are uncertain of their choice. This shows up in the gating task as the isolation point—that is, the point at which listeners identify the target but are unsure about their choice. In contrast, at the recognition point, listeners are confident about their choice, and this is also the point at which the target word is essentially the only remaining member of the initial cohort that is compatible with sensory and contextual information. Therefore, we suggest that the isolation point does not reflect a distinct phase in the processing of spoken words.

Finally, we end on a speculative note. It is clear from the estimated counts of possible word candidates at the isolation and recognition points that the base morpheme counts fit the predictions made by the cohort theory better than do the combined inflected and derived word-form counts—especially for the no-context and the no-semantic-context/weak-syntax conditions. One possible interpretation of this finding is that words are represented in the mental lexicon as base morphemes with inflectional and derivational markers attached to them. This implies that when the base morpheme is accessed, all inflectional and derivational variations of that morpheme are also accessed. Our hypothesis is that when a

word is heard in isolation the relative frequency of these different forms will play a part in determining which specific form is selected, with infrequent forms being "less good" candidates than forms of higher frequency. Thus, subjects will assume that they are hearing the most frequent forms of the base morpheme—such as first person singular, nouns, or infinitives—until they receive sensory information to the contrary. When words are heard in context, we propose that word frequency will play a less important role in selecting among different word forms derived from the same base, and that contextual constraints—in particular, syntactic—will serve this function.

REFERENCE NOTE

1. Tyler, L. K. *Sequential analysis processes in spoken word recognition*. Manuscript in preparation, 1983.

REFERENCES

- BLANK, M. A., & FOSS, D. J. Semantic facilitation and lexical access during sentence processing. *Memory & Cognition*, 1978, 6, 644-652.
- CAIRNS, H., COWART, W., & JABLON, A. Effects of prior context upon the integration of lexical information during sentence processing. *Journal of Verbal Learning and Verbal Behavior*, 1981, 20, 445-453.
- COLE, R., & JAKIMIK, J. Understanding speech: How words are heard. In G. Underwood (Ed.), *Strategies of information processing*. New York: Academic Press, 1978.
- COLE, R., & JAKIMIK, J. A model of speech perception. In R. A. Cole (Ed.), *Perception and production of fluent speech*. Hillsdale, N.J.: Erlbaum, 1980.
- FORSTER, K. Levels of processing and the structure of the language processor. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing*. Hillsdale, N.J.: Erlbaum, 1979.
- GROSJEAN, F. Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, 1980, 28, 267-283.
- KALIKOW, D. N., STEVENS, K. N., & ELLIOTT, L. L. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America*, 1977, 61, 1337-1351.
- MARSLÉN-WILSON, W. D. Speech recognition as a psychological process. In J. C. Simon (Ed.), *Spoken language generation and recognition*. Dordrecht: Reidel, 1980.
- MARSLÉN-WILSON, W. D. Function and process in spoken word-recognition. H. Bouma & D.G. Bouwhuis (Eds.), *Attention and performance: Control of language processes*. Hillsdale, N.J.: Erlbaum, 1983.
- MARSLÉN-WILSON, W. D., & TYLER, L. K. Processing structure of sentence perception. *Nature*, London, 1975, 257, 784-786.
- MARSLÉN-WILSON, W. D., & TYLER, L. K. The temporal structure of spoken language understanding. *Cognition*, 1980, 8, 1-71.
- MARSLÉN-WILSON, W. D., & WELSH, A. Processing interactions and lexical access during word-recognition in continuous speech. *Cognitive Psychology*, 1978, 10, 29-63.
- MILLER, G. A., HEISE, G., & LICHTEN, W. The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, 1951, 41, 329-335.
- MILLER, G. A., & ISARD, S. Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior*, 1963, 2, 217-228.
- MORTON, J. Interaction of information in word recognition. *Psychological Review*, 1969, 76, 165-178.
- MORTON, J., & LONG, J. Effect of word transitional probability on phoneme identification. *Journal of Verbal Learning and Verbal Behavior*, 1976, 15, 43-51.
- ODEN, G. C., & SPIRA, J. L. Influence of context on the activation and selection of ambiguous word senses. *Quarterly Journal of Experimental Psychology*, 1983, 35A, 51-64.
- QUIRK, R., & GREENBAUM, S. *A university grammar of English*. London: Longman, 1979.
- SEIDENBERG, M., TANENHAUS, M., LEIMAN, J., & BIENKOWSKI, M. Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, 1982, 14, 489-537.
- SWINNEY, D. Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 1979, 18, 645-660.
- TANENHAUS, M., LEIMAN, J., & SEIDENBERG, M. Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior*, 1979, 18, 427-440.
- UNDERWOOD, G. Contextual facilitation from attended and unattended messages. *Journal of Verbal Learning and Verbal Behavior*, 1977, 16, 99-106.

NOTES

1. The terms "interpretative" and "semantic" are used interchangeably throughout the text. They refer to the combination of linguistic and nonlinguistic knowledge necessary to interpret an utterance in terms of the listener's knowledge of the world and of the discourse context.

2. This is an estimate of the minimum amount of sensory input required to initiate a cohort and is based upon measurements of word-recognition times obtained from a variety of word monitoring and shadowing studies.

3. We are concerned here only with research in the domain of spoken language understanding. Therefore, we are specifically omitting any discussion of those experiments dealing with written language. We take this position because of the inherent temporal differences between the two modalities and the lack of any principled arguments for why the processes underlying recognition should be assumed to be the same in both modalities.

4. We chose semantic contexts that were minimally constraining in order to discourage listeners from developing a strategy of trying to predict the target. This does not imply that listeners normally try to predict the words a speaker will utter, but rather that it is possible to develop such a strategy when faced with particular experimental paradigms.

5. An adverbial was also possible in a few cases, but produced an extremely unusual construction in Dutch, given the verbs we used.

6. A literal English gloss of this example set follows. The syntactic constraints employed in the Dutch materials do not function in the same way in English.

Target word: to benefit.

- (a) The appointment with the dentist goes not through. John tries to ...
- (b) The appointment with the dentist goes not through. John can ...
- (c) The breath with the lie shuffles only through. The terrace tries to ...
- (d) The breath with the lie shuffles only through. The terrace will ...

7. One item had to be omitted from the analyses due to experimenter error.

8. We should dispose of a possible objection to these results on methodological grounds. According to this objection, the small effect of syntax observed in the present experiment could be simply due to subjects' preferences for infinitival forms when they produce word candidates in this task. While it is certainly the case that subjects restrict themselves to a subset of the permissible form classes when producing their responses, they do

not show a strong preference for infinitival forms. In fact, in the no-context condition, we find that subjects are more likely to produce nouns than verbs: 43% of responses were nouns, 27% were infinitive verbs, and 11% were inflected verbs. Other form classes produced were adverbs, adjectives, pronouns, and prepositions, but of these only the percentage of adjectives produced was higher than 3% (i.e., 12%).

9. Two tapes were constructed so that each subject heard a test item only once in either the no-semantic-context or the minimal-semantic-context/strong-syntax condition. The test items were excised from the original recordings and were irregularly interspersed with a number of filler items. Since subjects heard a maximum of only five words per test sentence, we minimized the possibility of overloading their memory and thus obtaining an incorrect record of what they had actually heard.

10. *Isolation point data*: (a) ANOVA on five experimental conditions, $F(4,64) = 13.95$, $p < .001$. (b) ANOVA on difference scores—syntax, $F(1,16) = 5.89$, $p = .027$; semantics, $F(1,16) = 22.4$, $p < .001$.

Recognition point data: (a) ANOVA on five experimental conditions, $F(4,64) = 14.3$, $p < .001$. (b) ANOVA on difference scores—syntax, $F(1,16) = 5.51$, $p = .032$; semantics, $F(1,16) = 21.84$, $p < .001$.

11. *Isolation point data*: (a) ANOVA on five experimental conditions, $F(4,24) = 2.181$, $p = .102$. (b) ANOVA on difference scores—syntax, $F(1,6) = .0038$, $p = .953$; semantics, $F(1,6) = 7.354$, $p = .035$.

Recognition point data: (a) ANOVA on five experimental conditions, $F(4,24) = 4.454$, $p = .008$. (b) ANOVA on difference scores—syntax, $F(1,6) = .0012$, $p = .97$; semantics, $F(1,6) = 14.54$, $p = .009$.

12. We ignored morphological compounds because of problems in determining how they were represented in the lexicon.

13. It is interesting that this pattern of results also holds when listeners are forced to produce their word choices rapidly. We carried out a version of this experiment, using the same materials, in which subjects produced their word choices as quickly as possible and their latencies were recorded. With latencies averaging 450 msec, exactly the same relative difference in the degree of facilitation due to syntactic and semantic constraints was observed (Tyler, Note 1).

(Manuscript received March 8, 1983;
revision accepted for publication August 8, 1983.)