

The Time-Limited Influence of Sentential Context on Function Word Identification

Petra van Alphen and James M. McQueen
Max Planck Institute for Psycholinguistics

Sentential context effects on the identification of the Dutch function words *te* (to) and *de* (the) were examined. In Experiment 1, listeners labeled words on a [tə]–[də] continuum more often as *te* when the context was *te* biased (*Ik probeer [?ə] schieten* [I try to/the shoot]) than when it was *de* biased (*Ik probeer [?ə] schoenen* [I try to/the shoes]). The effect was weaker in slower responses. In Experiment 2, disambiguation began later, in the second word after [?ə]. There was a weak context effect only in the slower responses. In Experiments 3 and 4, disambiguation occurred on the word before [?ə]: There was no context effect when one set of sentences was used, but there was an effect (larger in the faster responses) when more sentences were used. Syntactic processing affects word identification only within a limited time frame. It appears to do so not by influencing lexical access processes through feedback but, instead, by biasing decision making.

When English listeners hear the ambiguous word [?ə], which is somewhere between *to* and *the*, they tend to identify it as *to* in the context *We tried [?ə] go* but as *the* in the context *We tried [?ə] gold* (Isenberg, Walker, & Ryder, 1980). It might be tempting to think that this demonstration of contextual involvement in perceptual decision making is evidence for feedback in the speech recognition system. This context effect might suggest that there are feedback mechanisms that act so that higher levels of processing (those responsible for processing the sentential information) influence lower levels of processing (those responsible for the initial acoustic–phonetic processing necessary to provide the higher levels with an input). We argue, however, that sentential context effects are not evidence for feedback. Indeed, the analysis of the time course of such effects presented here suggests instead that there is no feedback in the speech recognition system.

Why, then, are context effects not evidence for feedback? We first illustrate our answer to this question with a context effect in speech perception that has been studied quite intensively: the demonstration of lexical involvement in phonetic categorization (Burton, Baum, & Blumstein, 1989; Burton & Blumstein, 1995; Connine, 1990; Connine & Clifton, 1987; Fox, 1984; Ganong, 1980; McQueen, 1991; Miller & Dexter, 1988; Pitt, 1995; Pitt & Samuel, 1993). Ganong (1980) was the first to show that listeners tend to label ambiguous phonemes (e.g., alveolar stops on a continuum between [d] and [t]) in the initial position of word–nonword (e.g., *deep–teep*) and nonword–word (e.g., *deach–teach*)

contexts so as to form words (e.g., as [d] in *deep–teep* but as [t] in *deach–teach*). Similar effects have been observed when the phonemes that listeners classify are word medial (e.g., [d] and [g] in *cradle–cragle* and *badel–bagel*; Connine, 1990) or word final (e.g., [s] and [ʃ] in *kiss–kish* and *fiss–fish*; McQueen, 1991).

One account of these effects is that there is a feedback loop in the speech recognition system. In most accounts of spoken word recognition, lexical knowledge is accessed through *prelexical* acoustic–phonetic processes that map the speech signal onto stored lexical representations. A feedback loop would allow lexical processes to modulate those prelexical processes. The top-down word-to-phoneme connections in the TRACE model (McClelland & Elman, 1986) instantiate just such a feedback loop. Lexical effects in categorization could be taken as support for this kind of feedback mechanism. The ambiguous signal [?ip], for example, activates the lexical representation of the word *deep*, which in turn feeds support back to the prelexical representation of the phoneme [d]. The result is more [d] responses in the [?ip] context than in the [?itʃ] context (in which feedback produces a bias toward [t]).

There is, however, no need to postulate that there is feedback in the speech recognition system to be able to explain lexical effects in phonetic categorization (Massaro & Oden, 1995; McQueen, 1991). Lexical knowledge can influence perceptual decision making without influencing prelexical processing. In the *Merge model* (McQueen, Norris, & Cutler, 1999; Norris, McQueen, & Cutler, 2000), phonemic decisions are made at a dedicated decision-making stage that receives input from both prelexical and lexical levels of processing. Information from these two sources is integrated. The Merge model accounts for lexical effects in categorization as follows. The signal [?ip], for example, activates the decision units (nodes) for both [d] and [t] through the connections from the prelexical level to the decision level. The same signal also activates the word node for *deep*, which in turn feeds support forward to the decision node for the phoneme [d], biasing the decision toward [d]. There will thus be more [d] responses in a [?ip] context than in a [?itʃ] context (in which the [t] decision node receives support from the word node for *teach*).

Part of this research formed the *afstudeer* (master's) project of Petra van Alphen at the University of Utrecht under the supervision of James M. McQueen, Anne Cutler, and Hugo Quené. Anne Cutler suggested Experiment 1 to us. We would like to thank her not only for initiating this project but also for her advice throughout. We would also like to thank Hugo Quené for his help during the project.

Correspondence concerning this article should be addressed to Petra van Alphen or James M. McQueen, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, the Netherlands. Electronic mail may be sent to Petra.vanAlphen@mpi.nl or James.McQueen@mpi.nl.

It is therefore possible to account for lexical effects in phonetic categorization in a model without feedback. The Merge model can also explain the detailed pattern of lexical context effects in other tasks that require phonemic decisions and, in fact, can do so better than models such as TRACE that include feedback loops (McQueen et al., 1999; Norris et al., 2000; Pitt & McQueen, 1998). Perhaps the strongest argument in favor of models without feedback is that, in such models, context is unable to distort the mental representation of the perceptual world. Context can certainly influence the phonemic decisions that listeners make, but it does so without modulating the prelexical representation of the speech signal; the coding of what actually occurred in the physical world is left intact. In contrast, models with feedback run the risk of hallucinating (Massaro, 1989; Norris et al., 2000). Feedback from the lexicon acts to change the prelexical representation of the speech signal, and this, in the extreme case, could result in the listener hearing a speech sound that was in fact not present in the physical signal.

Therefore, because models without feedback can account for lexical context effects, these effects cannot be taken as support for models with feedback. Effects of sentential context on phonetic categorization (Borsky, Tuller, & Shapiro, 1998; Connine, 1987; Connine, Blasko, & Hall, 1991; Miller, Green, & Schermer, 1984), however, might appear to offer a stronger case for feedback. In Miller et al.'s (1984) study, for example, listeners heard the sentence frames *She needs hot water for the . . .* and *She likes to jog along the . . .* with, as the final word, a continuum between *bath* and *path*. Listeners were asked to identify the final word. There were more *bath* responses in the semantically congruent *hot water* context than in the semantically incongruent *jogging* context. At first glance, such a result appears to suggest that the output of the interpretative processes that compute the semantic plausibility of a word in the overall meaning of a sentence can influence lower level processes, either lexical-level processes (biasing the activation of a word such as *bath*) or even prelexical processes (biasing the activation of a representation of the phoneme [b]). Is this not evidence of feedback?

The answer to this question is again no. Exactly the same argument applies in the case of sentential context effects as in the case of lexical context effects. Sentential context can influence the decision-making process, through which listeners make explicit perceptual judgments, without also influencing the lexical access process. On such a view, semantic congruity biases the decision process such that a congruous word tends to be selected rather than an incongruous word. But this bias need not influence word representations at the lexical level or the representations of the component phonemes of those words at the prelexical level. In other words, there is no need for feedback.

Detailed analysis of sentential context effects has in fact suggested that these effects are not caused by feedback. Miller et al. (1984) showed that the effects are susceptible to the demands of the experimental task. The effect was present when listeners had to identify both the final word of each sentence and the sentence frame itself but was absent when listeners had to identify only the final word of each sentence. This finding suggests that sentential context influences the decision process (and thus depends on the nature of that decision) and that sentential context does not exert a mandatory influence on the recognition process (and thus at least

rules out a strong version of a feedback account in which the feedback loop is always active).

Connine (1987) examined the reaction times (RTs) of categorization decisions in sentential contexts. She found that responses were faster when they were consistent with the sentence context (e.g., a [d] response to the sentence *She drives the car with the* [ʔent]) than when they were inconsistent (e.g., a [d] response to the sentence *She saw the show in the* [ʔent]). This RT consistency effect, however, was found only for endpoint stimuli (e.g., when the to-be-categorized phoneme was an unambiguous [d] or an unambiguous [t]), which most listeners labeled correctly most of the time. It was not found for responses in the most ambiguous region (at the *category boundary*), even though there was a sentential context effect (more consistent than inconsistent responses) at that boundary. This pattern of results is very similar to that obtained when categorization decisions are biased by a monetary payoff scheme (e.g., listeners are told they will receive a monetary reward for correct [d] decisions; Connine & Clifton, 1987). Connine (1987) therefore argued that sentential context effects are also the result of a bias, and thus that sentential context influences perceptual decision making rather than the process of speech encoding. Connine's findings are consistent with those of Samuel (1981), who examined the effect of sentential context on the phonemic restoration illusion. In a signal-detection analysis, Samuel (1981) showed that sentential context influenced the bias measure β and also argued that sentential context influences the decision-making stage rather than the encoding stage.

Thus, the results of Samuel (1981), Miller et al. (1984), Connine (1987), and Connine et al. (1991) all lead to the same conclusion: Sentential context influences decision making rather than encoding. This consensus, however, has recently been challenged by Borsky et al. (1998), who argued that sentential context does influence the encoding of speech. The results of Borsky et al. (1998) contradict those of Connine (1987). Borsky et al. presented listeners with sentences such as *The laughing dairyman hurried to milk the* [ʔot] *in the drafty barn* and visual probes (*goat* or *coat*) at the offset of the target word ([ʔot]). Listeners were required to decide whether the visual word matched what they had just heard. Borsky et al. found a sentential context effect (i.e., more *goat* identifications in *goat*-biased sentences), but, in contrast to Connine (1987), they also found that RTs at the category boundary were faster for responses that were consistent with the sentence context than for responses that were inconsistent with the sentence context. Borsky et al. argued that their results may have been different from those of Connine (1987) because of several procedural factors, such as the number of sentence contexts.

On the basis of their results, Borsky et al. (1998) claimed that sentential context influences initial phonological analysis. They thus argued for a form of feedback in the speech recognition system by which sentence-level processes modulate phonological processes (although they did not in fact use the term *feedback*). Their argument was that, because sentential context effects show the same RT pattern as lexical context effects, and because lexical effects are due to feedback, sentential context effects are also due to feedback. This argument is unconvincing. First, there is no standard RT pattern for lexical effects with which sentential effects can be compared. The sentential effects observed by Borsky et al. (1998) showed the same RT pattern as lexical effects in word-initial position (e.g., in *deep-teep* and *deach-teach* continua; Con-

nine & Clifton, 1987; Pitt & Samuel, 1993); that is, consistent responses were faster than inconsistent responses at the category boundary but not at the continuum endpoints. This is not the same pattern as is observed for lexical effects in word-final position (e.g., in *kiss-kish* and *fiss-fish* continua; McQueen, 1991; Pitt & Samuel, 1993). Word-final lexical effects in fact show the same RT pattern as that observed for sentential context effects by Connine (1987); that is, consistent responses are faster than inconsistent responses at the continuum endpoints but not at the category boundary. Thus, neither sentential nor lexical context effects show a consistent RT pattern. Second, as we have already argued, lexical effects are not necessarily due to feedback. The RT effect at the category boundary observed by Borsky et al. (1998) could be due to decision-level processes. Semantic congruity could have biased a decision process, not only such that congruous words tended to be selected rather than incongruous words but also such that congruent decisions were made more quickly than incongruent decisions.

The aim of the present experiments was to distinguish between these two accounts of sentential context effects: the account based on a decision bias (Connine, 1987; Samuel, 1981) and the account based on feedback (Borsky et al., 1998). We therefore examined the time course of sentential context effects through the use of RT range analyses in which categorization responses were divided into fast, medium, and slow RT bins. These analyses have been used quite extensively to examine the fine detail of the dynamics of lexical context effects (Burton et al., 1989; Burton & Blumstein, 1995; Fox, 1984; McQueen, 1991; Miller & Dexter, 1988; Pitt, 1995; Pitt & Samuel, 1993).

We reasoned that if sentential context effects are indeed due to feedback processes, they should be relatively stable over time. Simulations with the TRACE model reported by McClelland and Elman (1986, p. 26, Figure 6) and by McClelland (1987, p. 12, Figure 1.2) show how, in the case of lexical context effects, feedback builds up over time. In the McClelland (1987) simulations, for example, the input was /dar?/, where /?/ was an ambiguous phoneme midway between /d/ and /t/. It can clearly be seen that, although the top-down influence of the word *dart* takes some time to build up, as more time elapses, the activation of the node for the lexically consistent phoneme /t/ continues to rise and that for the /d/ levels off and then falls. The phoneme nodes thus tend to reflect the pattern of activation at the lexical level. As long as *dart* remains activated, /t/ will therefore be the dominant phonemic response. In other words, the lexical effect rises over time and then remains stable (or may even continue to increase in magnitude); it does not die away over time. We assumed that the same would be true for sentence context effects. That is, if they are due to feedback, they should take time to build up (higher level processes may need time to complete), but they should not die rapidly away. Once feedback from the sentential level of processing has acted to modify phonological representations (at either the lexical or pre-lexical level), those modifications should continue to influence responses made after that time.

If, on the other hand, sentential context effects are the result of a decision bias, predictions about their time course are much less clear. The effects could have a limited time course or could be very stable over time, or they could increase or decrease over time. The reason is that, although a decision bias can influence responses at one moment in time, this influence has no necessary effect on

responses later in time. Because the decision bias does not modify the bottom-up encoding of speech, the coding of the information in the signal remains available to influence later processing and, thus, influence decisions made later in time. In a model with feedback, however, the bottom-up information tends to be overwritten by the feedback from higher levels; this is why feedback models predict that context effects should not die away over time.

We also examined the time course of the effect at a more macroscopic level, through manipulation of the relative position of the biasing information in the sentential context and the target words. This analysis followed that of Connine et al. (1991). They asked listeners to categorize word-initial phonemes in continua such as *dent-tent*. However, in contrast to other studies on sentential context effects, the biasing sentential information appeared after the target word (e.g., in *When the [?ɛnt] in the fender was well camouflaged, we sold the car*). Connine et al. manipulated the distance between the target word and the biasing information. They found a sentential context effect when this information appeared three syllables after the target word but not when it appeared six to eight syllables after the target. In the present study, we sought to examine the temporal limitations of such effects further. As we discuss later, the temporal pattern that we observed places strong constraints on theoretical accounts of the effect.

We chose function words as the target words for this investigation. We believed that function words were likely to provide us with the strongest sentential context effects. The studies reviewed earlier used content words as targets and manipulated primarily semantic congruity effects. Although in some cases semantic congruity can exert a strong effect (it is a rare coat that can be milked; Borsky et al., 1998), the effect is not always absolute (if a tent is in a trailer, one can drive the car with the tent; Connine, 1987). Our use of function words as targets allowed us to manipulate the absolute syntactic requirements for certain function words in certain syntactic frames. In fact, the study of Isenberg et al. (1980) with which we began this article looked at exactly this case: *We tried [?ə] gold* versus *We tried [?ə] go*. One may predict very strong effects in this situation. If the context demands that *the* be present (and not *to*), not only for the sentence to make sense but also for it to be grammatical, why would one want to identify that word in any other way? Unfortunately, the study by Isenberg et al. remains unpublished, so the size of the effect they observed is unclear. Furthermore, it seems that they did not analyze the time course of the effect.

Experiment 1

Experiment 1 was therefore an attempt to replicate the Isenberg et al. (1980) study but with the addition of a time-course analysis. The experiment was carried out in Dutch, in which a function word continuum from *te* (to) to *de* (the) could be constructed by manipulating the voicing of the initial stop consonant: [tə]–[də]. We selected an unambiguous plural noun, *schoenen* (shoes), that can be preceded by the definite article *de* but not by the infinite marker *te*; an unambiguous verb infinitive, *schieten* (to shoot), that can be preceded by *te* but not by *de*; and another word, *schaatsen*, that can be either a plural noun (skates) or an infinitive verb (to skate) and thus can be preceded by both *de* and *te*. These three words were then placed in the same sentence context:

1. *Ik probeer te/de schoenen*
(I try to/the shoes; *de* biased)
2. *Ik probeer te/de schieten*
(I try to/the shoot; *te* biased)
3. *Ik probeer te/de schaatsen*
(I try to/the skate[s]; ambiguous).

We predicted that there would be a strong sentential context effect (more *te* decisions in the second sentence than in the first). As described earlier, we also predicted that if this effect were due to feedback, it would not become weaker in slower responses. If anything, an effect due to feedback should build up over time, as more contextual information becomes available. The final sentence was ambiguous; both the plural noun phrase reading (with *de*) and the infinitive reading (with *te*) are acceptable. We therefore predicted that there would be more *te* decisions in the final sentence than in the first sentence, but fewer than in the second sentence.

Method

Participants. Twenty-six volunteers from the Max Planck Institute participant pool were paid to take part in this experiment. All were native speakers of Dutch, and none reported any hearing loss.

Stimulus construction. The three sentences were recorded several times by Petra van Alphen (a female native speaker of Dutch) in a sound-damped booth onto digital audio tape (sampling at 48 kHz with 16-bit resolution). The utterances were digitized at a sample rate of 16 kHz through the ESPS speech editing system with Xwaves. One production of the function word *te* and one production of the function word *de* were excised and used to create an eight-step continuum from *te* to *de*. The difference between a Dutch [t] and [d] is mainly caused by a difference in voice-onset time (VOT) and by a difference in the duration of the burst: A [d] has a negative VOT and a short burst, whereas the burst for a [t] is longer and the VOT is positive. Step 1 of the continuum (a clear [tə]) consisted of 89 ms of silence, the release burst and following noise from the [t] (34.1 ms), and then the schwa of the excised *de* (36.6 ms). For each subsequent step along the continuum, the noise portion was reduced by 3.8 ms, and the silence was reduced by 11 ms (except for Step 8, in which the silence was reduced by 22 ms to obtain a [d] with no silence). From Step 5 onward, an additional voiced period of 4.5 ms excised from *de* was inserted between the silence and the release burst. That is, Step 8 had no silence, four periods of voicing, and 7.5 ms of noise followed by the schwa. To obtain smooth transitions between the combined parts, we spliced each part at a zero crossing.

All members of the continuum were placed in three different carrier sentences, each with the same beginning but with different endings. The beginning of each sentence (*ik probeer*) was acoustically identical. The ending was one of the three words *schoenen*, *schieten*, or *schaatsen*. In this way, two biased sentences and one ambiguous sentence were obtained. The combination of these three carrier sentences and the eight members of the continuum resulted in 24 sentences. All sentences sounded natural, and none contained any audible discontinuities.

Procedure. Sentences were presented binaurally over headphones at a comfortable listening level. Every sentence was presented 20 times to each listener. Two randomized lists of all 480 sentences were constructed. One of the lists was presented to half of the listeners, and the other list was presented to the other half. Listeners were tested in separate carrels in a quiet room in groups varying from 1 to 4. They were asked to decide as fast as possible whether they had heard *te* or *de* by pressing one of two buttons of an appropriately labeled response box. Half of the listeners used their writing hand to press the *de* button, and the other half used their writing hand to press the *te* button. There was an interval of 1,500 ms between the sentences in which listeners had to make their decisions. RTs were measured from the beginning of the schwa of the function word, that is, the

point moving through each step of the [tə]–[də] continuum at which all eight steps became acoustically identical. Before the experiment began, there was a training session of 12 sentences; each sentence was presented carrying the clear stimuli from the endpoints of the continuum and two ambiguous stimuli from the middle of the continuum. Listeners were informed that some of the sentences could be ungrammatical. They were offered a short rest break in the middle of the experiment.

Results and Discussion

The percentage of voiceless responses was computed for each listener as a function of stimulus continuum and sentence bias. The identification functions, pooled across participants, are shown in the upper panel of Figure 1. The endpoints of the continuum were labeled uniformly, demonstrating that they were clear, unambiguous exemplars of the target words *te* and *de*. As is evident from Figure 1, the responses of the listeners were influenced by the sentence context, especially in the category boundary region. There were more *te* responses in the *te*-biased sentence and more *de* responses in the *de*-biased sentence. The ambiguous sentence fell between the other two sentences.

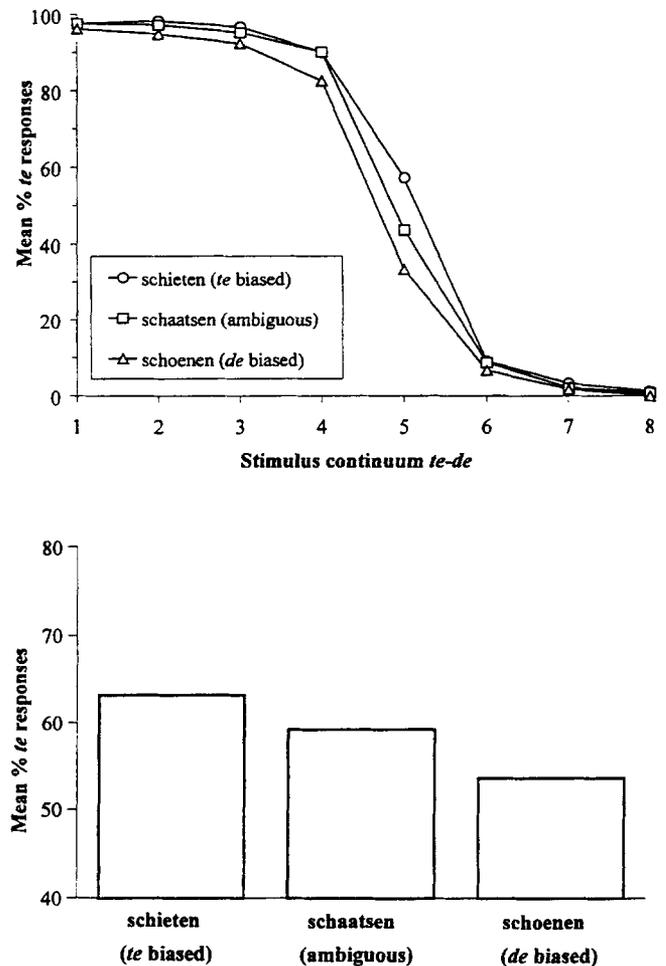


Figure 1. Mean percentage *te* responses in Experiment 1 in each of the three sentential contexts across the stimulus continuum (upper panel) and in the ambiguous region (Steps 3–6; lower panel).

The size of the shift of the curves was measured by calculating the area beneath each curve in the boundary region (see Pitt & Samuel, 1993). The boundary region was defined as extending from Step 3 to Step 6. For each sentence and for every participant, the proportion of *te* identifications for these four stimuli was calculated. The mean percentages of *te* responses in the ambiguous region in each sentence context are shown in the lower panel of Figure 1. Next, the proportions of *te* responses for each participant in each context were converted through an arcsine transformation (Studebaker, 1985). The mean values for the *de*-biased sentence, the *te*-biased sentence, and the ambiguous sentence were 1.64, 1.84, and 1.76, respectively. A one-way repeated measures analysis of variance (ANOVA) on these transformed proportions showed a significant effect of sentence bias, $F(2, 50) = 21.47, p < .001$. A post hoc Tukey honestly significant difference (HSD) test showed that all of the pairwise differences among the three conditions were significant at the .05 level.

Each listener's responses to each step along the stimulus continuum were ranked and divided into three groups: the fastest, medium, and slowest thirds. Mean RTs for these three groups were, respectively, 366 ms ($SD = 127$ ms), 500 ms ($SD = 155$ ms), and 714 ms ($SD = 270$ ms). As in the overall analysis, for each participant the proportion of *te* responses across Steps 3, 4, 5, and 6 was calculated for the sentences in each RT range. The mean proportions of *te* responses in the ambiguous region, for each sentence context and each RT range, are plotted in Figure 2. Each participant's proportional data in each condition and each RT range then underwent an arcsine transformation. The means of these values were 1.71 (*de* biased), 2.01 (*te* biased), and 1.84 (ambiguous) for the fast RT range; 1.64 (*de* biased), 1.86 (*te* biased), and 1.72 (ambiguous) for the medium RT range; and 1.59 (*de* biased), 1.69 (*te* biased), and 1.72 (ambiguous) for the slow RT range.

In the fast and medium ranges, there were significant effects of sentence bias, $F(2, 50) = 11.06, p < .001$, and $F(2, 50) = 7.44, p < .001$, respectively. The outcomes of Tukey HSD tests showed that all sentences differed significantly from each other in the fast range. Of the 26 participants, 22 (85%) made more *te* responses in the *te*-biased sentence than in the *de*-biased sentence. In the medium range, the *te*-biased sentence differed significantly from the *de*-biased sentence and from the ambiguous sentence. The difference between the *de*-biased sentence and the ambiguous sentence was not significant. Again, 22 participants produced more *te* responses in the *te*-biased sentence than in the *de*-biased sentence. In the slow range, the sentence bias effect was not significant. Only 16 participants (62%) made more *te* responses in the *te*-biased sentence than in the *de*-biased sentence. Thus, the effect was strongest in the fast RT range and became weaker in the slow RT range.

In summary, there were more *te* responses to the verb-final sentence than to the noun-final sentence, and responses to the ambiguous sentence were intermediate. Furthermore, the contextual effect changed over time: It was strongest in the fast RT range and became weaker as responses became slower (such that it was no longer statistically reliable in the slow RT range). One possible reason for this RT effect is that contextual information can be used by listeners only within a limited time window. If syntactic processing is complete before the identification decision is reached, syntactic bias may no longer be available to influence phonetic

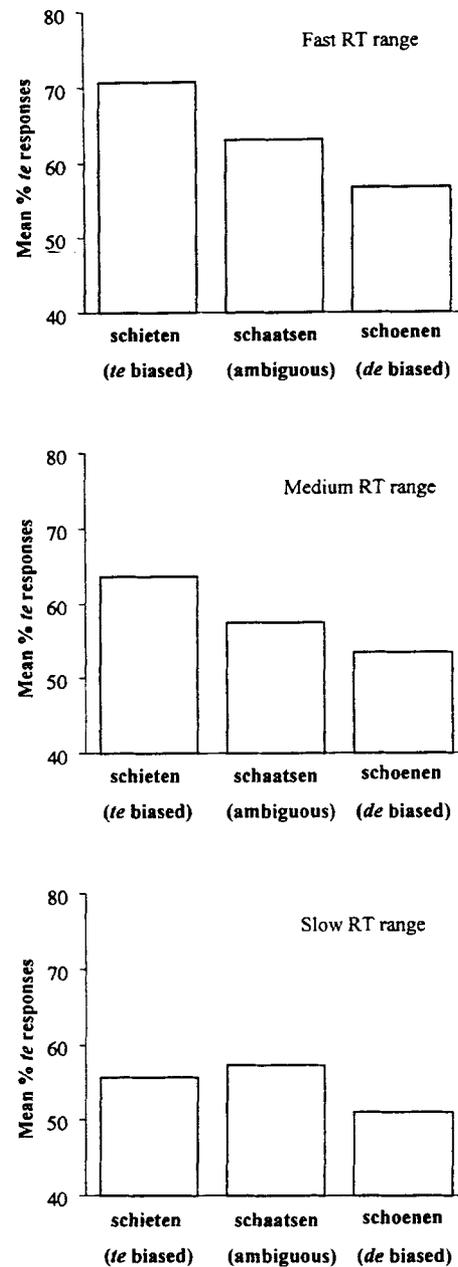


Figure 2. Mean percentage *te* responses in Experiment 1 in each of the three sentential contexts in the ambiguous region, plotted separately for each of the three reaction time (RT) ranges.

decisions. Note that it is not the case that the task became easier in slower responses; the ambiguous steps remained ambiguous (see Figure 2). This means that listeners could, in principle, still use the contextual information to bias their decisions, but they tended not to do so. Note also that it is not the case that the possible sentential bias becomes weaker in slower responses. By the time slow responses were being initiated, listeners had heard all of the sentences. Furthermore, the biases in these sentences did indeed influence the faster decisions. Thus, the weaker effects in slower responses must have occurred either because participants chose not

to use the information to help them with a difficult decision (which seems unlikely) or because the information simply was unable to influence the decision.

This account is inconsistent with the view that sentential context exerts its effect through feedback. If feedback had acted to modify lexical and phonological encoding processes (as the faster responses might suggest), then it would not be possible for the effects of this feedback to disappear in slower responses. After phonological representations have been altered, responses made on the basis of those representations should show the effects of that alteration whether they are fast or slow.

Another possibility is that slow responses reflect the operation of RT deadline processes. Participants are required to respond as fast as possible, that is, before some deadline. If no identification decision has been reached when the deadline occurs, particularly on the most ambiguous tokens of the stimulus continuum, responses may then be initiated at random. Such responses would tend to cancel out any contextual bias in the slow responses (see McQueen, 1991, for a similar argument about the disappearance of lexical context effects in slow word-final phonetic categorizations). Note, however, that the responses remained orderly in the slow range. Responses at the endpoints were more or less unanimously correct (*te* endpoint, 97% *te* responses on average; *de* endpoint, 1% *te* responses on average) and, as shown in Figure 2, those in the ambiguous region were close to chance. In other words, listeners were not responding at random in the slow RT range. It is therefore unlikely that the absence of a significant bias effect in the slow range was due only to the effects of deadline-based responses.

Experiment 2

If sentential context has no influence on slow responses because syntactic analysis is completed, it may also have no influence on fast responses when syntactic analysis cannot yet begin, that is, before disambiguating information has arrived. We therefore tested the time constraints of the syntactic bias further in Experiment 2, using sentences in which the disambiguation appeared later than in Experiment 1. We used the ambiguous word *schaatsen* from Experiment 1 but with three different following prepositional phrases. The first was biased toward the nominal reading of *schaatsen*, and the second was biased toward the verbal reading. The third phrase maintained the ambiguity. The three sentences were:

4. *We proberen te/de schaatsen van mijn broer*
(We try to/the skates of my brother; *de* biased)
5. *We proberen te/de schaatsen op noren*
(We try to/the skate on racing skates; *te* biased)
6. *We proberen te/de schaatsen zonder sokken*
(We try to/the skate[s] without socks; ambiguous).

We predicted that the contextual bias would build up over time; that is, no bias would be present in the fastest responses (before the disambiguation was processed), but bias would emerge in the slowest responses (once the disambiguation had been processed). It was, of course, possible that listeners would make all of their decisions without waiting for the following context. It was thus possible that there would be no contextual effects whatsoever.

Method

Participants. Twenty-six volunteers from the Max Planck Institute participant pool took part in this experiment, in exchange for pay. All were native speakers of Dutch without any hearing loss. None had participated in Experiment 1.

Stimulus construction. The three sentences were recorded and digitized through the same procedure as in Experiment 1. Because the continuum from *te* to *de* that was constructed for Experiment 1 was acoustically inappropriate in the Experiment 2 contexts, a new continuum had to be constructed. Other than differences in the exact duration of the different parts, stimulus construction was identical to that of Experiment 1, except that the amplitude of the voiced periods of the /d/ had to be attenuated (to obtain a smooth transition from the silent interval to the voiced part of the stimuli). Parts of the recorded sentences were combined such that the three carrier sentences were acoustically identical up to the end of *schaatsen*. The sentences differed only in their endings. As before, the eight steps of the continuum were placed in the three carrier sentences, resulting in 24 sentences.

Procedure. Except for the following changes, the procedure for Experiment 2 was identical to that of Experiment 1. Participants were tested in sound-damped booths. The 480 sentences were randomized in 20 blocks; each block contained the 24 different sentences in different random orders.

Results and Discussion

The identification curves for the responses in Experiment 2 are plotted in the upper panel of Figure 3. There are minimal visible differences among the three curves. The only effect is that the *de*-biased sentence is slightly shifted to the left relative to the other two sentences. As before, the proportion of *te* responses for Steps 3 to 6 of the stimulus continuum was calculated for each context and each participant (see the lower panel of Figure 3) and transformed by means of an arcsine transformation. The mean arcsine values were 1.75 (*de* biased), 1.79 (*te* biased), and 1.81 (ambiguous). An ANOVA revealed that the small bias effect was significant, $F(2, 50) = 3.31, p < .05$. A Tukey HSD test showed that only one shift was reliable: There were significantly fewer *te* responses in the *de*-biased sentence than in the ambiguous sentence.

The data for each participant were partitioned into thirds in the same way as in Experiment 1. The mean RTs of the responses in the fast, medium, and slow RT ranges were, respectively, 303 ms ($SD = 98$ ms), 424 ms ($SD = 125$ ms), and 663 ms ($SD = 315$ ms). The mean percentages of *te* responses in the ambiguous region for each context and each RT range are plotted in Figure 4. There were no differences among the three contexts in the fast RT range, but in the slow RT range there was a small difference between the *de*-biased sentence and the *te*-biased sentence: There were a few more *te* responses in the *te*-biased sentence than in the *de*-biased sentence. The mean arcsine-transformed proportions in each RT range were 1.92 (*de* biased), 1.92 (*te* biased), and 1.97 (ambiguous) for the fast RT range; 1.74 (*de* biased), 1.70 (*te* biased), and 1.79 (ambiguous) for the medium RT range; and 1.62 (*de* biased), 1.76 (*te* biased), and 1.68 (ambiguous) for the slow RT range. ANOVAs showed that there was no significant bias effect in any of the three RT ranges. However, in the slow RT range, the effect was almost significant, $F(2, 50) = 3.05, p = .056$; there was thus a trend in the expected direction. A Tukey HSD test showed that the difference between the *de*-biased sentence and *te*-biased sentence in the slow RT range was significant at the .05 level. The

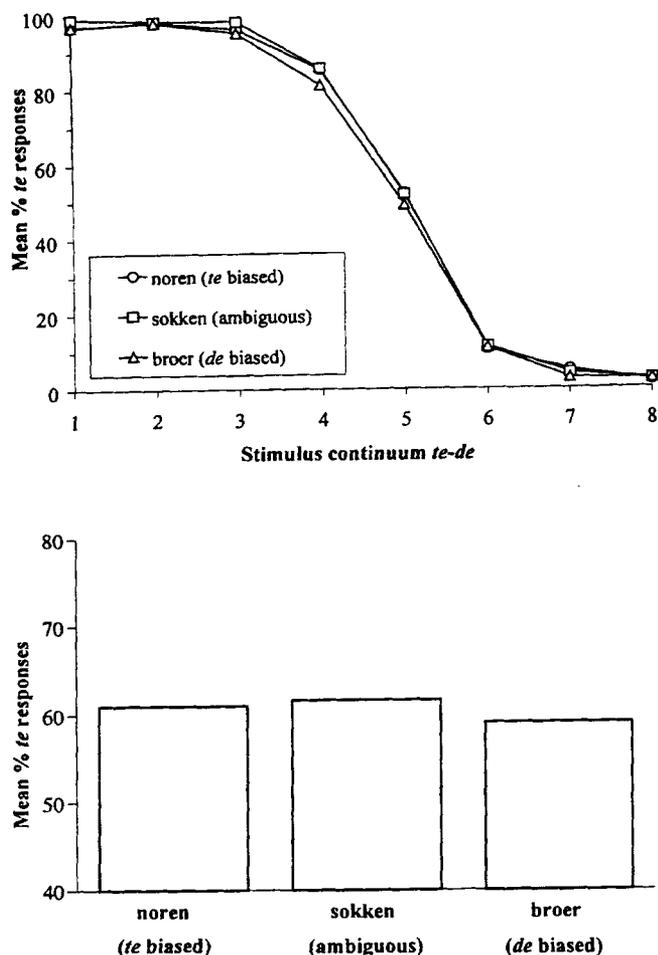


Figure 3. Mean percentage *te* responses in Experiment 2 in each of the three sentential contexts across the stimulus continuum (upper panel) and in the ambiguous region (Steps 3–6; lower panel).

numbers of participants showing a context effect (more *te* responses in the *te*-biased sentence than in the *de*-biased sentence) in each RT range were as follows: fast, 9 (35%); medium, 14 (54%); and slow, 18 (69%).

As was predicted, the sentential context effect was very small and was due only to responses in the slowest RT range. This finding suggests that the results of Experiment 1 were not due to deadline-based responses. The listeners in Experiment 2 were placed under the same amount of time pressure as those in Experiment 1, yet, in contrast to the Experiment 1 participants, they showed the strongest contextual bias effect in their slowest responses. The presence of an RT deadline therefore does not seem to play a large role in determining the strength of sentential bias effects.

The combined results of Experiments 1 and 2 suggest that syntactic bias has a limited time window of effectiveness. Although the listeners in Experiment 2 could have waited for the disambiguating context, they tended not to and so did not use sentential information in their faster responses. Likewise, although the listeners in Experiment 1 had already processed the full sen-

tence, they tended not to use the sentential information in their slower responses. The first of these temporal constraints almost certainly involves in part a strategic choice (i.e., a choice not to wait for the following information, even though listeners knew that it would be disambiguating in two thirds of the trials). The second constraint, however, seems not to involve a strategic choice: Why would one not use useful information to help make a difficult decision (especially when one does use that information in one's faster responses)? The temporal constraint that sentential context is

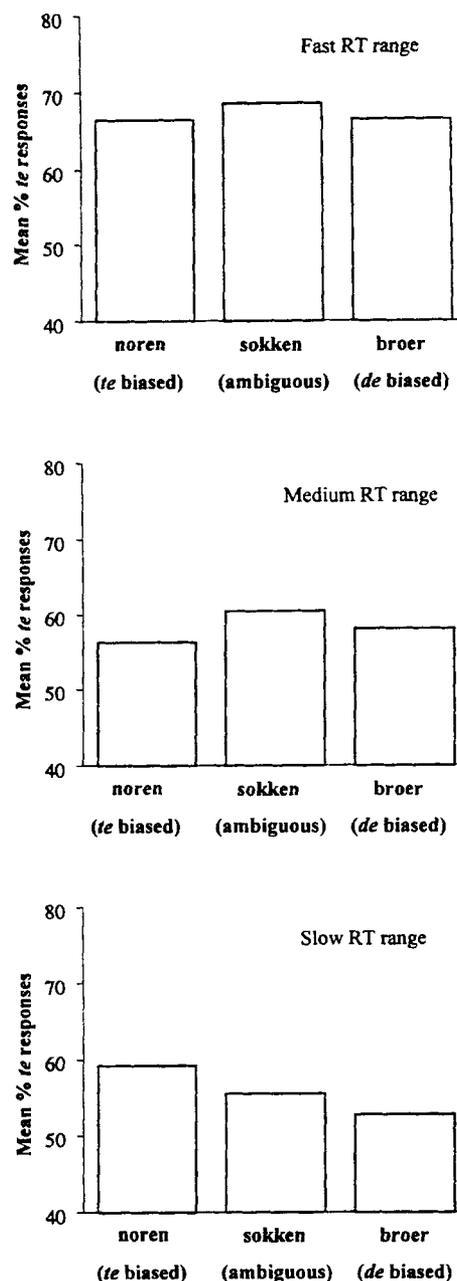


Figure 4. Mean percentage *te* responses in Experiment 2 in each of the three sentential contexts in the ambiguous region, plotted separately for each of the three reaction time (RT) ranges.

not used to bias slow decisions therefore appears to be a constraint of the automatic processing machinery; if syntactic processing is complete, it may be unable to bias identification decisions.

An interesting question therefore arises: What happens when the disambiguating context appears before the function words? If syntactic processing is completed before even the fastest responses can be initiated, there may be no effect (i.e., as tended to occur in the slow responses in Experiment 1). On the other hand, one might predict an effect in the faster responses that then dies out in the slower responses (i.e., as in Experiment 1). Yet another possibility, however, is that the bias effects will build up over time (as in Experiment 2); if syntactic processing takes time to complete, there may be no output available when the fastest decisions are made. A final possibility is that there will be strong and stable effects in all RT ranges. Because the disambiguating context precedes the function words, listeners are in a position to use it to maximum effect in their identification decisions.

Experiment 3

Experiment 3 therefore tested identification of *te* and *de* in three sentences in which the disambiguation appeared in the verb before the [ʔə] continuum:

7. *We verkopen te/de schaatsen*

(We sell to/the skates; *de* biased)

8. *We durven te/de schaatsen*

(We dare to/the skate; *te* biased)

9. *We proberen te/de schaatsen*

(We try to/the skate[s]; ambiguous).

After the verb *verkopen*, the definite article *de* is possible, but the infinitive marker *te* is ungrammatical. The only possible continuation with *te* would be that in which it is an intensifier (too), as in an adverbial phrase (*We verkopen te veel* [We sell too much]). But this reading of *te* is rendered ungrammatical in the context of the following plural-infinitive *schaatsen*. The disambiguation is therefore almost, but not entirely, due to *verkopen*. After the verb *durven*, both *de* and *te* are in fact possible, but the reading with *de* is made ungrammatical when the sentence ends with *schaatsen* (it would be possible to continue the sentence with a *de*, as in *We durven de schaatsen te verkopen* [literally, We dare the skates to sell]). Given that the sentences always ended with *schaatsen* and that *te* never appeared in the experiment as an intensifier, listeners were able to disambiguate [ʔə] on the basis of the preceding verbs in the first two sentences. The final sentence was ambiguous (as in Experiments 1 and 2, *proberen* can take both *de schaatsen* and *te schaatsen*).

Method

Participants. Twenty-six volunteers from the Max Planck Institute participant pool were paid for taking part. They were all native speakers of Dutch and did not report any hearing loss. None had taken part in the earlier experiments.

Stimulus construction. The three sentences for this experiment were recorded at the same time as the materials for Experiment 2. The carrier sentences were again constructed in such a way that those parts that were present in all three carrier sentences were acoustically identical. The verb of the sentences differed and determined the bias toward *te* or *de*. The verb *proberen* in the ambiguous sentence was acoustically the same as the verb used in all three carrier sentences in Experiment 2. The

continuum from Experiment 2 was slightly altered by using another production of the schwa to obtain a more natural intonation pattern, and it was placed in the three carrier sentences.

Procedure. The procedure was identical to that of Experiment 2.

Results and Discussion

Figure 5 shows the overall identification functions for the three carrier sentences (upper panel) and the mean percentages of *te* responses in the ambiguous region (Steps 3, 4, 5, and 6; lower panel). The figure displays a pattern different from any of those predicted; there were fewer *te* responses to stimuli in the ambiguous carrier sentence than in the *de*-biased and *te*-biased sentences. The curves do not show any difference between the *de*-biased and *te*-biased sentences. The arcsine values that were used in the ANOVAs were calculated in the same way as in previous experiments. The mean values were 1.76 (*de* biased), 1.76 (*te* biased), and 1.56 (ambiguous). The bias effect was significant, $F(2, 50) = 19.12$, $p < .001$. A Tukey HSD test showed that the ambiguous sentence differed significantly from the two biased sentences.

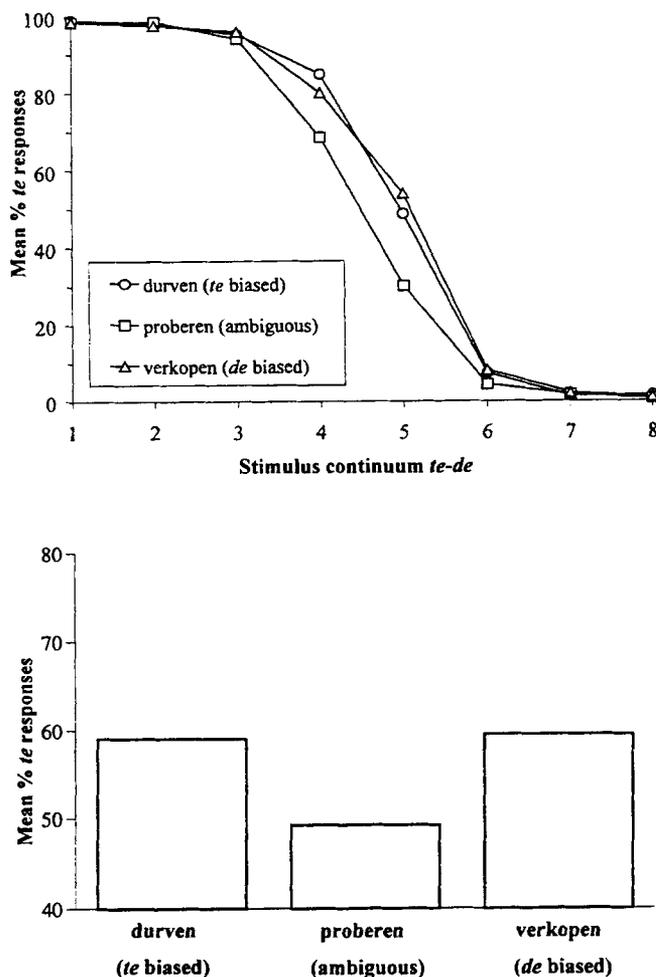


Figure 5. Mean percentage *te* responses in Experiment 3 in each of the three sentential contexts across the stimulus continuum (upper panel) and in the ambiguous region (Steps 3–6; lower panel).

There was no significant difference between the *de*- and *te*-biased sentences.

The data for each participant were partitioned into thirds in the same way as in the previous experiments. The mean RTs for the fast, medium, and slow RT ranges were, respectively, 325 ms ($SD = 119$ ms), 438 ms ($SD = 140$ ms), and 621 ms ($SD = 248$ ms). The mean percentages of *te* responses in the ambiguous region for each context and each RT range are plotted in Figure 6. The data in the three RT ranges show the same pattern as in the overall analysis: There were more *te* responses in the two biased

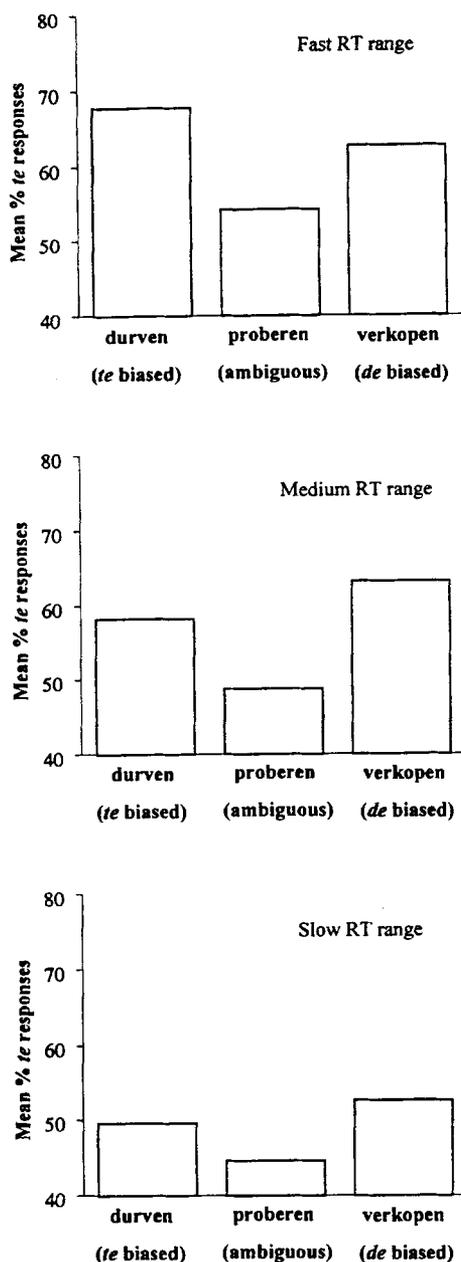


Figure 6. Mean percentage *te* responses in Experiment 3 in each of the three sentential contexts in the ambiguous region, plotted separately for each of the three reaction time (RT) ranges.

sentences than in the ambiguous sentence. Mean arcsine values (calculated in the same way as before) for each sentence and each RT range were 1.84 (*de* biased), 1.95 (*te* biased), and 1.66 (ambiguous) for the fast RT range; 1.84 (*de* biased), 1.74 (*te* biased), and 1.55 (ambiguous) for the medium RT range; and 1.62 (*de* biased), 1.56 (*te* biased), and 1.46 (ambiguous) for the slow RT range. The bias effect was significant in all three RT ranges: $F(2, 50) = 6.03, p = .005$, in the fast RT range; $F(2, 50) = 14.62, p < .001$, in the medium RT range; and $F(2, 50) = 5.15, p < .01$, in the slow RT range. Tukey HSD tests showed that these effects were due to significant differences between the ambiguous sentence and the two biased sentences in all three RT ranges.

There was no effect of sentential context between the *de*- and *te*-biased sentences in any RT range. There was, however, a bias toward *de* in the ambiguous sentence in all three RT ranges. Why might this have occurred? One possibility is that the verb *proberen* was not perfectly ambiguous. It may be that *proberen* followed by a direct object noun phrase (the skates) is preferred over *proberen* with an infinitive verb complement (to skate), perhaps because the direct object reading occurs more often. A post hoc test examined this possibility.

In this test, 15 participants were asked to complete ten written sentences in which the word *te* or *de* was deleted. All ten of the sentences were of the same structure used in Experiment 3 (i.e., they all ended with a word that was ambiguous between a noun plural and an infinitive verb). Three of the sentences were ambiguous, including *We proberen _____ schaatsen*. The other seven sentences contained main verbs that rendered the sentences unambiguous, including *verkopen* and *durven*. Results showed that all participants wrote down *de* after *verkopen* and *te* after *durven*. In the ambiguous sentence *We proberen _____ schaatsen*, two thirds of the participants preferred the word *de*. The test showed that *verkopen* and *durven* were well chosen, in that they had the predicted biases, but that *proberen* did indeed have a *de* bias. This is not sufficient, however, to explain the pattern of results in Experiment 3. If both *proberen* and *verkopen* had *de* biases, both of these conditions should have been different from the *durven* condition in Experiment 3.

An important difference between *proberen* and *verkopen*, however, is that although both have a *de* bias, *proberen* does not mandate that the word before *schaatsen* be *de*, whereas *verkopen* does. Syntactic processing could be completed early given unambiguous items, particularly in an experiment with many repetitions of the same verbs. In principle, after listeners had learned the small set of test sentences, they could select the appropriate verb on the basis of the first phoneme after *we* and then immediately begin syntactic processing. Syntactic analysis of the unambiguous verbs may thus have been completed early and hence may not have influenced the *te-de* decision. But because of the inherent ambiguity of *proberen*, syntactic processing may have been slower to complete so that there was a *de* bias in categorization spread out over the RT ranges. If this argument is correct, then it should be possible to find an effect with unambiguous preceding contexts if these contexts are more varied. In other words, if syntactic processing cannot be completed unusually early, an effect on *te-de* labeling should be found. Experiment 4 tested this hypothesis using the same three kinds of sentences as in Experiment 3 but with more verbs in each condition.

Experiment 4

Method

Participants. Twenty-six native speakers of Dutch from the Max Planck Institute participant pool were paid for taking part. None reported any hearing loss, and none had participated in any of the earlier experiments.

Stimulus construction. Twenty-five sentences were recorded. They were similar to those used in Experiment 3 and are listed in the Appendix. Ten of the sentences contained a *de*-biased verb, 10 sentences contained a *te*-biased verb, and 5 sentences contained an ambiguous verb. Among these were the verbs used in Experiment 3 (*verkopen*, *durven*, and *proberen*). The sentences were recorded and digitized in the same way as before. In each sentence, the first two words were excised and used as the sentence onset. One exemplar of the word *schaatsen* was excised from one of the recorded sentences and served as the last word for all 25 carrier sentences. The *te-de* continuum used in Experiments 2 and 3 was also used here, but with a different production of the schwa. Not all steps of the continuum were used. As a means of preventing the experiment from becoming too long, only the clear endpoints and the three most ambiguous steps (Steps 1, 4, 5, 6, and 8) were used. There were 125 different sentences in total.

Procedure. The procedure was identical to that of Experiments 2 and 3, except for the number of sentences. Listeners heard every step in each of the *de*-biased and *te*-biased sentences three times, and they heard each step in each of the ambiguous sentences six times. Thus, each condition involved 150 responses per participant.

Results and Discussion

The upper panel of Figure 7 shows the overall identification functions for the three conditions. There are clearly more *te* responses to stimuli in the *te*-biased carrier sentences than in the *de*-biased sentences; the percentage of *te* responses to stimuli in the ambiguous sentences lies in between. The proportion of *te* responses for Steps 4, 5, and 6 of the stimulus continuum was calculated for each participant (see the lower panel of Figure 7). As before, an arcsine transformation was applied to these proportions. The mean values for each condition were 1.58 (*de* biased), 1.90 (*te* biased), and 1.71 (ambiguous). An ANOVA indicated that the bias effect was significant, $F(2, 50) = 20.77, p < .001$. A Tukey HSD test showed that all of the pairwise differences among the three conditions were significant at the .05 level. Interestingly, an examination of performance on the individual sentences showed that, in contrast to Experiment 3, there were more *te* responses in the *durven* context than in the *verkopen* context. The *durven* context, however, was not the most strongly *te*-biased context, nor was the *verkopen* context the most strongly *de* biased. It is also important to note that, as in Experiment 3, *proberen* produced a *de* bias relative to both *durven* and *verkopen*. There was some variability among the ambiguous sentences, suggesting that, as with *proberen*, the verbs used in these sentences were not completely unbiased. The average percentages of *te* responses in the ambiguous sentences nevertheless fell between the averages for the two types of unambiguous sentences.

As before, the responses to each step along the stimulus continuum were divided into thirds for each participant. Mean RTs for the fast, medium, and slow RT ranges were, respectively, 351 ms ($SD = 105$ ms), 484 ms ($SD = 139$ ms), and 684 ms ($SD = 243$ ms). As in the overall analyses, the proportions of *te* responses for Steps 4, 5, and 6 for each condition and participant in each RT range were computed; the mean percentages are plotted in Figure

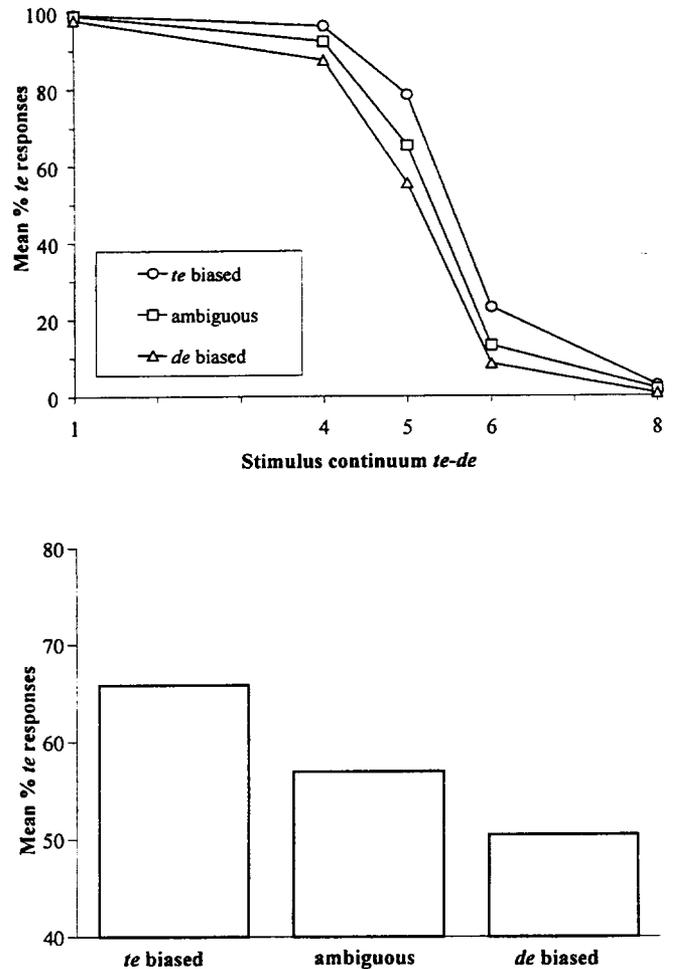


Figure 7. Mean percentage *te* responses in Experiment 4 in each of the three sentential contexts across the stimulus continuum (upper panel) and in the ambiguous region (Steps 4–6; lower panel).

8. These proportions again underwent an arcsine transformation. The mean arcsine values in each condition were 1.51 (*de* biased), 2.19 (*te* biased), and 1.86 (ambiguous) for the fast RT range; 1.62 (*de* biased), 1.92 (*te* biased), and 1.65 (ambiguous) for the medium RT range; and 1.57 (*de* biased), 1.62 (*te* biased), and 1.62 (ambiguous) for the slow RT range. ANOVAs showed that there were significant bias effects in the fast and medium RT ranges, $F(2, 50) = 24.99, p < .001$, and $F(2, 50) = 18.07, p < .001$, respectively. A Tukey HSD test indicated that, in the fast RT range, all pairwise differences among the three conditions were significant. Of the 26 participants, 25 (96%) made more *te* responses to the *te*-biased sentences than to the *de*-biased sentences in the fast range. In the medium RT range, the *te*-biased condition differed significantly from the *de*-biased and ambiguous conditions. Twenty-one participants (81%) made more *te* responses to the *te*-biased sentences than to the *de*-biased sentences in the medium range. There was no significant bias effect in the ANOVA for the slow RT range. Only 12 participants (46%) produced more *te* responses to the *te*-biased sentences than to the *de*-biased sentences in the slow range.

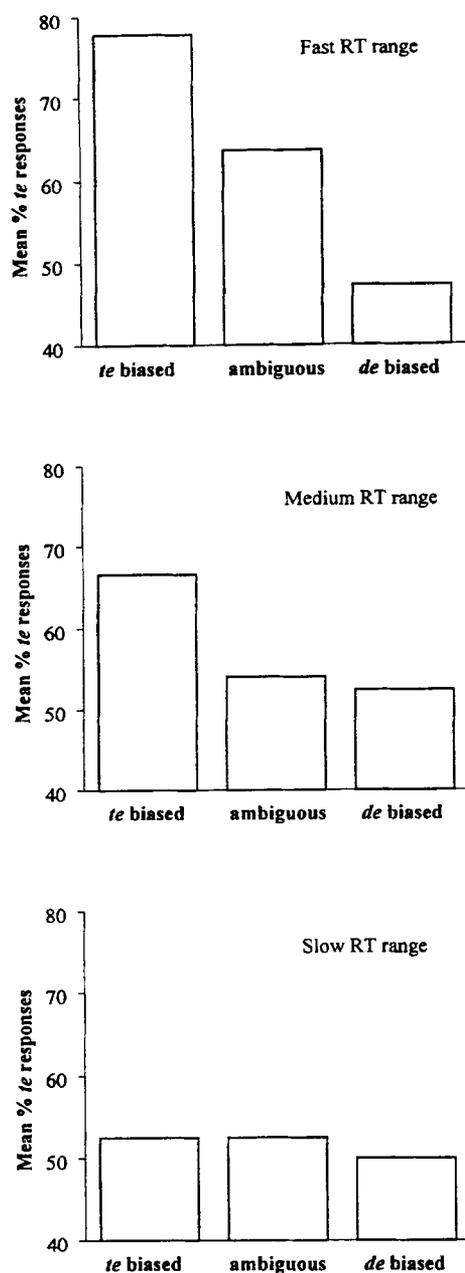


Figure 8. Mean percentage *te* responses in Experiment 4 in each of the three sentential contexts in the ambiguous region, plotted separately for each of the three reaction time (RT) ranges.

There was thus a significant effect of sentence context with the unambiguous sentences that tended to weaken in slower responses. The lack of such an effect in Experiment 3 appears to have been due to the constrained nature of the task situation, which allowed listeners to complete syntactic processing very early. In the more varied situation in Experiment 4, syntactic processing was able to influence the *te-de* decision, but again (as in Experiment 1) only for a limited time.

General Discussion

Sentential context influences the identification of function words, but only within a limited time frame. In Experiment 1, when the disambiguating information in the sentence appeared in the word following the ambiguous function word, listeners used that information in their identification decisions but did so to a greater extent in their fast responses than in their slow responses. There were more *te* (to) decisions when the following word was an infinitive verb (*schieten* [shoot]) than when the following word was a plural noun (*schoenen* [shoes]). This effect was significant in the fast and medium RT ranges but not in the slow RT range. However, when the disambiguating information appeared one word (two syllables) later in the sentences (in Experiment 2), listeners used this information only in their slowest responses.

The last of these findings follows naturally from an analysis of the task demands of Experiment 2 and thus probably reflects a strategic choice on the part of the listeners. They were encouraged to respond as fast as possible, and thus they tended to make their *te-de* decisions without waiting to process the context after the ambiguous word *schaatsen* (skate[s]). The biasing context in fact began 541 ms after the beginning of the schwa of the function word (i.e., 541 ms after the point at which RTs were measured). The bulk of the responses in the fast and medium ranges ($M = 303$ ms and $M = 424$ ms, respectively) were therefore initiated before any of the biasing context had been heard. Nevertheless, on those trials in which listeners made their slowest responses, a context effect was observed. That is, when they had heard information that disambiguated *schaatsen*, they used that information to bias their identification of the function word (the mean RT in the slow range was 663 ms). Sentential context thus appears to bias word identification automatically, even when listeners appear to have adopted a strategy of not waiting for the following context. We suspect, however, that if more material than two syllables had intervened between the function word and the onset of the disambiguation, no context effect would have been found. Connine et al. (1991) found no sentence context effect when disambiguation began six syllables after the target word. Under these circumstances, an insufficient number of responses are likely to be initiated after the disambiguating information has been heard (if there are any such responses at all).

In Experiment 1, listeners used the information in the following context in their faster responses; however, even though they clearly could have also used this information in their slower responses (i.e., it was not the case that the information was not available, as in the fast responses in Experiment 2), they did not do so reliably. This result again suggests that the sentential context effect is automatic rather than strategic. If listeners were using the sentence context in a strategic way in their fast decisions, they should have used it as well in decisions that happened to be slower. It is also unlikely that this result reflects deadline-based responses. Listeners identified the endpoint tokens accurately even in the slow RT range. This suggests that the majority of slow responses were not made at random, in response to an approaching RT deadline. Furthermore, listeners were put under the same time pressure in Experiment 2 and produced the largest contextual effect in their slowest responses.

The results of Experiments 1 and 2 therefore suggest that the effect of sentential context on function word identification is time

limited, because the syntactic processing on which the effect depends itself has a natural time course. If the disambiguation has been fully integrated into the listener's interpretation of an utterance, that information may then not be available to bias identification decisions. On this account, sentential processing decisions can be dissociated from decisions about the identity of ambiguous words. A listener can thus decide that the appropriate word in a sentence frame is *de*, for example, in *Ik probeer* [ʔə] *schoenen*, given the need for a definite article before the plural noun. But, at the same time, he or she can decide on the basis of the information in the speech signal that the ambiguous word is sometimes *te* and sometimes *de*. In other words, just as with the endpoint tokens (a clear *te* will be identified as such even in a *de*-biased context), identification of the phonological form of ambiguous words can be independent of the sentential context. This appears to be the tendency in the case of slow reactions in Experiment 1. We suggest that when the faster responses in Experiment 1 were being initiated, however, syntactic processing was usually not complete, so the contextual bias in the sentence was able to influence the identification decisions.

The results of Experiments 3 and 4, with preceding contexts, support this account. One might have predicted that when the disambiguation preceded the function words, strong effects of context would be observed. There was, however, no effect due to the biased contexts (*verkopen* and *durven*) in Experiment 3. We argued that this was a consequence of the limited set of experimental materials. When listeners could identify the biasing words on the basis of their initial phoneme, and (in contrast to Experiments 1 and 2) those words appeared before the function words, a syntactic decision about which function word was appropriate could be made very early, that is, early enough for that decision to fail to influence the identification decision. The results in the ambiguous condition of Experiment 3 were consistent with this account. In all other experiments, the identification functions in the ambiguous condition usually fell somewhere between the functions for the two biased conditions. In Experiment 3, however, there was a strong bias toward *de* responses in this condition. The post hoc test showed that the ambiguous sentence did in fact have a *de* bias; Dutch speakers preferred the sentence *We proberen de schaatsen* (We try the skates) to *We proberen te schaatsen* (We try to skate). But this bias is less extreme than with a verb such as *verkopen*; *te* is syntactically incorrect in the *verkopen* sentence but is grammatical and only dispreferred in the *proberen* sentence. We suggested that the inherent ambiguity of *proberen* delayed syntactic resolution and, thus, that the *de* bias was able to influence the identification decisions.

In Experiment 4, a greater variety of sentences were used in each condition. We hypothesized that increased variety in the sentences would delay completion of syntactic processing and would thus allow a sentence context effect to be observed in the *te*- and *de*-biased conditions. This hypothesis was confirmed. There were indeed more *te* responses in the *te*-biased sentences than in the *de*-biased sentences, with an intermediate proportion of *te* responses in the ambiguous sentences. Furthermore, as in Experiment 1, the sentential context effect became weaker in the listeners' slower responses. It would again appear that the use of sentential information in word identification is time limited. Even though listeners could have used the sentence bias to assist in their

decisions about ambiguous words, they tended not to in their slower responses.

Taken together, the present results suggest that sentential context acts to bias decision making. This is the same conclusion drawn by a number of other researchers who have examined the influence of sentential context on the identification of content words (Connine, 1987; Connine et al., 1991; Miller et al., 1984). It also concurs with the conclusion drawn by Samuel (1981) on the basis of data from phonemic restoration in content words. One might have expected stronger effects with function words, given that sentential context can impose an absolute constraint on the grammaticality or ungrammaticality of a function word. Although sentential context can impose quite strong constraints on the semantic plausibility of particular content words, such constraints are likely to be less "all or none." A sentential frame could of course impose a syntactic constraint on the suitability of a content word in a particular position, but such constraints have not yet been tested. The fact that the context effects in the present experiments were not particularly strong, along with the fact that the present results lead to the same conclusion as that drawn from experiments with content words, suggests that the current results do not stem from the special status of function words.

We suggest that sentential context biases identification decisions made at a dedicated decision-making stage. This stage of processing is analogous to the level in the Merge model at which phonemic decisions are made (McQueen et al., 1999; Norris et al., 2000). In response to the experimental task at hand, listeners could build representations for *te* and *de* in addition to the lexical representations of these words used in normal word recognition. The decision-level representations could receive input from the lexical level but also from the sentential processing level. Given a perfectly ambiguous input [ʔə], the lexical representations of *te* and *de* would be activated equally and would pass equal amounts of activation on to both decision units. Sentential context could then bias the activation of these decision units. Note that although, on this view, sentence context acts as a decision bias, this bias is considered to be an automatic component of perceptual decision making. It is not considered to be a postperceptual strategic bias. Although it can be difficult to distinguish between different types of bias, several aspects of our data suggest that sentence context effects are automatic rather than strategic. If they were strategic, one would expect to see a bias that was independent of the stimulus continuum (i.e., a bias at the continuum endpoints that was as large as the bias at the category boundary), and, as we have already suggested, one would expect to see a bias that did not vary as a function of RT. Neither of these effects were observed.

Our results suggest that sentence-level processing can bias the *te*-*de* decision, but only if the identification decision is being made at the same time as the syntactic structure of the utterance is being resolved. If syntactic resolution occurs earlier than identification (as in the biased sentences in Experiment 3 and in the slower RT ranges in Experiments 1 and 4), or if resolution occurs later than identification (as in the faster RT ranges in Experiment 2), no bias in identification will be observed. One way of thinking about this is in terms of an activation-based model (for example, Merge; Norris et al., 2000). If the activation level of either the *te* or *de* decision unit is below asymptote when syntactic resolution occurs, contextual information can increase the activation of the unit that is appropriate in that context, thus biasing identification. But if

syntactic resolution occurs much earlier—that is, before the contextually appropriate unit has been activated—there will be no sentence-level influence on the activation of that unit. Likewise, if syntactic resolution occurs relatively late, identification responses will either already have been initiated or the activation level of the contextually appropriate decision unit will be so close to asymptote that the contextual influence will no longer be able to modulate that decision.

Note that, on this view, bottom-up information is given priority in perceptual identification. It has a stronger influence on the activation of the decision units than does contextual information. This assumption is also made in the Merge model with respect to the relative weighting of bottom-up and lexical information in phonemic decision making (Norris et al., 2000). In the present study, we observed no influence of sentential context on the stimuli at continuum endpoints in any experiment or any RT range. This, we argue, is also because of the priority of bottom-up information: The *te* or *de* decision unit is activated so strongly by the unambiguous signal that sentential bias is unable to modify that activation.

The implications of the current findings for models of spoken word recognition are clear. Sentential context does not appear to modulate the phonological encoding of spoken language, contrary to the claims of Borsky et al. (1998). In other words, there is no feedback loop by which the output of sentential processing can influence processes at lexical or prelexical levels. The account of the results in terms of a limited bias in decision making that we have just outlined is consistent with this claim. In such a model, there is no need for sentence-level processes to modify lower levels of processing. Information needs only to flow bottom-up through the recognition system. On this point, we are in agreement with the majority of other authors who have worked on this topic (Connine, 1987; Connine et al., 1991; Miller et al., 1984; Samuel, 1981) but not with the authors of the most recent article on the topic (Borsky et al., 1998). As we argued in the introduction, however, the results of Borsky et al. are in fact consistent with models of spoken word recognition in which there is no feedback.

The present results challenge accounts of sentential context effects based on feedback. If feedback acted to alter lexical or prelexical processes (e.g., by changing the activation of word or phoneme representations), then sentential context effects should not have become weaker in slower responses (as in Experiments 1 and 4). If, for example, the activation level of the lexical representation for *te* were boosted by a sentence context biased toward *te*, the effects of that change in activation should have been seen in all responses made after the change. In particular, if such an activation boost were the cause of the shift in identification seen in fast responses, there should have been a similar shift (if not a stronger one as more time elapsed with more feedback) in slow responses. This is analogous to the effects of lexical context on the activation of prelexical phoneme representations in the TRACE model (McClelland, 1987; McClelland & Elman, 1986), in which the benefit in activation of the lexically consistent phoneme due to feedback remains strong until well after the acoustic offset of the word.

It would, of course, be possible to build a model in which feedback operated only for a limited time window. But note that, in feedback models, the representations responsible for perceptual

decisions are also responsible for encoding speech. Once those representations have been altered by feedback, the coding of the input is distorted. Given [ʔə] in a *te*-biased context, for example, the activation of *te* would be increased, and the activation of *de* would be suppressed. This activation pattern would in turn bias the activation of [t] and [d] nodes through feedback to the prelexical level. The representations at the lower levels of the speech perception system are thus modified by feedback such that they agree with decisions made at higher levels in the system. In this way, feedback tends to overwrite the information that was actually present in the speech signal (Massaro, 1989; Norris et al., 2000). It is therefore difficult for such a system to recover from the effects of feedback. One way to undo these effects (and thus explain how contextual effects tend to disappear in slow responses) would be to store an additional representation of the speech signal in a separate memory buffer and to feed this information back into the speech recognition system once top-down feedback has switched off. This does not strike us as very plausible.

An alternative would be to argue that feedback from sentential levels of processing influences the lexical level but that the lexical level does not then feed back information to the prelexical level. The prelexical level would then be able to encode an undistorted representation of the input that could be used to re-inform the lexical level about the input when feedback was switched off. Feedback, however, is usually seen as something that benefits the word-recognition process (but see Norris et al., 2000, for counterarguments). If feedback were beneficial, it is unclear why it would operate at one stage of the recognition system (from sentential to lexical levels) but not at another (from lexical to prelexical levels) and why it would be time limited.

The reason why the decision-bias account is not challenged by the weakening of the sentential context effect in slower responses is that the account makes a clear distinction between the representations used for perceptual decision making and those that encode the speech signal (i.e., the prelexical and lexical representations used in normal word recognition). A decision unit can be biased by sentential context information at one moment in time, but this bias can be undone as the bottom-up information is processed further. In addition, the bias is seen as something that modifies explicit identification decisions rather than something that modifies the normal process of word recognition. It is thus quite plausible that the bias would be time limited. Such a model can account for the weakening of context effects in slower responses more readily than a model with feedback, because there is no need in the bias model to undo the effects of top-down feedback on lower level representations.

We therefore suggest that there is bottom-up priority in perceptual decision making. Sentential constraints can certainly be used by listeners in their interpretation of an utterance and, within time limits, can also be used to bias word identification decisions. But if the listener must make a decision about what a speaker actually said (i.e., a phonological rather than syntactic-semantic decision), then the speech signal should be king. This is possible in models of spoken word recognition with no feedback. In such models, decisions at the sentential level cannot distort representations of the perceptual world.

References

- Borsky, S., Tuller, B., & Shapiro, L. P. (1998). "How to milk a coat": The effects of semantic and acoustic information on phoneme categorization. *Journal of the Acoustical Society of America*, *103*, 2670–2676.
- Burton, M. W., Baum, S. R., & Blumstein, S. E. (1989). Lexical effects on the phonetic categorization of speech: The role of acoustic structure. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 567–575.
- Burton, M. W., & Blumstein, S. E. (1995). Lexical effects on phonetic categorization: The role of stimulus naturalness and stimulus quality. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1230–1235.
- Connine, C. M. (1987). Constraints on interactive processes in auditory word recognition: The role of sentence context. *Journal of Memory and Language*, *26*, 527–538.
- Connine, C. M. (1990). Effects of sentence context and the lexicon during speech processing. In G. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistics and computational perspectives* (pp. 281–294). Cambridge, MA: MIT Press.
- Connine, C. M., Blasko, D. G., & Hall, M. (1991). Effects of subsequent sentence context in auditory word recognition: Temporal and linguistic constraints. *Journal of Memory and Language*, *30*, 234–250.
- Connine, C. M., & Clifton, C. (1987). Interactive use of lexical information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 291–299.
- Fox, R. A. (1984). Effect of lexical status on phonetic categorization. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 526–540.
- Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 110–125.
- Isenberg, D., Walker, E. C. T., & Ryder, J. M. (1980, November). A top-down effect in the identification of function words. Paper presented at the meeting of the Acoustical Society of America, Los Angeles, CA.
- Massaro, D. W. (1989). Testing between the TRACE model and the fuzzy logical model of speech perception. *Cognitive Psychology*, *21*, 398–421.
- Massaro, D. W., & Oden, G. C. (1995). Independence of lexical context and phonological information in speech perception. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1053–1064.
- McClelland, J. L. (1987). The case for interactionism in language processing. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 3–36). Hillsdale, NJ: Erlbaum.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86.
- McQueen, J. M. (1991). The influence of the lexicon on phonetic categorization: Stimulus quality in word-final ambiguity. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 433–443.
- McQueen, J. M., Norris, D., & Cutler, A. (1999). Lexical influence in phonetic decision-making: Evidence from subcategorical mismatches. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1363–1389.
- Miller, J. L., & Dexter, E. R. (1988). Effects of speaking rate and lexical status on phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 369–378.
- Miller, J. L., Green, K., & Schermer, T. M. (1984). A distinction between the effects of sentential speaking rate and semantic congruity on word identification. *Perception & Psychophysics*, *36*, 329–337.
- Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging phonetic and lexical information in phonetic decision-making. *Behavioral and Brain Sciences*, *23*, 299–325.
- Pitt, M. A. (1995). The locus of the lexical shift in phoneme identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1037–1052.
- Pitt, M. A., & McQueen, J. M. (1998). Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language*, *39*, 347–370.
- Pitt, M. A., & Samuel, A. G. (1993). An empirical and meta-analytic evaluation of the phoneme identification task. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 699–725.
- Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General*, *110*, 474–494.
- Studebaker, G. A. (1985). A "rationalized" arcsine transform. *Journal of Speech and Hearing Research*, *28*, 455–462.

Appendix

Materials for Experiment 4

Sentences With a *de*-Biased Verb

We poetsen *te/de* schaatsen. (We polish to/the skates.)
 We stelen *te/de* schaatsen. (We steal to/the skates.)
 We lenen *te/de* schaatsen. (We borrow to/the skates.)
 We ruilen *te/de* schaatsen. (We exchange to/the skates.)
 We zoeken *te/de* schaatsen. (We look for to/the skates.)
 We verkopen *te/de* schaatsen. (We sell to/the skates.)
 We verstoppen *te/de* schaatsen. (We hide to/the skates.)
 We dragen *te/de* schaatsen. (We wear to/the skates.)
 We begraven *te/de* schaatsen. (We bury to/the skates.)
 We breken *te/de* schaatsen. (We break to/the skates.)

Sentences With a *te*-Biased Verb

We durven *te/de* schaatsen. (We dare to/the skate.)
 We beginnen *te/de* schaatsen. (We start to/the skate.)
 We hopen *te/de* schaatsen. (We hope to/the skate.)
 We besluiten *te/de* schaatsen. (We decide to/the skate.)
 We beweren *te/de* schaatsen. (We claim to/the skate.)
 We behoren *te/de* schaatsen. (We ought to/the skate.)
 We denken *te/de* schaatsen. (We think to/the skate.)^{A1}

We trachten *te/de* schaatsen. (We attempt to/the skate.)
 We schijnen *te/de* schaatsen. (We appear to/the skate.)
 We lijken *te/de* schaatsen. (We seem to/the skate.)
 We plannen *te/de* schaatsen. (We plan to/the skate.)

Ambiguous Sentences

We proberen *te/de* schaatsen. (We try to/the skate[s].)
 We vergeten *te/de* schaatsen. (We forget to/the skate[s].)
 We prefereren *te/de* schaatsen. (We prefer to/the skate[s].)
 We weigeren *te/de* schaatsen. (We refuse to/the skate[s].)
 We beloven *te/de* schaatsen. (We promise to/the skate[s].)

^{A1} Literal translations are given throughout. In this case, the sentence means "We think we are skating."

Received September 28, 1999

Revision received October 6, 2000

Accepted December 11, 2000 ■

Low Publication Prices for APA Members and Affiliates

Keeping you up-to-date. All APA Fellows, Members, Associates, and Student Affiliates receive—as part of their annual dues—subscriptions to the *American Psychologist* and *APA Monitor*. High School Teacher and International Affiliates receive subscriptions to the *APA Monitor*, and they may subscribe to the *American Psychologist* at a significantly reduced rate. In addition, all Members and Student Affiliates are eligible for savings of up to 60% (plus a journal credit) on all other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the American Association for Counseling and Development, Academic Press, and Human Sciences Press).

Essential resources. APA members and affiliates receive special rates for purchases of APA books, including the *Publication Manual of the American Psychological Association*, and on dozens of new topical books each year.

Other benefits of membership. Membership in APA also provides eligibility for competitive insurance plans, continuing education programs, reduced APA convention fees, and specialty divisions.

More information. Write to American Psychological Association, Membership Services, 750 First Street, NE, Washington, DC 20002-4242.