

J38 A

Can linguistic boundaries change the effectiveness of silence as a phonetic cue?

Bruno H. Repp

Haskins Laboratories, 270 Crown Street, New Haven, CT 06511-6695, U.S.A.

Received 10th May 1985, and in revised form 18th November 1985

This study investigated the influence of three kinds of linguistic boundaries—word boundaries, prosodic breaks, and syntactic breaks—on the perception of a silent interval at the boundary site as a cue to the presence of a labial stop consonant. The experimental technique involved cross-splicing of portions of four naturally produced pairs of sentences, as well as presentation of excerpts from these sentences. Although one sentence pair showed a pronounced syntactic boundary effect, the other three (including two that were better controlled for semantic bias) did not, which points to a different, stimulus-specific origin of the effect obtained. Prosodic boundary effects were also generally absent, presumably because the stimuli were constructed such that prosodic variation ceased 78 ms prior to the critical silent interval. Only introduction of a word boundary effected a systematic reduction in stop consonant percepts, although this manipulation was confounded with other contextual factors. On the whole, the data provide little evidence for any direct effects of structural linguistic variables on phonetic segment perception; such effects seem to be restricted to the level of word recognition.

1. Introduction

One fundamental issue in speech perception research concerns the relative importance of physical signal properties (“bottom-up” information) versus the listener’s expectations and interpretations (“top-down” processes). There is little doubt that phonotactic, semantic, and pragmatic factors can influence word perception, particularly when the speech signal is ambiguous (see, e.g., Fox, 1984; Ganong, 1980; Massaro & Cohen, 1983). Whenever a listener has internally generated or contextually induced expectations about the likelihood of certain phonological or lexical alternatives, these expectancies will help reduce any uncertainty introduced by insufficient physical information.

It is much less clear whether a listener’s apprehension of structural factors that do *not* affect the likelihood of phonological or lexical alternatives can have repercussions at the level of phonetic segment perception. Specifically, the question is whether linguistic boundaries (syllabic, lexical, or syntactic) can *reduce the phonetic coherence* of an utterance at the boundary site, with possible consequences for the perceived segmental composition. Such an interaction, if it were to occur, would be theoretically interesting,

for it would suggest that higher-level processes of lexical access and syntactic analysis can exert a direct influence on the internal representation of the bottom-up information, or at least generate expectations about its detailed acoustic structure. It should be kept in mind, however, that the effects under investigation, unlike the top-down effects studied extensively in research on word recognition (see, e.g., Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Welsh, 1978), are rather special phenomena that, even if real, probably play only a very minor role in real speech understanding.

The evidence for such effects, however, is not compelling so far. Previous studies on this topic have been concerned with the function of silence as a phonetic cue. There is much evidence that short periods of silence in speech are not perceived as gaps or interruptions but as carriers of articulatory information about closure of the vocal tract, as occurs in connection with stop and affricate consonants (see, e.g., Dorman, Raphael & Liberman, 1979; Repp, Liberman, Eccardt & Pesetsky, 1978). One particular situation investigated in several recent studies involves the effect of a short interval of silence preceding a fricative noise as a cue to the contrast between a word-initial fricative and affricate (Dechovitz, 1979, 1980, 1981; Price & Levitt, 1983; Rakerd, Dechovitz & Verbrugge, 1982). The hypothesis tested in these studies was that introduction of a coincident linguistic break might reduce the perceptual effectiveness of the silence, either because the silence could be interpreted as a hesitation associated with the break rather than as an articulatory closure associated with a stop consonant, or because the linguistic boundary has a direct disruptive influence on the coherence of the signal portions preceding and following the silence, so that the presence and precise duration of the closure interval become perceptually irrelevant. Dechovitz (1979, 1980, 1981) claimed to have found such an effect due to syntactic structure alone; i.e., he found a significant reduction of silence-cued affricate percepts when a syntactic boundary at the critical location was created by remote context under semantically neutral and constant local acoustic conditions. However, these data have not been published and Price & Levitt (1983) have failed to replicate the effect.

Nevertheless, all these previous studies found that the introduction of clause- or sentence-final prosody—including a falling intonation contour and final syllable lengthening—reduced the perceptual effect of a following silent interval. Although prosodic changes usually accompany changes in linguistic structure and thus carry considerable lexical and syntactic information, they do involve acoustic changes in the immediate vicinity of the silent interval. Since this may alter some of the local phonetic cues, the observed prosodic effects may not represent an influence of perceived linguistic structure on phonetic perception but may have more direct causes.

The present experiment extended these earlier studies by further investigating the influence on phonetic perception of syntactic and prosodic breaks, and by also considering the possible role of word boundaries. Stimulus materials were chosen in which the critical silence served as a cue for a labial stop consonant following a fricative and preceding a liquid (see Dorman *et al.*, 1979; Fitch, Halwes, Erickson & Liberman, 1980). The fricative-affricate contrast used previously is characterized by a rather sharp category boundary at a very short silence duration, which raises the possibility of psychoacoustic interactions that are immune to contextual influences. On the other hand, the type of contrast employed here typically has its category boundary at relatively longer silence durations, so low-level psychoacoustic interactions are unlikely (see Pastore, Szczesiul & Rosenblum, 1984; Repp, 1985), and it also has a larger region of ambiguity, which makes it more sensitive to influences of all

kinds.¹ The critical silence was embedded in plausible, natural sentences, which constitutes an improvement over the somewhat contrived and limited materials used in earlier investigations. Syntactic and prosodic factors were varied independently by swapping the two words surrounding the critical silent interval between syntactically different sentence frames. (Prosodic variation beyond these two words was confounded with syntactic structure.) Prosodic variation in the word immediately preceding the silence included the duration and amplitude envelope of the final [s] noise segment which, judging from earlier findings and from informal observations during stimulus construction, would certainly have had a strong perceptual effect. Because of this foregone conclusion, it was decided to neutralize this segment and to examine only whether prosodic information *beyond* the immediately preceding acoustic segment can influence phonetic perception.

It should be pointed out that the role of silence as a cue to stop consonant perception is twofold. If the closure silence is too short (less than about 60 ms in the fricative-liquid context), no stop consonant may be perceived even when other cues are available (e.g., Dorman *et al.*, 1979; Fitch *et al.*, 1980). If the silence is longer (roughly 100–300 ms), a (labial) stop consonant will often be perceived even when there are no other cues (Dorman *et al.*, 1979; Repp, 1985). These two effects may be called “stop suppression” and “stop generation”, respectively (Repp, 1985). The stop suppression effect may in part be due to psychoacoustic interactions (such as forward masking) between the closely adjacent signal portions (however, see Pastore *et al.*, 1984), whereas such interactions are much less likely in the case of the stop generation effect. Therefore, if there are any effects of linguistic boundaries on phonetic perception, they are more likely to occur at longer closure intervals, where psychoacoustic interactions play no role. The specific hypothesis tested was that, compared to a no-boundary condition, introduction of a linguistic boundary at the point of the critical silence would *decrease* the number of stop consonant responses at relatively long closure durations. To the extent that stop suppression is *not* caused by psychoacoustic interactions, an *increase* of stop responses might be predicted at short closure durations, because a linguistic boundary might then reduce the (negative) cue value of short silences as well.

Following some piloting, two full-size experiments were conducted which were very similar in design. Because stimulus parameters were still not optimal, the first experiment inadvertently focused exclusively on the region of stop consonant suppression, where little sensitivity to linguistic boundaries was expected (and obtained). Therefore, only the results of the second experiment will be reported which, due to additional stimulus adjustments, successfully encompassed both regions of stop consonant generation and suppression. Where the two designs overlapped, the results of the first experiment were consistent with the findings reported below.

2. Methods

2.1. Subjects

Ten paid volunteers participated. All were Yale undergraduates and native speakers of American English.

¹ However, the effects studied here do not *require* phonetic ambiguity, as do most other contextual effects in speech perception. Rather, these purely structural effects, if extant, should disrupt the perceptual contribution of closure silence even at its *optimal*, least ambiguous setting. (That there is often some ambiguity even at that setting is due to the fact that closure duration is only a secondary cue to stop manner; see Repp, 1984a).

TABLE I. Stimulus sentences

-
- | | |
|----|---|
| 1. | a. The royal tomb was protected by <i>six</i> (<i>b</i>) <i>locks</i> of solid gold. |
| | b. When the clock strikes <i>six</i> , (<i>b</i>) <i>lock</i> the gate. |
| 2. | a. The girl tried to <i>kiss</i> (<i>B</i>) <i>Radford</i> on the cheek. |
| | b. After giving his wife a <i>kiss</i> , (<i>B</i>) <i>Radford</i> boarded the train. |
| 3. | a. Will you please welcome <i>Miss</i> (<i>B</i>) <i>Lackman</i> to the office. |
| | b. Enraged by a spectacular <i>miss</i> , (<i>B</i>) <i>Lackman</i> quit the game. |
| 4. | a. To the maid's dismay, <i>worse</i> (<i>b</i>) <i>rooms</i> could hardly be imagined. |
| | b. What made matters <i>worse</i> , (<i>b</i>) <i>rooms</i> were difficult to find. |
-

2.2. Stimulus preparation

The stimulus sentences are shown in Table I. Four pairs of sentences were constructed. The members of each pair contained the same two critical words in succession; the first word ended in [s], whereas the second word either did or did not begin with [b], so that there were two versions of each sentence. In one sentence of each pair (version b), a clause boundary intervened between the two critical words, whereas in the other sentence (version a) the two words formed a syntactic unit. The second critical word, which either did or did not begin with [b], represented fictitious surnames in two instances (Nos. 2 and 3) and real words in the other two (1 and 4). Orthogonal to this distinction, the consonant following the optional [b] was [l] in two words (1 and 3) and [r] in the other two (2 and 4). Because of the two possible versions of the second critical word, there was a total of 16 sentences.

These 16 sentences were recorded by a male speaker of American English in a sound-insulated booth using high-quality equipment. The recordings were low-pass filtered at 4.9 kHz and digitized at a 10 kHz sampling rate. Using a waveform editor in conjunction with careful listening, each sentence was divided into four sections which were stored in separate computer files: preceding context (C1), first critical word (W1), second critical word (W2), and following context (C2). All cuts were made at zero crossings. In those sentences in which W2 had an initial [b], the stop closure was edited out and discarded. Thus, W1 ended at the beginning of the stop closure and W2 began at its end. In sentences without a W2-initial [b], the end of W1 and the beginning of W2 coincided, except in two sentences in which a lateral noise burst occurring at an [s-l] juncture was edited out.

For each sentence pair listed in Table I, each of the two different context frames (C1 + C2) existed in two distinct productions. Only one of these was retained—that deriving from sentences in which W2 had been articulated with an initial [b] (an arbitrary choice). The first critical word (W1) existed in four recorded versions; only those two versions that were *not* followed by a W2-initial [b] were used (another arbitrary choice). In these two remaining versions, the clause-final [s] noises were much longer in duration (ranging from 144 to 201 ms across the four sentences) than the non-clause-final [s] noises (range 55–109 ms). For reasons outlined in Section 1, these noises were removed and replaced by a constant [s] noise excerpted from the same talker's production of the word "spectacular" (Sentence 3b). (For an explanation of this choice, see below and footnote 2.) This [s] noise, originally only 54 ms in duration, was artificially lengthened to 78 ms by duplicating a 24-ms central section of the waveform.

Finally, the onsets of the W2 words, which existed in four recorded versions, were examined and edited. Words articulated with an initial [b] all had labial release bursts ranging in duration from 12 to 18 ms. These bursts, which provided strong stop manner cues (see, e.g., Repp, 1984a) were eliminated, leaving only potential coarticulatory cues in the periodic stimulus portion. The words without an initial [b] had no bursts and were retained without change.

In summary, then, for each of the four sentence pairs listed in Table I, there were two different context frames C1 + C2, each in a single recorded version; two versions of W1, a clause-final one and a non-clause-final one, with a common final [s] noise; and four versions of W2, two that had originally started with [b] and two that had not, and orthogonal to this distinction, two clause-initial and two non-clause-initial ones.

These components were re-assembled into sentences, with four different silent closure intervals introduced between the W1 and W2 words: 40, 80, 120, and 160 ms. All possible combinations of sentence components were employed in the *sentence* test, leading to a total of 4 (sentence types) \times 2 (contexts) \times 2 (W1) \times 4 (W2) \times 4 (silences) = 256 sentences. They were recorded in four blocks of 64, randomized within each block in groups of 16, with interstimulus intervals (ISIs) of 3 s, and intervals of 10 s between groups. The first and third blocks contained sentences in which the prosody of W1 was appropriate for the syntactic context, whereas the second and fourth blocks contained the sentences in which W1 had the inappropriate prosody. These latter sentences sounded somewhat odd but not bizarre; they were deemed appropriate for an assessment of prosodic factors.

In addition to this lengthy main test, four shorter test tapes were recorded. The first of these was a *pretest* containing 16 sentences. The first eight sentences represented the eight different contexts, with prosodically appropriate W1 and W2; W2 was either the "stronger" version (i.e., that originally began with [b]) preceded by the second-shortest silence (80 ms), or the "weaker" version (that originally began with [l] or [r]) preceded by the longest silence (160 ms). The second set of eight sentences contained the context-W2 combinations not contained in the first set. All 16 sentences were arranged in a quasi-counterbalanced sequence, with ISIs of 20 s. The purpose of this pretest was to assess the listeners' response to the test sentences on first hearing.

The second test contained the W1-silence-W2 *word pairs* in all 128 possible combinations, without their sentential context. They were recorded in four blocks of 32, with ISIs of 4 s. The purpose of this test was to provide a baseline for assessing the contribution of the contextual frame, regardless of its syntactic implications, and to examine prosodic effects in this more restricted context (cf. Price & Levitt, 1983).

In the third test, the W2 words were preceded only by the constant [s] noise plus silence, to provide a baseline for testing the hypothesis that a word boundary following the [s] reduces the likelihood of silence-cued labial stop percepts. In this test, the [s] was to be perceived as the initial segment of a *nonsense word* (e.g., "splock"). The constant [s] noise was taken from a word-initial position (see above) to facilitate this task.²

Finally, the excerpted W2 words were assembled into a *single-word* test. The 16 W2 words (4 words \times 4 versions) were recorded in four different random orders with ISIs

²In the author's judgment, word-final [s] noises were not acceptable as word-initial segments, whereas the word-initial [s] seemed acceptable both as a word-initial or word-final segment. In any case, in the sentences and word pairs lexical and semantic constraints were assumed to exert sufficient pressure on listeners to consider the [s] as W1-final, even if its acoustic characteristics were more appropriate for a word-final position.

of 4 s. This test was to provide a baseline against which the effect of closure silence in the other tests could be compared.

2.3. Procedure

The subjects listened to all tests in a single session, using TDH-39 earphones in a quiet room. The tests were presented in a fixed sequence: the pretest was followed by the sentence test, the word pair test, the nonsense word test, and the single word test.

In the pretest, the subjects' task was to write each sentence down verbatim on a blank sheet of paper. Subjects were informed that the sentences were meaningful, that some of them contained proper names, and that the second set of eight would be very similar, but not necessarily identical, to the first set of eight.

For the sentence test, the subjects were provided with printed answer sheets. Each page listed all the stimulus sentences on top, arranged as in Table I, without the italics but with two words in each sentence capitalized. The first of those words was a keyword in the first clause (e.g., ROYAL) identifying the context; the second was W2. For each item the answer sheets listed the four pairs of possible key words and W2 below each pair, with the initial B in parentheses. The subjects' task was, for each sentence heard, to first circle the appropriate keyword and then to indicate, by either circling or crossing out the parenthetical B in the word below, whether W2 did or did not begin with a [b]. Since the sentences came at a fairly brisk rate, the subjects were encouraged to circle the keyword before the sentence was over, and to skip the keyword if the time seemed too short. Some subjects omitted a few keyword responses in the beginning but soon found their rhythm. The only purpose of the keyword responses was to keep the subjects' attention on the context and thus to prevent an overly selective listening strategy.

For the word pair test, answer sheets listed for each item the four possible W1-W2 pairs, with the W2-initial B in parentheses. The subjects' task was to find the appropriate word pair and to either circle or cross out the B. For the nonsense word test, the answer listed for each item the four possible choices with a parenthetical P following the initial S (i.e., S(P)LOCK, S(P)RADFORD, S(P)LACKMAN, S(P)ROOMS). Subjects were asked to try their best to consider the stimuli as [s]-initiated nonsense words and to either circle or cross out the P in the correct alternative. Their attention was drawn to the unfamiliar [sr] cluster as a possible beginning of a nonsense word. Finally, the answer sheet for the single word test listed the four possible W2 choices for each item, and subjects located the correct alternative and either circled or crossed out the parenthetical word-initial B.

3. Results and discussion

3.1. General contextual effects

Averaging over different versions of W1 and W2, Figure 1 shows the results in terms of percent labial stop responses to W2 onset, separately for each sentence pair (S1-S4), as a function of silent closure duration. The various response functions compare sentences with (S-b) and without (S-a) a syntactic break preceding W2, word pairs (WP), and nonsense words (NW). The percentage of "b" responses to single W2 words (SW) is indicated by the arrows at the right-hand side of each panel.

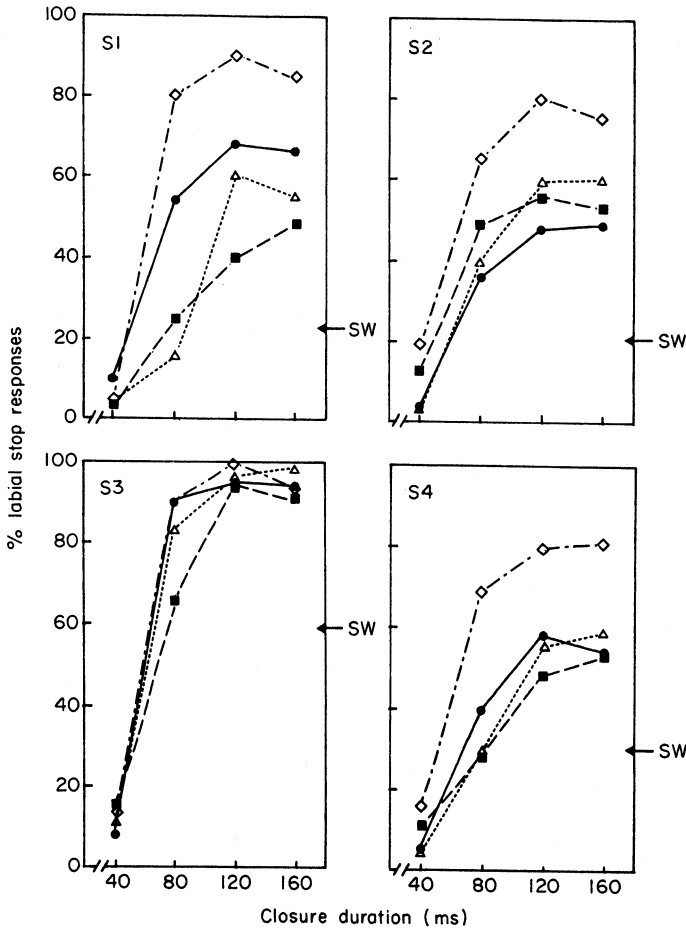


Figure 1. Percent stop responses, separately for the four sentences (S1–S4), as a function of closure duration. Separate response functions are shown for sentences without a syntactic boundary (S-a, ●), with a syntactic boundary, (S-b, ■), for isolated word pairs (WP, △), and for nonsense words (NW, ◇). The arrows represent the percentages for single words (SW).

The first thing to note is that the percentage of labial stop percepts increased as closure duration increased. Repeated-measures analyses of variance on the separate tests showed that this expected effect was extremely significant and also interacted strongly with the sentence factor, as is evident from the different slopes of the response functions (all effects at least $p < 0.001$). A visual comparison with the single-word (SW) percentages shows that labial stop responses at the longer closure durations exceeded those to single W2 words by a considerable margin (the stop generation effect), whereas the opposite relationship held at the shortest silent interval (the stop suppression effect).

The next finding to note in Fig. 1 is that the response functions for word pairs (WP) were not systematically different from those for sentences (S-a and S-b combined); thus, having some sentential context around the W1–silence–W2 constellation did not influence the subjects' criterion for reporting a "b". By contrast, the percentages of labial

TABLE II. W1 durations (not including the final [s] noise) and terminal fundamental frequencies (F_0)

Sentence	W1	Duration (ms)			Terminal F_0 (Hz)	
1. a.	<i>six</i>	[s]	[l]	[k]	Total	98
		75	46	51	172	
b.					Total	62
		135	62	92	289	
2. a.	<i>kiss</i>	[k ^h]	[l]		Total	86
		46	51		97	
b.					Total	53
		63	87		150	
3. a.	<i>miss</i>	[m]	[l]		Total	82
		39	50		89	
b.					Total	50
		106	91		197	
4. a.	<i>worse</i>	(not segmentable)			Total	87
					151	
b.					Total	54
					233	

Definitions: [s] = fricative noise; [l] = voiced portion; [k] = silent closure interval; [k^h] = release burst and aspiration; [m] = nasal murmur; terminal F_0 = average F_0 of the last three complete pitch periods.

4. Summary and conclusions

In the present study it was attempted to create a perceptual discontinuity at the point of a critical silent interval by purely linguistic means in a relatively natural speech processing situation. The effect of word boundaries was studied, as well as the effects of (slightly removed) prosodic and syntactic breaks, following earlier studies by Dechovitz (1979, 1980, 1981), Rakerd *et al.* (1982), and Price & Levitt (1983).

There was a clear effect of introducing a word boundary. Although this effect was confounded with the presence vs. absence of preceding word context and therefore must be interpreted with care, it does suggest the possibility that within-word silence is more tightly integrated into the speech stream than is between-word silence. The reason for this may lie in subjects' expectations based on experience with real speech, in which interword intervals tend to be less reliable indicators of phonetic distinctions than intraword silences.

In contrast to several previous studies, there were no effects of prosodic discontinuity. The most likely explanation for this is the fact that the fricative noise immediately preceding the critical silence was not allowed to vary, so that the distinctive prosodic information ended 78 ms before the silent interval. If this interpretation is correct, it indicates that prosodic effects of the kind demonstrated by Price & Levitt (1983) and Rakerd *et al.* (1982) are extremely local in character and are probably caused by the duration of the acoustic segment preceding the silence, which acts as a secondary stop manner cue. Similarly restricted effects have been observed in related experiments on the perception of vowel duration in sentence context (Luce & Charles-Luce, 1985; Nooteboom & Doodeman, 1980) and on the perceptual consequences of varying speaking rate (e.g., Summerfield, 1981). Rather than constituting a direct influence of suprasegmental variation on segmental perception, these effects may be mediated by changes in local acoustic signal properties serving as segmental cues.

There were no consistent effects of syntactic structure *per se* on phonetic perception. The anomalous results for one sentence pair were probably due to a semantic bias. These

negative results confirm the conclusions of Price & Levitt (1983) and cast further doubt on the replicability of Dechovitz's (1979, 1980, 1981) unpublished findings showing a "purely syntactic" effect on phonetic perception. It seems likely that syntactic processes operate exclusively at a level beyond that of segmental phonetic classification.

This research was supported by NICHD Grant HD01994 and BRS Grant RR-05596 to Haskins Laboratories. I am grateful to Andrea Levitt and Patti Price for critical comments.

References

- Dechovitz, D. (1979). Effects of syntax on the perceptual integration of segmental features. In *Speech Communication Papers Presented at the 97th Meeting of the Acoustical Society of America* (J. J. Wolf & D. H. Klatt, editors), pp. 319–322. New York: Acoustical Society of America.
- Dechovitz, D. R. (1980). Constraints imposed on phonetic processing by syntax, *Journal of the Acoustical Society of America*, **67**, Suppl. 1, S99.
- Dechovitz, D. R. (1981). Phonetic integration as a potential test for controversial surface structures, *Journal of the Acoustical Society of America*, **69**, Suppl. 1, S115.
- Dorman, M. F., Raphael, L. J. & Liberman, A. M. (1979). Some experiments on the sound of silence in phonetic perception, *Journal of the Acoustical Society of America*, **65**, 1518–1532.
- Fitch, H. L., Halwes, T., Erickson, D. M. & Liberman, A. M. (1980). Perceptual equivalence of two acoustic cues for stop consonant manner, *Perception & Psychophysics*, **27**, 343–350.
- 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., III (1980). Phonetic categorization in auditory word perception, *Journal of Experimental Psychology: Human Perception and Performance*, **6**, 110–125.
- Luce, P. A. & Charles-Luce, J. (1985). Perception of word-final voicing in two sentence contexts. Paper presented at the 109th meeting of the Acoustical Society of America, Austin, Texas, April 1985.
- Marslen-Wilson, W. D. & Welsh, A. (1978). Processing interactions and lexical access during word-recognition in continuous speech, *Cognitive Psychology*, **10**, 29–63.
- Marslen-Wilson, W. D. & Tyler, L. K. (1980). The temporal structure of spoken language understanding, *Cognition*, **8**, 1–71.
- Massaro, D. W. & Cohen, M. M. (1983). Phonological context in speech perception, *Perception & Psychophysics*, **34**, 338–348.
- Nooteboom, S. G. & Doodeman, G. J. N. (1980). Production and perception of vowel length in spoken sentences, *Journal of the Acoustical Society of America*, **67**, 276–287.
- Pastore, R. E., Szczesiul, R. & Rosenblum, L. (1984). Does silence simply separate speech components? *Journal of the Acoustical Society of America*, **72**, 1904–1907.
- Price, P. J. & Levitt, A. G. (1983). The relative roles of syntax and prosody in the perception of the /s/–/ɛ/ distinction, *Language and Speech*, **26**, 291–304.
- Rakerd, B., Dechovitz, D. R. & Verbrugge, R. R. (1982). An effect of sentence finality on the phonetic significance of silence, *Language and Speech*, **25**, 267–282.
- Repp, B. H. (1984a). The role of release bursts in the perception of [s]-stop clusters, *Journal of the Acoustical Society of America*, **75**, 1219–1230.
- Repp, B. H. (1984b). Effects of temporal stimulus properties on perception of the [sl]–[spl] distinction, *Phonetica*, **41**, 117–124.
- Repp, B. H. (1985). Perceptual coherence of speech: Stability of silence-cued stop consonants, *Journal of Experimental Psychology: Human Perception and Performance*, **11**, 799–813.
- Repp, B. H., Liberman, A. M., Eccardt, T. & Pesetsky, D. (1978). Perceptual integration of acoustic cues for stop, fricative, and affricate manner, *Journal of Experimental Psychology: Human Perception and Performance*, **4**, 621–637.
- Summerfield, Q. (1981). On articulatory rate and perceptual constancy in phonetic perception, *Journal of Experimental Psychology: Human Perception and Performance*, **7**, 1074–1095.
- Summerfield, Q., Bailey, P. J., Seton, J. & Dorman, M. F. (1981). Fricative envelope parameters and silent intervals in distinguishing 'slit' and 'split', *Phonetica*, **38**, 181–192.