Article

Language and Speech

Language *and* Speech
54(2) 181–197

DOI: 10.1177/0023830910397494
las.sagepub.com
SAGE

Processing of No-release Variants in Connected Speech

Paul C. LoCasto
Quinnipiac University

Cynthia M. Connine
Binghamton University, State University of New York

Abstract

The cross modal repetition priming paradigm was used to investigate how potential lexically ambiguous no-release variants are processed. In particular we focus on segmental regularities that affect the variant's frequency of occurrence (voicing of the critical segment) and phonological context in which the variant occurs (status of the following word-initial segment). Primes consisted of carrier words ending in a segment likely (voiced; e.g., BAND) or unlikely (voiceless; PLANT) to be produced in no-release form followed by a consonant or vowel onset context word. Each carrier word had an embedded lexical competitor (embedded word) formed by the segments prior to its final consonant (e.g., plan in plant). Productions of these embedded words (true embedded word) were also used as primes. Both true embedded words (Experiments 1a–1c) and carrier words (Experiments 2a–2c) were used as visual targets. The results are discussed in terms of the contribution of probabilistic speech events to theories of spoken word recognition and lexical competition.

Keywords

auditory processing, phonological variation, speech perception

1 Processing of no-release variants in connected speech

The production of connected speech often results in missing, reduced or altered phonetic detail due to the influence of surrounding phonemes. The resultant variability presents a challenge for theories to adequately describe how words that vary in their physical form are represented and processed with minimal disruption to the spoken word recognition system. Further, recognition of words that are missing acoustic events may be more problematic to resolve if the consequence is an ambiguous sequence that can be a version of two distinct words. The present investigation

Corresponding author:
Paul C. LoCasto, Department of Psychology, Quinnipiac University, 275 Mt. Carmel Ave., CLAC1, Hamden, CT 06518, USA; tel: +203-582-3725
Email: paul.locasto@quinnipiac.edu

focuses on variation that systematically occurs in specific contexts (licensing environments). The systematic nature of this variability in lexical form makes their realization predictable. This predictability represents an ideal way to investigate listeners' use of knowledge concerning licensing environments as well as other distributional factors when processing phonological variation.

Specifically, the current set of experiments focuses on words that end in stop consonants that can occur with or without a final release burst (no-release variants). We investigate two factors that may affect the processing of word final no-release variants: the distributional frequency of release occurrence across segment categories (voiced vs. voiceless) and the distributional frequency of release occurrence as a function of following phonological context (vowel vs. consonant). As discussed below, we investigate these factors in the presence of embedded words in order to investigate the general issue of lexical ambiguity and parallel activation.

Recent studies have investigated rule-governed variation that occurs in word final stops where reduction or deletion results in missing information. Sumner and Samuel (2005) have shown that while systematic variation of word-final /t/ is processed fluidly at a semantic level, it presents form processing difficulties in the long term. Processing of similar word-final stop variants have been shown to be sensitive to distributional probabilities of occurrence. In Dutch for instance, /t/ reduction is more likely to occur after /s/ than /n/ (Mitterer & Ernestus, 2006). Listeners have been shown to perceive the presence of a /t/ in /t/-reduced stimuli more often when presented after /s/ than /n/ (Mitterer & Ernestus, 2006; Mitterer, Yoneyama, & Ernestus, 2008).

In American English, one factor that influences the distributional frequency of a release is the voicing category of the segment. A corpus analysis of connected speech (Crystal & House, 1988) found that word-final voiceless stops were released 42% of the time, whereas voiced stops were released only 18% of the time. Thus, the presence or absence of the final release patterned probabilistically with the voicing status of the segment, with a higher likelihood of release occurring for voiceless segments. There is some evidence that listeners use this distributional asymmetry during spoken word recognition. Deelman and Connine (2001) found that phoneme monitoring reaction times for voiced segments were equivalent for the release and no-release variants but responses to voiceless segments showed an advantage for the more frequent release-bearing form. These results suggest that listeners are sensitive to the probabilistic patterning of releases for words presented in isolation. However, there are a number of acoustic-phonetic properties that co-occur with voiced and voiceless segments. Vowels preceding voiced segments are typically longer than vowels preceding voiceless segments and listeners use this information in speech perception (Raphael, 1972; Warren & Marslen-Wilson, 1988). There is also evidence that the co-articulatory cues in a vowel preceding a voiceless consonant are reduced compared to a voiced consonant (Marslen-Wilson & Warren, 1994; but see Whalen, 1984). Therefore, the processing of no-release variants may draw upon acoustic-phonetic information as well as probabilistic knowledge about word-final releases for successful recognition.

A second factor influencing the distribution of word final releases is the nature of the following word. Crystal and House (1988) found that releases tend to be produced more often when followed by a vowel (a range of 18–96%, depending on syllable stress patterns) than when followed by a consonant (4–21%, depending on the following consonant's manner of articulation). These differences in patterning based on subsequent acoustic-phonetic context suggest a second source of probabilistic information for processing no-release variants that is only available in connected speech. Indeed, the likelihood of listeners perceiving the presence of a reduced /t/ has been shown to pattern with distributional probabilities of following segments. Listeners reported the presence of /t/ more often if subsequent word began with /m/ compared to /k/ (Mitterer & Ernestus, 2006).

The nature of connected speech also raises the possibility that co-articulatory cues for the final segment may be available in the following word onset. Co-articulatory effects on spoken word processing are well documented (e.g., Streeter & Nigro, 1979; Martin & Bunnell, 1981; Warren & Marslen-Wilson, 1987; Whalen, 1991). It is therefore important to differentiate between context probability effects and word boundary co-articulation effects.

Of further interest is the more general issue of embedded words (McClelland & Elman, 1986; Norris, McQueen, & Cutler, 2000.) The primary way theories of spoken word recognition attempt to accommodate embedding is to assume that lexical activation occurs in parallel (Zwitserlood, 1989; Connine, Blasko, & Wang, 1994; but see Nooteboom, Janse, Quené, & te Riele, 2000). TRACE simulations (Frauenfelder & Peeters, 1990) have demonstrated that a carrier word reaches a higher level of activation than a target embedded word. Such differences have been explained in terms of the activation/inhibition relationship between lexical competitors. Activation levels of an embedded word decrease when a carrier word is presented due to the mismatching information present in the input; decreased activation of the embedded word renders it a less effective source of inhibition for the carrier word. In the present context, the recognition of words that are missing acoustic events may be more problematic to resolve if the consequence is an ambiguous sequence that can be a version of two distinct words. For example, a production of the word PLANT in its no-release variant form would require disambiguation because its no-release counterpart is potentially similar to the word PLAN. In the same way, the lack of a final release in the word BAND is potentially homophonous with its embedded word BAN. However, the sequence forming an embedded word (e.g., 'plan' produced as part of PLANT) may be acoustic-phonetically distinct from token forms produced in isolation (i.e., PLAN). Recent evidence suggests that these acoustic differences are used by listeners during word recognition (e.g., Davis, Marslen-Wilson, & Gaskell, 2002; Salverda, Dahan, & McQueen, 2003). For a word embedded in a no-release variant form (e.g., PLAN in the no-release carrier of PLANT), acoustic-phonetic properties unique to the carrier word may limit activation of the embedded word. In one study investigating degrees of /t/-deletion in Dutch (Janse, Nooteboom, & Quené, 2007), parallel activation of representations (e.g., kas and kast) was observed. However, comparable priming observed for *kast* given canonical (kast) and no-release /t/ (kas/t/) primes in a cross-modal form priming paradigm was interpreted as evidence that lexical competition was resolving toward the intended word (i.e., kast) as the time course progressed. Interestingly, a corpus analysis performed by the authors showed that the frequency of occurrence for such variants in this particular environment was 85%; that is, more frequent than the canonical form. This differs from the present investigation where frequency rates are both lower and more variable in context.

Two complementary series of experiments investigated the role of the frequency of release across segment categories (voiced vs. voiceless) and across following phonological context (vowel vs. consonant) in the recognition of no-release variants. Stimuli ending with voiced (e.g., band) or voiceless (e.g., plant) segments (referred to as no-release carrier words) were produced in sentence contexts and contained an embedded word that ended with the segment preceding the final stop (e.g., ban, plan). The word-final stop consonants used were /d/ for the voiced condition, and /t/ for the voiceless condition as this contrast provided a sufficient number of lexical items containing embedded words. For all no-release carrier words, the word-final release was excised (such stimuli will be visually represented as ban[d] and plan[t] where the [d] and [t] are not present but were part of the original recording environment). A second set of stimuli included both the voiced and voiceless embedded words (e.g., ban, plan) produced and presented unaltered (referred to as voiced or voiceless true embedded words). Experiments 1a–1c measured activation of the embedded word to investigate whether voiceless no-release stimuli were more likely to permit embedded word activation relative to voiced no-release stimuli. Manipulation of the following context (vowel vs. consonant) evaluated whether embedded word activation was reduced given a context favoring a no-release variant (consonant context). The true embedded word condition served as a baseline priming effect against which to compare the no-release carrier words.

2 Experiment 1a

2.1 Method

2.1.1 Participants. A total of 80 undergraduate students received course credit for their participation in the experiment. All participants were native speakers of English and reported no hearing or visual problems.

2.1.2 Materials and design. Two groups of monosyllabic words were chosen: 32 had final /t/ (voiceless) and 32 had final /d/ (voiced) segments. These carrier words contained an embedded word that ended before the final /t/ or /d/ segment. Carrier words were matched for frequency with their embedded word as much as was permitted by language constraints[1].

For each of the stimuli, sentences were constructed that were semantically plausible when the carrier or embedded word was inserted in a medial position. Three versions of the same sentence were recorded corresponding to the three experimental conditions. The first experimental condition included the carrier word followed by a vocalic phonological context. The second condition included the carrier word followed by a consonant context. The third experimental condition included the embedded word followed half of the time by a vowel context and half of the time by a consonant context. Finally, a set of control words were recorded in the same sentence frames in place of the carrier or embedded words. The control words were semantically unrelated to their respective carrier and embedded word but were semantically plausible in the sentence context. Four experimental lists were created. Stimuli were counterbalanced across lists so each participant heard the same number of stimuli in each condition and were rotated across lists so that each participant was exposed to only one of the four possible conditions of each target stimulus.

The sentences were read and recorded in a sound-attenuated booth by a male native speaker of American English. A waveform editor was used to create the no-release versions of the /t/ and /d/ words. Releases, identified by locating visual and auditory evidence of the release after the preceding vowel or consonant offset, were excised along with any closure and aspiration. The waveform was edited at zero crossings such that the token was free of disfluencies. An inaudible time mark was set at the offset of the critical words in each condition; the no-release word in the first and second experimental condition, the embedded word in the third experimental condition, and the unrelated word in the control condition. The timing marks initiated both presentation of the visual target and measurement of lexical decision response latencies.

Sixty-four additional filler sentences were created using other monosyllabic words in medial position within the sentence. Time marks were placed at the offset of these words to indicate when a filler visual target would be presented. All filler targets were non-words that varied from a true monosyllabic word by a single phoneme. Half of the non-word filler targets were related to the auditory probe (i.e., created by changing one phoneme of the auditory probe), and the other half of the non-word filler targets were unrelated to the auditory probe (i.e., created by changing one phoneme of a word that was not in the sentence).

Table 1. Reaction times (and percentage correct) for lexical decision task as a function of prime condition and voicing status. Experiment 1a. Visual target was the embedded word (e.g. BAN, PLAN).

	No-release		True embedded word	Unrelated control
	Vowel context	Consonant context		
Voiced (BAN)	707 (94)	706 (97)	694 (97)	732 (95)
Voiceless (PLAN)	715 (97)	698 (97)	681 (99)	731 (96)

Note. For all experiments, the auditory primes in the *No-release* conditions (Vowel and Consonant context) were unreleased versions of the carrier words (e.g., PLANT/BAND). See Experiment 1a for description of stimulus construction.

2.1.3 Procedure. Experiments were conducted in a sound-attenuated room for one or two participants. Auditory stimuli were presented over headphones at a comfortable listening level. Word and nonword visual targets were presented on a computer screen for 250ms. Participants were instructed to decide whether the visual target was a word or nonword by pressing an appropriately marked button on a response box. A 16-trial practice session was conducted for each participant after being read the instructions. An experimental session lasted approximately 30 minutes.

2.2 Results

Response latencies greater than 1500ms and less than 200ms were excluded from all analyses (less than 1% of responses). Twelve participants who did not meet the criterion of at least 80% accuracy were excluded from all analyses. Two items responded to with less than 75% accuracy were also excluded from all analyses. A significance level of $p < .05$ was adopted for all analyses.

As is evident from Table 1, priming of the embedded word (e.g., BAN and PLAN) was found for all experimental conditions. Priming effects were larger for the true embedded word (e.g., ban and plan) compared to the no-release (e.g., ban[d], plan[t]) conditions. There was no difference between voicing conditions, and only a slight advantage for the no-release consonant context over the vowel context. The results were analyzed using a 2 (voicing: voiced and voiceless) × 4 (prime type: no-release vowel context, no-release consonant context, true embedded word, and unrelated control) ANOVAs for both subjects (*F1*) and items (*F2*). There was a main effect of prime type, $F1(3, 201) = 12.7$, $MSE = 3678$; $F2(3, 180) = 7.5$, $MSE = 2828$. No other effects were significant.

A series of planned comparisons evaluated the prime type effect. The vowel-context stimuli showed a 21ms priming effect, (711ms vs. 732ms), $t1(67) = –2.9$, $SD = 60$; $t2(62) = –2.5$, $SD = 66.2$, the consonant-context stimuli showed a 30ms priming effect, (702ms vs. 732ms), $t1(67) = –4.0$, $SD = 61.8$; $t2(62) = –2.8$, $SD = 81.9$, and the true embedded word showed the most priming: 45ms (687ms vs. 732ms), $t1(67) = –6.3$, $SD = 58.2$; $t2(62) = –4.6$, $SD = 76.1$. *T*-tests on difference scores showed that both no-release types (vowel-context and consonant-context) had smaller priming effects than the embedded word (vowel context: $t(67) = 3$, $SD = 63.6$; consonant context: $t(67) = 2.1$, $SD = 57$). Further, the no-release context conditions (vowel vs. consonant) did not differ. A comparable ANOVA performed on percentage correct showed a main effect of prime type, $F1(3,201) = 4.13$, $MSE = 40$; $F2(3, 180) = 4.16$, $MSE = 24$, indicating fewer errors in the true embedded word condition.

In summary, no-release variants showed priming for embedded word targets albeit less than true embedded word primes. This indicates that the no-release carrier words effectively activated the embedded word but retained properties of the carrier word (rendering the edited token a less effective embedded word). Neither voicing nor phonological context influenced priming effects and suggests that listeners did not use knowledge concerning the probability of occurrence of a release to 'fill in' the missing release[2]. However, the visual targets occurred immediately at the offset of the auditory prime and it is possible that additional processing time would permit access and utilization of the segmental properties of the context. Experiment 1b examined this possibility by presenting the visual target 150ms downstream from the offset of the auditory prime.

3 Experiment 1b

3.1 Method

3.1.1 Participants. A total of 83 participants were recruited from the same population used for Experiment 1a.

3.1.2 Materials and design. The materials used were the same as those in Experiment 1a. The only difference was that visual targets were presented 150ms subsequent to the offset of the auditory prime.

3.1.3 Procedure. All aspects of the procedure were identical to Experiment 1a.

3.2 Results

Response latencies greater than 1500ms and less than 200ms were excluded from all analyses (less than 1% of responses). Three participants were excluded who failed to meet the criterion of at least 80% accuracy. Four items responded to with less than 75% accuracy were also excluded from all analyses. Table 2 shows reaction times and accuracy as a function of voicing and prime type. A 2 (voicing) × 4 (prime type) ANOVA showed a main effect of prime type, $F1(3, 237) = 22.8$, $MSE = 3886$; $F2(3, 174) = 13.9$, $MSE = 2567$. A series of planned comparisons showed priming for the true embedded word (665ms vs. 716ms), $t1(79) = -7.1$, $SD = 64.5$; $t2(60) = -5.5$, $SD = 72.0$. For the no release variants, priming was observed in the consonant (41ms) context only, $t1(79) = -5.7$, $SD = 64.1$; $t2(60) = -4.2$, $SD = 76.8$. The difference between true embedded word and no-release variant consonant context priming was marginal by subjects, $t1(79) = 1.7$, $SD = 53.7$, $p = .09$, but not significant by items, $p > .25$. The main effect of voicing was significant by participants, $F1(1, 79) = 5.6$, $MSE = 2527$, but not by items. A comparable ANOVA performed on percentage correct showed a main effect of prime type, $F1(3, 237) = 4.9$, $MSE = 52$; $F2(3, 174) = 3.8$, $MSE = 26$, with somewhat lower accuracy in the control and vowel contexts relative to the other conditions. No other effects were significant.

Table 2. Reaction times (and percentage correct) for lexical decision task as a function of prime condition and voicing status. Experiment 1b. Visual target was the embedded word (e.g., BAN, PLAN).

	No-release		True embedded word	Unrelated control
	Vowel context	Consonant context		
Voiced	704 (95)	669 (97)	660 (98)	707 (96)
Voiceless	700 (96)	682 (97)	670 (98)	726 (95)

Table 3. Reaction times (and percentage correct) for lexical decision task as a function of prime condition and voicing status. Experiment 1c. Visual target was the embedded word (e.g., BAN, PLAN).

	No-release		True embedded word	Unrelated control
	Vowel context	Consonant context		
Voiced	669 (97)	658 (96)	626 (97)	701 (95)
Voiceless	678 (95)	672 (95)	635 (97)	719 (93)

In summary, priming was observed for the no-release variants in consonant contexts but not vowel contexts. This pattern of priming is unexpected based on corpus statistics that suggest vowel contexts would result in more embedded word priming than the consonant contexts. Experiment 1c sought to localize the source of the unexpected asymmetry by removing the following context. If listeners were relying on acoustic-phonetic information contained in the following context, this manipulation should neutralize this effect. Experiment 1c used the same set of materials with the context following the auditory primes excised. The visual probe was presented 150ms downstream of the auditory prime to maintain temporal consistency with Experiment 1b.

4 Experiment 1c

4.1 Method

4.1.1 Participants. A total of 91 participants were recruited from the same population as the previous experiments.

4.1.2 Materials and design. The materials used were the same as the materials used for Experiment 1b, with the sentence context following the critical auditory prime word excised.

4.1.3 Procedure. The experiments were conducted in the same manner as the previous two experiments, and lasted approximately 20 minutes.

4.2 Results

Response latencies greater than 1500ms and less than 200ms were excluded from all analyses (less than 1% of responses). Data from eleven participants who failed to meet the criterion of at least 80% accuracy were excluded from the analyses. Three items responded to with less than 75% accuracy were also excluded from all analyses. As Table 3 shows, priming was observed across all experimental conditions.

A 2 (voicing) × 4 (prime condition) ANOVA showed a main effect of prime type $F1(3, 237) = 44.1$, $MSE = 3887$; $F2(3, 171) = 20.6$, $MSE = 2871$. Planned comparisons showed significant priming effects for all three prime types (vowel context: $t1(79) = –5.0$, $SD = 65.1$; $t2(61) = –3.6$, $SD = 79.3$; consonant context: $t1(79) = –6.4$, $SD = 62.7$; $t2(61) = –4.9$, $SD = 68.5$; true embedded word: $t1(79) = –9.7$, $SD = 73.8$; $t2(61) = –7.4$, $SD = 81.1$). The difference between the vowel and consonant context priming effects (36ms vs. 45 ms) was not significant. A main effect of voicing was found across participants, $F1(1, 79) = 7.7$, $MSE = 3265$, but not items.

A comparable ANOVA performed on percentage correct showed a main effect of prime type, $F1(3, 237) = 3$, $MSE = 50$, $F2(3, 177) = 2.6$, $MSE = 22$, $p = .052$. As in the previous experiments, the control condition showed slightly less accurate responses than the no-release and true embedded word primes.

In summary, true embedded word primes showed the most priming while the no-release vowel and consonant contexts showed comparable priming. Furthermore, removing the following context resulted in comparable priming across the consonant and vowel contexts for no-release words. This suggests that the vowel context (Exp. 1b) provided better support for the carrier word and thereby was a more effective inhibitor of the embedded word.

5 Discussion of Experiments 1a–1c

Three experiments examined activation of embedded words by no-release-variant carrier words presented in vowel and consonant contexts. There was no evidence that listeners used probabilistic knowledge concerning the likelihood of release during processing. In fact, the observed asymmetry across consonant and vowel context priming when the visual probe was presented 150ms into the following word (Exp. 1b) was in the direction opposite predicted by corpus statistics. However, presentation of the stimuli with the following context excised (Exp. 1c) resulted in comparable vowel and consonant priming. This suggests that the vowel context provided co-articulatory information to support the carrier word interpretation (and thereby inhibited the embedded word interpretation).

A second major finding was that embedded word activation for no-release carrier primes (ban[d] plan[t]) was generally smaller than for true embedded primes (ban, plan). The no-release carrier words were uncharacteristic tokens of the embedded words and these tokens simply functioned as less effective primes. This finding is consistent with other work showing that subtle acoustic details map onto lexical representations and modulate activation (Salverda et al., 2003). On the other hand, the no-release carrier tokens were sufficiently similar to embedded words that such primes did activate the (embedded word) target. The probability of a release associated with the carrier words or context does not appear to modulate activation of the embedded word. To anticipate the next series of results (see Exp. 2a–2c below), listeners appear to activate both the embedded and carrier word lexical hypotheses in parallel but strong activation for a carrier word can inhibit activation of the embedded word.

Experiments 1a–1c provided a profile of embedded word activation levels. In the following experiments (Experiments 2a–2c) we assess carrier word activation given no-release variant and true embedded word primes. To this end, the carrier words (e.g., PLANT and BAND) were presented as visual targets. Of interest is whether activation of the carrier word is influenced by voicing and/or by following context. True embedded word primes in this series of experiments serve a different purpose than Experiments 1a–1c. Here, the true embedded words provide a means to evaluate carrier word activation in a condition without any possible acoustic support for the carrier word. Experiment 2a paralleled Experiment 1a and presented the sentences intact and the visual target at the offset of the prime.

6 Experiment 2a

6.1 Method

6.1.1 Participants. A total of 80 participants were recruited from the same population used for Experiments 1a–1c.

6.1.2 Materials and design. The auditory materials were the same as those used for Experiment 1a. The visual targets were the carrier word.

Table 4. Reaction times (and percentage correct) for lexical decision task as a function of prime condition and voicing status. Experiment 2a. Visual target was the carrier word (e.g., BAND/PLANT).

	No-release		True embedded word	Unrelated control
	Vowel context	Consonant context		
Voiced	660 (97)	670 (98)	715 (95)	741 (94)
Voiceless	677 (97)	704 (96)	760 (93)	742 (92)

6.1.3 Procedure. All aspects of the procedure were identical to Experiment 1a.

6.2 Results

All response latencies greater than 1500ms and less than 200ms were excluded from analyses (less than 1% of responses). Data from eight participants who failed to meet the criterion of at least 80% accuracy were excluded from the analyses. Seven items responded to with less than 75% accuracy were also excluded from all analyses.

As Table 4 shows, the voicing conditions showed a vowel/consonant context asymmetry; for voiceless targets (e.g., PLANT), vowel context priming was larger than consonant context but no such difference was apparent for the voiced targets (e.g., BAND). The true embedded word primes also showed voicing differences: the voiced condition showed priming relative to the unrelated control, whereas the voiceless condition did not.

A 2 (voicing) × 4 (prime type) ANOVA showed main effects of voicing, $F1(1, 71) = 24.3$, $MSE = 3476$; $F2(1, 55) = 3.9$, $MSE = 10936$, and prime type, $F1(3, 213) = 44.3$, $MSE = 4298$; $F2(3, 165) = 32.9$, $MSE = 2567$. However, these effects were mediated by an interaction that was significant by participants but not by items, $F1(3, 213) = 3.3$, $MSE = 4126$; $F2(3, 165) = 1.7$, $MSE = 2567$, $p = .17$.

An asymmetry in the priming effect across phonological context was observed. For voiceless no-release stimuli (e.g., plan[t]), priming effects were larger in the vowel context compared to the consonant context, $t1(71) = –3.0$, $SD = 79.9$; $t2(27) = –2.0$, $SD = 71.1$, $p = .06$, whereas for voiced no-release stimuli (e.g., ban[d]), priming in consonant and vowel contexts was nearly identical.

For true embedded words, the voiced true embedded stimuli led to significant priming by participants but marginal by items, $t1(71) = –2.2$, $SD = 103$; $t2(28) = –1.6$, $SD = 75.6$, $p = .12$, that was smaller than both the no-release vowel and no-release consonant context conditions, $t1(71) = –5.4$, $SD = 86.1$; $t2(28) = –3.5$, $SD = 85.7$; $t1(71) = –4.5$, $SD = 84.2$; $t2(28) = –3.4$, $SD = 69.8$ respectively. Voiceless true embedded words showed no priming.

A comparable ANOVA performed on percentage correct showed a main effect of prime type $F1(3, 213) = 9$, $MSE = 69$; $F2(3, 165) = 6.4$, $MSE = 6.4$ with slightly less accurate responding in the control and true embedded word conditions. The main effect of voicing was significant across participants, $F1(1, 71) = 4.2$, $MSE = 83$, but not items. No other effects were significant.

In summary, there were three major findings from Experiment 2a. First, there was a context asymmetry for voiced and voiceless no-release carrier stimuli. Voiceless stimuli (e.g., plan[t]) showed more priming in the vowel context while voiced stimuli (e.g., ban[d]) showed equivalent priming across consonant and vowel contexts. Second, there was a voicing asymmetry for true embedded words. Voiced embedded words (e.g., ban) resulted in priming but voiceless embedded words (e.g., plan) did not. Third, the no-release carrier words served as better primes than embedded words. Two additional experiments investigate these effects. Experiment 2b paralleled

Table 5. Reaction times (and percentage correct) for lexical decision task as a function of prime condition and voicing status. Experiment 2b. Visual target was the carrier word (e.g., BAND/PLANT).

	No-release		True embedded word	Unrelated control
	Vowel context	Consonant context		
Voiced	667 (98)	679 (98)	713 (96)	746 (94)
Voiceless	680 (96)	708 (95)	762 (93)	760 (93)

Experiment 1b and examined the role of increasing the availability of the context information by presenting the visual target 150 ms downstream from offset of the prime.

7 Experiment 2b

7.1 Method

7.1.1 Participants. Sixty-four participants were recruited from the same population as previous experiments.

7.1.2 Materials and design. The materials used were the same as those used for Experiment 2a except that the visual probe occurred 150ms subsequent to the offset of the prime.

7.1.3 Procedure. This experiment was conducted in the same manner as the previous experiments.

7.2 Results

Response latencies greater than 1500ms and less than 200ms were excluded from all analyses. Data from eight participants who failed to meet the criterion of at least 80% accuracy were excluded from the analyses. Nine items responded to with less than 75% accuracy were also excluded from all analyses

As Table 5 shows, and similar to Experiment 2a, a vowel/consonant context priming asymmetry was observed for the voiceless, but not voiced, no-release carrier primes.

A 2 (voicing) × 4 (prime type) ANOVA showed a main effect of voicing status that was marginal across items, $F1(1, 55) = 11.9$, $MSE = 3915$; $F2(1, 56) = 2.1$, $MSE = 12351.9$, $p = .10$, and a main effect of prime type, $F1(3, 165) = 42.9$, $MSE = 3784$; $F2(3, 168) = 35.8$, $MSE = 2679$. The interaction between voicing status and prime type was marginally significant across both subjects and items, $F1(3, 165) = 2.1$, $MSE = 3544.3$, $p = .10$; $F2(3, 168) = 1.8$, $MSE = 2679.2$, $p = .14$.

An asymmetry across contexts similar to Experiment 2a was observed. Priming effects in the voiceless vowel context were larger than the voiceless consonant context, $t1(55) = –3.3$, $SD = 67$; $t2(28) = 28$, $SD = 69$, whereas priming effects across the voiced consonant and vowel contexts did not differ. Also similar to Experiment 2a, only voiced embedded word primes showed significant priming, $t1(55) = –3.1$, $SD = 83$; $t2(28) = –2.7$, $SD = 70$.

A comparable ANOVA performed on percentage correct showed a main effect of voicing that was marginal across items $F1(1, 55) = 13.9$, $MSE = 50$; $F2(1, 56) = 3$, $MSE = 54$, $p = .08$. Consistent with previous experiments, the main effect of prime type was significant $F1(1, 165) = 8.8$, *MS*E

= 50; $F2(3, 168) = 4.4$, $MSE = 54$, with somewhat lower accuracy in the control and true embedded word conditions. No other effects were significant.

In summary, Experiments 2a and 2b showed a context asymmetry for the voiced and voiceless carrier words; comparable consonant and vowel context priming for the voiced targets (e.g., BAND) and larger vowel context priming for the voiceless targets (e.g., PLANT). Similarly, embedded words showed a voicing asymmetry – voiced embedded words (e.g., ban) activated their carrier word counterpart (e.g., BAND) but voiceless words did not. While the voicing × prime type interactions were marginal for both Experiments 2a and 2b, a post-hoc voicing × prime type × experiment (Exp. 2a vs. 2b) ANOVA showed a significant voicing × prime-type interaction for both participants $F1(3, 756) = 4.9$, $MSE = 3973$, and items $F2(3, 333) = 3.5$, $MSE = 2623$. Further, the variable of experiment did not enter into any significant interactions.

Experiment 2c uses logic similar to Experiment 1c – the following context was removed and the visual carrier word target was presented 150ms from the offset of the prime word.

8 Experiment 2c

8.1 Method

8.1.1 Participants. Sixty participants were recruited from the same population as the previous experiments.

8.1.2 Materials and design. The auditory stimuli used for this experiment were the same as those used for Experiment 1c (sentences without the following context).

81.3 Procedure. All aspects of the experiment were identical to previous experiments.

8.2 Results

Response latencies greater than 1500ms and smaller than 200ms were excluded (less than 1% of the data). Six items responded to with less than 75% accuracy were also excluded from all analyses.

Table 6 shows that all experimental conditions facilitated responding compared to the unrelated controls. The no-release conditions show a larger overall advantage compared to the true embedded word prime condition. For both voiced and voiceless true embedded word conditions, priming seems to be equivalent.

A 2 (voicing) × 4 (prime condition) ANOVA showed a main effect of prime type, $F1(3, 177) = 56.2$, $MSE = 4455$; $F2(3, 168) = 29.0$, $MSE = 4003$. Planned comparisons showed significant priming for the no-release vowel (87ms; $t1(659) = –9.9$, $SD = 71.1$; $t2(57) = –8.3$, $SD = 82.5$) and

Table 6. Reaction times (and percentage correct) for lexical decision task as a function of prime condition and voicing status. Experiment 2c. Visual target was the carrier word (e.g., BAND/PLANT).

	No-release		True embedded word	Unrelated control
	Vowel context	Consonant context		
Voiced	662 (98)	658 (96)	729 (95)	767 (91)
Voiceless	702 (97)	696 (97)	743 (92)	772 (89)

consonant condition (92ms; $t1(59) = -10.2$, $SD = 72.2$; $t2(57) = -8.4$, $SD = 82.2$). While the priming effect for the true embedded words was significant (34ms; $t1(59) = -3.9$, $SD = 73.2$; $t2(57) = -2.7$, $SD = 92.8$), it was reduced compared to the no-release carrier words (vowel context: $t1(59) = -6.8$, $SD = 61.6$; $t2(57) = -4.3$, $SD = 102$; consonant context $t1(59) = -7.1$, $SD = 64$; $t2(57) = -4.8$, $SD = 91.6$.). There was also a main effect of voicing by participants and marginally by items, $F1(1,59) = 25.3$, $MSE = 4074$; $F2(1,56) = 3.1$, $MSE = 14172$, $p = .09$.

A comparable ANOVA for percentage correct showed a main effect of prime, $F1(3, 177) = 14.9$, $MSE = 86$; $F2(3, 168) = 14.8$, $MSE = 41$, with lower accuracy in the control and embedded word conditions. No other effects were significant.

In summary, the asymmetry across context as a function of voicing observed in Experiments 2a and 2b was no longer evident in Experiment 2c. Instead, voiced and voiceless no-release carrier words showed similar priming patterns in the absence of following context. Additionally, the asymmetry across voicing for embedded word primes observed in Experiments 2a and 2b was similarly not evident. Instead, true embedded words showed similar priming across voicing.

9 Discussion of Experiments 2a–2c

Overall, activation of the carrier word was influenced by voicing and by the following context. Voiced targets (e.g., BAND) showed comparable priming for consonant and vowel contexts while voiceless targets (e.g., PLANT) showed larger priming for vowel contexts. There was a voicing asymmetry for true embedded words; voiced embedded words primed the carrier word but voiceless embedded words did not. Excising the following context neutralized the differences across voicing stimuli for both carrier words and for embedded words[3]. In all experiments, no-release carrier words showed more priming than true embedded words.

10 General discussion

Two series of experiments investigated how no-release variants in embedded word contexts are accommodated during spoken word recognition. Of particular interest was the role of frequency of occurrence based on voicing status and connected speech context on activation of carrier and embedded words. The interplay of activation between an embedded and carrier word was used to investigate the role of these variables and permitted an investigation of the time course of parallel activation for carrier words and embedded word competitors.

Does release frequency influence activation levels in potentially lexical ambiguous contexts? Priming of carrier word targets by true embedded word primes suggests that it does. When following context was included (Exp. 2a, 2b), the voiced embedded words (e.g., ban) showed more priming of the carrier word target (e.g., BAND) than the voiceless embedded words (e.g., plan/PLANT). Therefore, greater priming was observed for stimuli whose distributional frequency statistics favor a carrier word interpretation of the primes – despite the fact that these words do not possess obvious markers for the presence of a word-final stop. Is this asymmetry in voicing found for carrier word activation due to the influence of phonological inference based on distributional probabilistic knowledge? There are a number of patterns in the data that suggest that it is. First, the true embedded word primes showed the same pattern as the no-release carrier word primes, that is, greater priming for the voiced targets. This implicates the role of phonological inference in processing the true embedded word primes because these stimuli (produced as true embedded words) did not carry acoustic-phonetic markers of a missing final consonant release. Therefore, it implicates a process that infers the intended representation.

Second, the observed differences across voicing (Exp. 2a and 2b) were not due to acoustic information in the primes since the voicing asymmetry disappeared when following context was removed (Exp. 2c). In other words, the voicing asymmetry disappeared when the target stimulus was no longer part of connected speech. This is consistent with other work on phonological variants suggesting that connected speech is a necessary condition for listeners to use knowledge concerning speech events that happen in connected speech (Marslen-Wilson, Nix, & Gaskell, 1995). In the absence of this necessary condition, inferential mechanisms regarding connected speech events are not engaged. Third, word boundary co-articulatory information in the no-release carrier primes influenced activation levels of both the carrier and embedded targets: activation of the no-release carrier word representations in the vowel context (Exp. 2a and 2b) blocked activation of the voiced embedded word representations (Exp. 1a and b).

These combined patterns of results suggest a mechanism underlying the voicing asymmetries for the carrier words that is similar in kind to that found for the embedded words. Both patterns converge on a role for phonological inference of the missing information that is tied to the experienced frequency of the missing information. Voiced and voiceless consonants missing releases are not treated as a class but are processed in a way that is sensitive to the probability of a missing release.

Our findings provide new information concerning the problem of embedding in connected speech as well as confirm theoretical claims (and empirical demonstrations) of parallel activation. Parallel activation of the embedded and carrier words was demonstrated. The potential ambiguity introduced by a process in connected speech (no-release variants) was further accommodated by sensitivity to the relative frequency of the variant. This suggests that listeners are able to simultaneously accommodate both variability in form and resolve ambiguity introduced by embedding by appealing to frequency-sensitive information about lexical form. In the case of missing releases, embedded word activation is modulated as a function of carrier word likelihood based on probability of the release. In this case then, acoustic-phonetic properties unique to the carrier words did not limit embedded word activation. The present results are consistent with findings in Dutch that no-release variants (kas[t]#b) activate both the carrier and embedded word representations (Janse et al., 2007). Interestingly, embedded representations were shown to be inhibited subsequent to the offset of the following context (i.e., /b/) suggesting that the high probability for these no-release variants (85%) was an effective constraint. Both patterns of data are consistent with the suggestion by Davis et al. (2002) that the problem of embedding may be overestimated in connected speech.

The present study also investigated whether information occurring subsequent to word offset was used to accommodate the processing of no-release variants. Our results suggest that co-articulatory information contained in the context following the no-release variants influenced processing. However, the observed asymmetry between the vowel and consonant contexts was not mediated by differences in their release probabilities. Instead, co-articulation was a stronger determinant in the pattern of activation observed, as is evidenced by the absence of any asymmetry when the following context was not present. These findings are consistent with other research that demonstrates the contribution of co-articulatory information across word boundaries (e.g., Grosjean, 1985).

Why, then, are context probabilities not used in the same way as voicing probabilities during the processing of no-release variants? The most obvious explanation is that context probabilities about the likelihood of a release are not used because the contextual occurrence does not force the surface form. In contrast to place assimilation in which an assimilated segment occurs only in the context of a licensing context, American English no-release variants are much more variable (18–96% for vowel context and 4–21% for consonant context). Further, the co-articulatory spillover advantage for a vowel context is presumably a general characteristic of speech (at least in our sample). Given this as a typical state of affairs, the context probabilities would have to work in the opposite direction to the co-articulatory influences in the signal. The possibility that listeners would develop the knowledge necessary to utilize the weak context probabilities given a co-articulatory influence precisely in the opposite direction seems remote. In contexts where a given surface form is almost always encountered (e.g., in Dutch, 85% no-release /t/ frequency in /st#b/ contexts), context probabilities have been shown to guide interpretation (Janse, Nooteboom, & Quené, 2007; see also Mitterer & Ernestus, 2006). This comparison with the present results shows the importance of investigating such variation in its full phonological context.

Our results are also consistent with recent work on another context conditioned variant, word final flaps. In a corpus analysis, Ranbom, Connine, and Yudman (2009) found that the word final flap variant occurs relatively frequently compared to the citation form [t] variant and is only probabilistically constrained by phonemic context (a word final flap is more likely to be found when followed by a word beginning with a vowel but the word final flap is also found in the absence of a vowel context). The experienced distribution of the flap is reflected in lexical processing in that lexical activation is not influenced by contextual constraint. Ranbom and Connine (2007) argue that a phonological inference mechanism is not utilized precisely because the context is not sufficiently predictive of the flap.

What is the nature of the phonological inference process for missing release forms? This question can be considered in light of fundamental issues in spoken word recognition. Is this process a pre-lexical or post-lexical process? A pre-lexical process would imply that statistical likelihoods apply prior to lexical knowledge concerning its relevance. One issue that a pre-lexical process must address is the potential for excessive errors – in the absence of lexical knowledge concerning the relevance of a pre-lexical inference, an inferred voiced or voiceless segment would often be an error. In a post-lexical process, phonological inference would come into play after a word has been activated. At issue for a post-lexical process is the correct characterization of the phonological knowledge and its use. Traditionally, phonological knowledge has been characterized as a phonological rule within a larger linguistic system of knowledge. Satisfaction of the relevant constraints serves to implicate the utilization of the general rule. On this view, phonological knowledge is represented by a listener as an abstraction across a class of words and its use is triggered when a word satisfies the constraints of the rule. Alternatively, knowledge concerning phonological variation may be represented *within* lexical entries for individual items that are experienced in this variant form (see LoCasto & Connine, 2002). The advantage of this representational scheme is that the dimension of lexical processing (lexical activation) that triggers its use co-exists with its application. This would effectively bind the experienced surface frequency of a word with its lexical representation. For no-release variants, the experienced frequency of a form will determine the extent to which it is easily recognizable as the intended word. On this view, knowledge of surface variation is captured in the representational structure of the lexicon. Recognition of a variant, in either its more frequent or less frequent form, would not require any special or additional mechanisms. A functional consequence of this scheme is differential effects for individual lexical items– with stronger effects observed for items that are consistently experienced as a no-release variant.

What seems to be clear from the available literature is that an attempt to generalize the question (i.e., such as "*at what level is rule-governed variation resolved?*") is likely to lead to the oversimplification of the issues involved. First, it is unlikely that the various forms of rule governed variation have a single solution. Solutions may be sensitive to the frequency of contact with a given form of variation (see Connine, 2004; Ranbom & Connine, 2007) and distributed probabilities associated with its realization (e.g., Deelman & Connine, 2001; Janse et al., 2007). Depending on the nature of the variation, it may be resolved at a language-independent auditory level. For instance, there is evidence that listeners do not need experience with processing assimilation in order to process it successfully (e.g., Gow & Im, 2004; Mitterer, Csépe, & Blomert, 2006; Mitterer et al., 2008). However, additional experience may subsequently influence how assimilated speech is processed (see Coady, Kluender, & Rhode, 2003; Mittereret al., 2006, for similar discussion). Additionally, it is probable that for a given form of variation, the method of resolution depends on mechanisms that span various levels of analysis. For instance, investigation of /t/ reduced variants in Dutch highlights the role of general auditory processes, lexical constraints and phonological learning (Mitterer et al., 2008). Finally, even across languages that exhibit similar processes, different forms of resolution may be utilized depending on other structural differences across languages. For instance, investigation of voicing assimilation in Dutch (Quené, van Rossum, & van Wijck, 1998) has shown that listeners use the voicing value of an assimilated segment to predict the voicing of the following consonant. The dynamics in processing voice assimilation in Dutch may be different than place assimilation in English because word-final voiced plosives do not occur other than as a consequence of regressive voicing assimilation in Dutch.

Finally, as discussed above, the solution adopted for the processing of phonological variation seems to implicate both general auditory and language-specific processes. The degree to which processes are relied upon may depend on both of these factors and may change as one's experience with a given type of variation increases. Of particular interest here are potential dissimilarities in the pattern of results that would be produced by listeners whose native language's phonological properties regarding a final-word release vary from those of the English language. Spanish is an example of such a language. Voiceless stop consonants do not occur in word-final position (with a few exceptions, consisting of borrowed words); and the only voiced stop consonant that appears at word offset is the phoneme /d/, which is never released when found in word-final position. Investigating this difference between American English and Spanish in the context of native speakers of Spanish learning English as a second language could profitably be used to track, on-line, the degree to which general auditory processes can accommodate such variation, and how additional language experience of English may subsequently affect the processing of word final no-release variants.

Acknowledgements

This research was supported by NIDCD Grant R01 DC02134 to Cynthia Connine. Additional support was provided by the Center for Psycholinguistic and Cognitive Science at Binghamton University. Portions of this research were presented at the 142nd Meeting of the Acoustical Society of America. We thank David Patterson for his help throughout the project. We also thank Marna Schwartz, Julie Rosenfeld, Kit Yeung, and Katie Chesley for their help with data collection.

Notes

1 The average frequency (Francis & Kucera, 1982) of the embedded and carrier words were: Voiced embedded word = 170.8 (range = 0–899, *SD* = 208.7), Voiced carrier word = 90.1 (range = 0–356, *SD* = 83.8), voiceless embedded word = 103.2 (range = 1–2003, *SD* = 380), voiceless carrier word = 60 (range = 0–400, *SD* = 113). The observed difference between the embedded and carrier frequencies were due mainly to two voiceless embedded words (frequencies of 772 and 2003) and two voiced embedded words (frequencies of 689 and 899). Consideration of the remaining stimuli moved the frequency average more into line with their carrier word counterparts (89.6 and 57.1, voiced and voiceless respectively). A series of paired comparisons (including all stimuli) showed no statistical differences in frequency between embedded words and their carrier word counterparts, between voiced and voiceless embedded words and between voiced and voiceless carrier words.

2 The current stimuli were artificially created and so the distributional probabilities do not necessarily directly apply to our stimuli. First, we should point out that such 'artificial' stimuli have been shown to pattern with said distributional frequencies (see Deelman & Connine, 2001). Further, such variation is not categorical but rather free to vary. Therefore our stimuli can be seen to represent the extreme end of such variation. Finally, using natural stimuli (words recorded without their releases) loses some control about the presence/degree of the release and also leaves free to vary any other acoustic-phonetic event that may pattern systematically with the lack of a release. In an attempt to ask specifically whether phonological knowledge (viewed as abstracted rules) is used during no-release variant processing, it was important to, as much as possible, limit any such systematic acoustic-phonetic cues that might signal 'no-release variant' status.

3 As a final test, we conducted a pilot experiment where the co-articulation in the contexts was neutralized by cross-splicing sentence fragments. In these sentences, the following contexts were produced with a segment other than a stop and thus contained no useful co-articulatory spill over into the following context. A group of 38 participants were run on a subset of the materials (48 items, half from each voicing category) and conditions (no-release words with consonant and vowel contexts and an unrelated control). As in Experiment 2c, a priming effect was found, $F(2, 76) = 20.1$, $MSE = 5863$, but there was no evidence for utilization of the following context in assessing release probability. There was no main effect of context type, $F < 1$; 735ms, 751ms, 808ms; vowel, consonant and unrelated control, respectively.

References

Coady, J .A., Kluender, K. R., & Rhode, W. S. (2003) Effects of contrast between onsets of speech and other complex spectra. *Journal of the Acoustical Society of America, 114*, 2225–2235.

Connine, C. M. (2004). It's not what you hear but how often you hear it: On the neglected role of phonological variant frequency in auditory word recognition. *Psychological Bulletin and Review, 11*, 1084–1089.

Connine, C. M., Blasko, D. G., & Wang, J. (1994). Vertical similarity in spoken word recognition: Multiple lexical activation, individual differences, and the role of sentence context. *Perception & Psychophysics, 56*, 624–636.

Crystal, T. H., & House, A. S. (1988). The duration of American-English stop consonants: an overview. *Journal of Phonetics, 16*, 285–294.

Davis, M. H., Marslen-Wilson, W. D., & Gaskell, M. G. (2002). Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 218–244.

Deelman, T., & Connine, C. M. (2001). Missing information in spoken word recognition: Nonreleased stop consonants. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 656–663.

Francis, W. N., & Kucera, H. (1982). *Frequency Analysis of English Usage: Lexicon and Grammar*. Boston, MA: Houghton Mifflin.

Frauenfelder, U. H., & Peeters, G. (1990). Lexical segmentation in TRACE: An exercise in simulation. In G. T. M. Altman (Ed.), *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives* (pp. 50–86). MIT Press.

Gow, D. W., & Im, A. M. (2004). A cross-linguistic examination of assimilation context effects. *Journal of Memory and Language, 51*, 279–296.

Grosjean, F. (1985). The recognition of words after their acoustic offset: evidence and implications. *Perception & Psychophysics, 38*, 299–310.

Janse, E., Nooteboom, S. G., & Quené, H. (2007). Coping with gradient forms of /t/-deletion and lexical ambiguity in spoken word recognition. *Language and Cognitive Processes, 22*, 161–200.

LoCasto, P. C., & Connine, C. M. (2002). Rule-governed missing information in spoken word recognition: Schwa vowel deletion. *Perception & Psychophysics*, *64*, 208–219.

Marslen-Wilson, W., Nix, A., & Gaskell, G. (1995). Phonological variation in lexical process: Abstractness, inference and English place assimilation. *Language and Cognitive Processes, 10*, 285–308.

Marslen-Wilson, W. D., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: words, phonemes, and features. *Psychological Review, 101*, 653–675.

Martin, J. G., & Bunnell, H. T. (1981). Perception of anticipatory coarticulation effects. *Journal of the Acoustical Society of America, 69*, 559–567.
McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1–18.
Mitterer, H., Csépe, V., & Blomert, L. (2006). The role of perceptual integration in the perception of assimilation word forms. *Quarterly Journal of Experimental Psychology, 59*, 1395–1424.
Mitterer, H., & Ernestus, M. (2006). Listeners recover /t/s that speakers reduce: Evidence from /t/-lenition in Dutch. *Journal of Phonetics, 34*, 73–103.
Mitterer, H., Yoneyama, K., & Ernestus, M. (2008). How we hear what is hardly there: Mechanisms underlying compensation for /t/-reduction in speech comprehension. *Journal of Memory and Language, 59*, 133–152.
Nooteboom, S., Janse, E., Quené, H., & te Riele, S. (2000). *Proceedings from the Spoken Word Access Processes workshop* (pp. 15–18). Nijmegen, The Netherlands.
Norris, D., McQueen, J., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Brain and Behavioral Sciences, 23*, 299–370.
Quené, H., van Rossum, M., & van Wijck, M. (1998). Assimilation and anticipation in word perception. *Paper presented at the 5th International Conference on Spoken Language Processing.* Sydney, Australia.
Ranbom, L. R., & Connine, C. M. (2007). Lexical representation of phonological variation. *Journal of Memory and Language, 57*, 273–298.
Ranbom, L. R., Connine, C. M., & Yudman, E. (2009). Is phonological context always used to recognize variant forms in spoken word recognition? The role of variant frequency and context distribution. *Journal of Experimental Psychology: Human Perception & Performance, 15*, 1205–1220.
Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *Journal of the Acoustical Society of America, 51*, 1296–1303.
Salverda, A. P., Dahan, D., & McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition, 90*, 51–89.
Streeter, L. A., & Nigro, G. N. (1979). The role of medial consonant transitions in word perception. *Journal of the Acoustical Society of America, 65*, 1533–1541.
Sumner, M., & Samuel, A. G. (2005). Perception and representation of regular variation: The case of the final /t/. *Journal of Memory and Language, 52*, 322–338.
Warren, P., & Marslen-Wilson, W. (1987). Continuous uptake of acoustic cues in spoken word recognition. *Perception & Psychophysics, 41*, 262–275.
Warren, P., & Marslen-Wilson, W. (1988). Cues to lexical choice: Discriminating place and voice. *Perception & Psychophysics, 43*, 21–30.
Whalen, D. H. (1984). Subcategorical phonetic mismatches slow phonetic judgments. *Perception & Psychophysics, 35*, 49–64.
Whalen, D. H. (1991). Subcategorical phonetic mismatches and lexical access. *Perception & Psychophysics, 50*, 351–360.
Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition, 32*, 25–64.