Representation of Lexical Form

Conor T. McLennan
Cleveland State University, c.mclennan@csuohio.edu

Paul A. Luce
University at Buffalo, The State University of New York, luce@buffalo.edu

Jan Charles_Luce
Univesity at Buffalo, The State University of New York, cdscluce@buffalo.edu

Recommended Citation
McLennan, C. T., Luce, P. A., & Charles-Luce, J. (2003). Representation of lexical form. Journal of Experimental Psychology: Learning, Memory, and Cognition, 29(4), 539-553. doi:10.1037/0278-7393.29.4.539

Representation of Lexical Form

Conor T. McLennan, Paul A. Luce, and Jan Charles-Luce
University at Buffalo, The State University of New York

The authors attempted to determine whether surface representations of spoken words are mapped onto underlying, abstract representations. In particular, they tested the hypothesis that flaps—neutralized allophones of intervocalic /t/s and /d/s—are mapped onto their underlying phonemic counterparts. In 6 repetition priming experiments, participants responded to stimuli in 2 blocks of trials. Stimuli in the 1st block served as primes and those in the 2nd as targets. Primes and targets consisted of English words containing intervocalic /t/s and /d/s that, when produced casually, were flapped. In all 6 experiments, reaction times to target items were measured as a function of prime type. The results provide evidence for both surface and underlying form-based representations.

Information-processing theories have typically characterized spoken word perception as being composed of a series of linguistic stages of analysis, with form-based (or sound-based) representations becoming successively more abstract at each stage of processing. Studdert-Kennedy (1974) provided one of the earliest explicit articulations of this kind of *mediated lexical access model*, which itself drew inspiration from linguistic theory (see Bloomfield, 1933; Chomsky & Halle, 1968; Harris, 1955; Kenstowicz & Kisseberth, 1979; see also Jusczyk & Luce, 2002, for a discussion). More contemporary examples of mediated access can be found in computational models of spoken word recognition such as TRACE (McClelland & Elman, 1986), Shortlist (Norris, 1994), and PARSYN (Luce, Goldinger, Auer, & Vitevitch, 2000).

Recently, these mediated access models have been challenged by proposals that eschew the postulation of intermediate representations. According to *direct access models*, after the initial recoding of sensory data, information is mapped directly onto form-based lexical representations. For example, Stevens's lexical access from features (LAFF) model (see Klatt, 1989) and Marslen-Wilson and Warren's (1994; see also Lahiri & Marslen-Wilson, 1991) direct access featural model propose that lexical representations are accessed directly from binary phonetic features. Similarly, Klatt's (1989) LAFS (Lexical Access from Spectra) model proposes that only context-sensitive spectra are computed in the process of mapping waveform onto word. In short, although both mediated and direct access theories assume that sensory information is initially recoded in some manner, they differ as to whether additional levels of representation intervene between sensory recoding and lexical representation.

Examples of direct and mediated models are illustrated in Figure 1. According to *extreme* direct access models, auditory representations based on initial sensory recoding are mapped directly onto form-based representations, which are then used to contact lemmas. In contrast, mediated access models posit some form of intermediate representations between initial recoding and lexical representation, illustrated in Figure 1 as allophones, phonemes, and syllables.

Evidence in support of direct access models comes from a series of experiments reported by Marslen-Wilson and Warren (1994). Building on earlier work by Whalen (1984, 1991) and Streeter and Nigro (1979), Marslen-Wilson and Warren generated a set of cross-spliced words and nonwords, creating *subcategorical mismatches*. For example, the initial consonant and vowel of the word *jog* were spliced onto the final consonant of the word *job,* resulting in a mismatch between the information in the vowel (which is consistent with a final /g/) and the spliced final consonant /b/. Both mediated and direct access theories predict processing costs when words are cross spliced with other words because of conflicting information at the lexical level. In the *jog/job* example, although information in the vowel is consistent with the word *jog*, the actual final consonant (/b/) is consistent with the word *job*. Thus, both *jog* and *job* may be activated and compete for recognition.

Marslen-Wilson and Warren also spliced nonwords with other nonwords. For example, the initial consonant and vowel of the nonword *smod* were spliced onto the final consonant of the nonword *smob*. Although both direct and mediated access theories predict conflicts when two words are cross spliced, only mediated access theories predict processing costs when nonwords are cross spliced with other nonwords. This prediction is based on the assumption that there will be conflicting cues at a *sublexical* level. In the *smod/smob* example, although information in the vowel is consistent with a final /d/, the actual final consonant is /b/. Poten-

Conor T. McLennan and Paul A. Luce, Department of Psychology and Center for Cognitive Science, University at Buffalo, The State University of New York; Jan Charles-Luce, Department of Communicative Disorders and Sciences and Center for Cognitive Science, University at Buffalo, The State University of New York.

Portions of this work were presented at the 14th International Congress of Phonetic Sciences, August, 1999, San Francisco. This research was supported in part by Research Grant R01 DC 0265801 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

We thank Steve Goldinger, Cynthia Connine, Ken Forster, and LouAnn Gerken for helpful discussions regarding various aspects of this project and Theofanis Pantazis and Melissa Pfennig for their help preparing stimuli and running participants.

Correspondence concerning this article should be addressed to Conor T. McLennan or Paul A. Luce, Language Perception Laboratory, 245 Park Hall, Department of Psychology, University at Buffalo, The State University of New York, Buffalo, New York 14260. E-mail: mclennan@buffalo.edu or luce@buffalo.edu

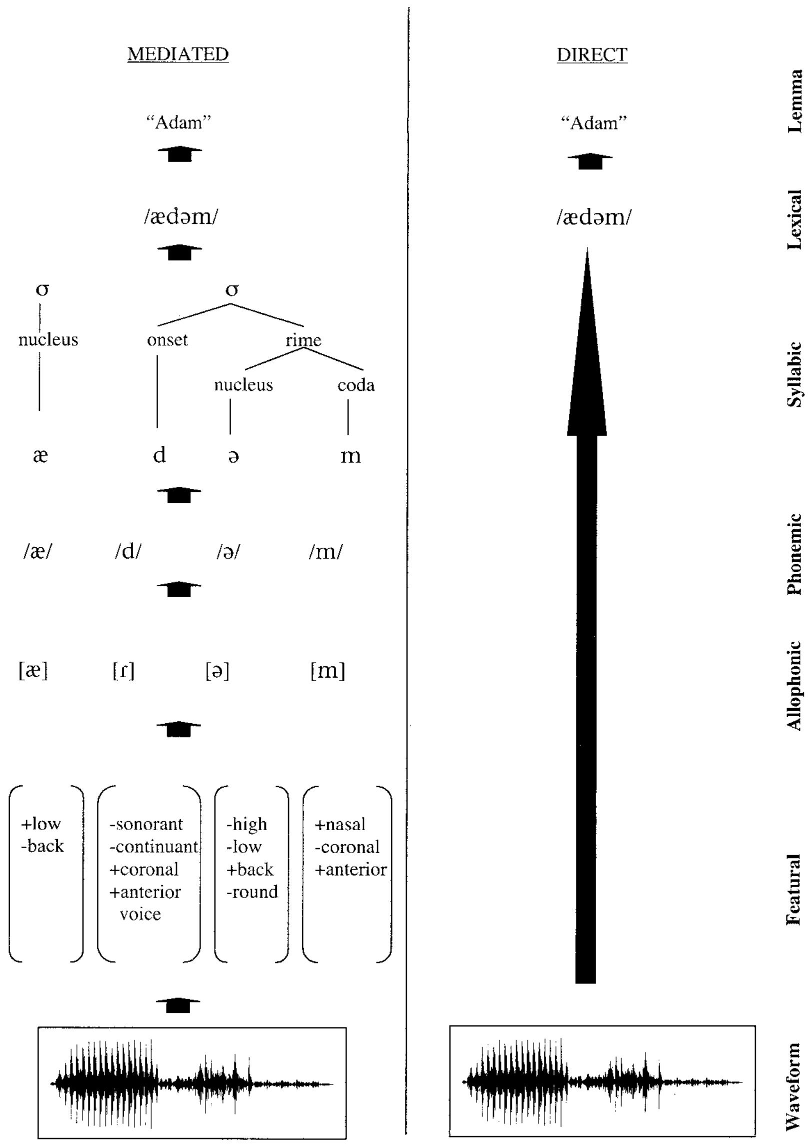

Figure 1. Illustration of mediated and direct access theories.

tial conflicts may arise because some information in the input is consistent with the sublexical unit /d/, whereas other information is consistent with the sublexical unit /b/. Direct access theories predict no processing cost when nonwords are cross spliced with other nonwords because, according to these theories, no intermediate representations exist and therefore no representations are in conflict.

Marslen-Wilson and Warren's results supported the predictions of direct access theories of spoken word recognition: They observed processing costs only when mismatching coarticulatory information involved words. The critical finding was that nonwords cross spliced with other nonwords failed to exhibit processing costs associated with subcategorical mismatch. Marslen-Wilson and Warren concluded that the failure to find effects of subcategorical mismatch for nonwords is due to the absence of intermediate representations.

Recently, McQueen, Norris, and Cutler (1999) challenged Marslen-Wilson and Warren's finding. They found that the crucial distinction between words cross spliced with other words and nonwords cross spliced with other nonwords could be made to come and go as a function of task demands. Moreover, they found that models with a phonemic level of representation could simulate the data pattern obtained by Marslen-Wilson and Warren, thus calling into question the claim that mediated models should always show effects of conflicting information at a sublexical level. Nonetheless, a lack of positive evidence for sublexical representations persists. As a result, the debate between mediated and direct access theories remains unresolved.

We examined the status of intermediate representations in more detail by exploring the perceptual consequences of allophonic variation. More specifically, we examined flapping in American English. A flap (/ɾ/) is a neutralized version and allophone of intervocalic /t/ and /d/. In casual American English, when a /t/ or a /d/ is produced between two vowels, as in *greater* or *Adam*, it is often realized as a flap, a segment that is neither exactly a /t/ nor exactly a /d/ (see Patterson & Connine, 2001). We attempted to determine whether flaps map onto their underlying, abstract phonemic counterparts, /t/ and /d/. Mediated access theories predict that allophonic variation occurring on the surface should map onto more abstract, underlying phonological representations (see, e.g., Pisoni & Luce, 1987). However, according to direct access theories, allophonic variation occurring on the surface should map directly onto lexical representations. Therefore, examining the perceptual consequences of allophonic variation may help to distinguish between these competing theories.

The current study examined the representational status of flaps in memory using a repetition priming paradigm. In this paradigm, participants are presented with a block of spoken words to which they must respond (the study phase). After this initial exposure, participants are presented with another block of words (the test phase). In the second block, some of the words from the first block are repeated. Typically, repeated words are responded to more quickly and accurately than new words (Church & Schacter, 1994; Goldinger, 1996; Kempley & Morton, 1982; Luce & Lyons, 1998). This *repetition priming effect* presumably arises because repeated activation of form-based representation facilitates processing.

Repetition priming can be used to determine whether two nominally different stimuli activate the same mental representation. In particular, the priming paradigm may be used to determine whether flapped segments are mapped onto underlying intermediate form-based representations of /t/s, /d/s, or both, or whether flaps are represented veridically as they appear in casual speech as /ɾ/. Specifically, is there a recoding of the surface allophonic representation, /ɾ/, to the underlying phonological representation, /t/ or /d/, as predicted by mediated access theories of spoken word recognition?

In the present experiments, two blocks of stimuli containing carefully and casually articulated versions of words (and nonwords) were presented. Casually articulated (hypoarticulated) words are produced in a relaxed manner, whereas carefully articulated words are more clearly articulated. Intervocalic /t/s and /d/s are flapped in casually articulated words but not in carefully articulated words. We hypothesize that priming of casually articulated stimuli by carefully articulated stimuli (or vice versa) indicates the presence of a mediating underlying representation in memory. We refer to any significant attenuation in priming for stimuli that mismatch in articulation style as evidence for *specificity*. The presence of specificity effects indicates the absence of intermediate representations, consistent with direct access theories. Conversely, lack of specificity effects indicates the presence of intermediate representations, consistent with mediated access theories.

To review, traditional information-processing theories assume that access to the lexicon is mediated by intervening representations. Direct access theories assume that, after initial sensory registration, access to the lexicon is direct. These classes of theories make opposite predictions regarding the perceptual consequences of allophonic variation. To evaluate these theories, we conducted a series of long-term repetition priming experiments in which flapped and carefully articulated words served as both primes and targets. The basic logic of all of the experiments is that if flapped words (e.g., [æɾəm] activate underlying phonemic representations, they should prime—and be primed by—carefully articulated words (e.g., [ædəm], [ætəm]). On the other hand, if flapped words are not mapped onto underlying phonemic representations, they should fail to prime their carefully articulated counterparts. This latter outcome would parallel the results of many previous studies (e.g., Church & Schacter, 1994; Goldinger, 1996) in which changes in surface properties of words caused a marked attenuation of long-term priming (dubbed *specificity*).

Experiment 1

Method

Participants. Twenty-four participants were recruited from the University at Buffalo community. They were paid $5 or received partial credit for a course requirement. Participants were right-handed native speakers of American English, with no reported history of speech or hearing disorders.

Materials. The stimuli consisted of bisyllabic spoken words containing alveolar and non-alveolar medial consonants. The *alveolar* stimuli consisted of 12 sets of spoken words. Each set contained three stimuli: a minimal pair of carefully produced words that differed only on the voicing of the medial alveolar stop (e.g., [ædəm] and [ætəm] and a casually produced flapped version of the minimal pair (e.g., [æɾəm]). The *non-alveolar* stimuli consisted of 12 words containing a medial [b], [p], [g], or [k] (e.g., *bacon*). The non-alveolar stimuli were also casually and carefully produced. Casually and carefully produced stimuli differed primarily in speed of articulation. In addition, casual alveolar stimuli were produced with a flap. Note that flapped stimuli may be ambiguous (i.e., [æɾəm] may refer to *atom* or *Adam*), whereas casually produced non-alveolar stimuli are not. A complete list of the stimuli used in all experiments is presented in the Appendix.

The final 12 sets of alveolar stimuli were chosen from 24 sets of carefully and casually articulated words containing intervocalic /t/s and /d/s. As a means of ensuring that the casually articulated alveolar stimuli contained fully ambiguous flapped segments (and not clear /t/s or /d/s), the 72 stimuli composing the 24 sets in the original master list were randomized and presented to 10 listeners in a forced-choice (/t/–/d/) identification task. Twelve flapped stimuli were then chosen that were not identified consistently as containing a /t/ or /d/ by more than 6 listeners. On average, half of the participants identified the 12 flapped stimuli as containing /d/ and half as /t/. Thus, flapped words were perceived to be ambiguous. For the carefully articulated stimuli, 9 or more listeners identified the stimuli as containing the intended segment.

The /t/ and /d/ members of the stimulus pairs were matched on average log frequency of occurrence (Kučera & Francis, 1967). The mean log frequencies for /t/ and /d/ words were .53 and .30, respectively. This difference was not significant, $t(11) = 1.67$, $p = .12$. The mean durations for /d/ and /t/ carefully articulated words were 529 ms and 515 ms, respectively. This difference was not significant, $t(11) = 0.53$, $p = .61$. The mean duration for the flapped stimuli was 387 ms. The difference in duration between the casual (flapped) and careful stimuli reflects articulation style; no attempt was made to equate the durations of the flapped and careful stimuli.

The stimuli were recorded in a sound-attenuated room by a phonetically sophisticated male speaker of a midwestern dialect, low-pass filtered at 10 kHz, and digitized at a sampling rate of 20 kHz with a 16-bit analog-to-digital converter. All words were edited into individual files and stored on computer disk.

Design. Two blocks of stimuli were presented. The first constituted the *primes* and the second the *targets*. The carefully and casually produced alveolar and non-alveolar stimuli served as both primes and targets. For both the primes and targets, half of the alveolar and non-alveolar stimuli were casually articulated and half were carefully articulated. Primes *matched*, *mismatched*, or were *unrelated* to the targets. Matching primes and targets were identical (e.g., [ædəm] and [ædəm]). Mismatching primes and targets differed in articulation style only (e.g., [ædəm] and [æɾəm]). The prime block consisted of 8 alveolar, 8 non-alveolar, and 8 unrelated (i.e., control) stimuli. The target block consisted of 12 alveolar stimuli and 12 non-alveolar stimuli. In the target block, 8 stimuli were matching, 8 were mismatching, and 8 were control.

Orthogonal combination of the three levels of prime type (match, mismatch, and control) and two levels of target type (casual and careful) resulted in six conditions, shown in Table 1. Across participants, each careful and casual item was present in every possible condition. However, no single participant heard more than one version of a given word within a block. For example, if a participant heard the word [ædəm] in one of the blocks, he or she did not hear [ædəm], [ætəm], or [æɾəm] again in the same block.

Procedure. Participants were tested individually in a quiet room and were not told at the beginning of the experiment that there would be two blocks of trials. Participants performed a single-word shadowing task in which they attempted to repeat (or shadow) the stimulus word as quickly and accurately as possible. In both the prime and target blocks, the stimuli were presented binaurally over headphones. The headphones had an attached microphone that was placed approximately 1 in. (2.5 cm) from the participant's lips. A Centris 650 computer controlled stimulus presentation and recorded shadowing times. Stimulus presentation within each block was random for each participant.

A given trial proceeded as follows. A light at the top of the response box was illuminated to indicate the beginning of the trial. The participant was then presented with a stimulus word binaurally over the headphones. The participant was instructed to shadow the stimulus word as quickly and accurately as possible. Reaction times (RTs) were measured from onset of the presentation of the stimulus word to onset of the participant's shadowing response. After the participant responded, the next trial was initiated. If the maximum RT (5 s) expired, the computer automatically recorded an incorrect response and presented the next trial.

Results

RTs less than 200 ms or greater than 2,000 ms were replaced with the appropriate condition mean. Less than 1% of the RTs were replaced. Any participant whose overall mean RT fell two standard deviations beyond the grand mean was excluded, resulting in the elimination of 1 participant.

Prime Type (match, mismatch, or control) × Target Type (careful or casual) participant (F_1) and item (F_2) analyses of variance (ANOVAs) were performed on RTs for correct responses and percentages correct for the alveolar and non-alveolar target stimuli.[1] Effects are significant at the .05 level unless otherwise indicated. Accuracy was greater than 97% and produced no significant effects.

Table 1
Experimental Conditions and Examples

Condition	Example: Block 1: prime	Example: Block 2: target
Match		
Careful prime → careful target	ætəm	ætəm
Casual prime → casual target	æɾəm	æɾəm
Mismatch		
Casual prime → careful target	æɾəm	ætəm
Careful prime → casual target	ætəm	æɾəm
Control		
Unrelated prime → careful target	pepɚ	ætəm
Unrelated prime → casual target	pepɚ	æɾəm

Alveolar stimuli. RTs for the alveolar stimuli as a function of prime and target type are plotted in the upper left panel of Figure 2. Mean RTs as a function of condition and magnitudes of specificity and priming for all six experiments are shown in Table 2. Magnitude of specificity is indicated by the difference in RT between the matching and mismatching conditions. Magnitude of priming is indicated by the difference in RT between the matching and control conditions.

Casually articulated (i.e., flapped) items were responded to more quickly than carefully articulated items, $F_1(1, 22) = 56.08$, $MSE = 5{,}934.54$, $F_2(1, 11) = 24.53$, $MSE = 7{,}004.35$, presumably because of the differences in duration. There was also a significant effect of prime type, $F_1(2, 44) = 3.72$, $MSE = 6{,}402.81$, $F_2(2, 22) = 3.26$, $MSE = 5{,}532.03$. Prime type and target type did not interact, F_1 and $F_2 < 1$.

Planned comparisons based on the main effect of prime type revealed significant differences between match and control conditions and between mismatch and control conditions, $F_1(1, 44) = 6.34$, $F_2(1, 22) = 5.47$ and $F_1(1, 44) = 4.70$, $F_2(1, 22) = 4.23$, respectively. There was no difference between match and mismatch conditions, F_1 and $F_2 < 1$.

Both matching and mismatching prime types produced significant facilitative effects on shadowing times. Furthermore, matching primes facilitated target shadowing as much as mismatching primes. These results are consistent with theories that posit underlying intermediate representations.

Non-alveolar stimuli. RTs for the non-alveolar stimuli as a function of prime and target type are plotted in the upper right panel of Figure 2. Magnitudes of specificity and priming are shown in Table 2. Casually articulated items were responded to more quickly than carefully articulated items, $F_1(1, 22) = 24.32$, $MSE = 9{,}764.20$, $F_2(1, 11) = 11.27$, $MSE = 10{,}663.38$. There was also a significant effect of prime type, $F_1(2, 44) = 3.78$, $MSE = 7{,}246.919$, $F_2(2, 22) = 4.83$, $MSE = 3{,}696.95$. Prime type and target type did not interact, $F_1(2, 44) = 1.13$, $MSE = 8{,}858.18$, $F_2 < 1$.

[1] In this and all subsequent experiments, Prime Type × Target Type × Voicing (/t/ vs. /d/) analyses were first performed. In no instance did voicing enter into any significant interactions. Thus, in all analyses, we collapsed across /t/ and /d/ stimuli. In addition, when appropriate (i.e., Experiments 1, 2, 3, and 4), analyses were performed that compared relative changes in RTs from prime to target block. These analyses were consistent with all analyses on the targets alone in this and subsequent experiments. Finally, for a number of reasons, item analyses may not be appropriate for the current experiments. First, the stimuli used exhaust the (small) universe of items that meet our specific criteria, making the need for generalization beyond the present set of stimuli unnecessary. Second, the stimuli are matched on all variables known to affect the dependent variables under scrutiny, thus calling into question the suitability of performing traditional ANOVAs with items as random factors (see Raaijmakers, Schrijnemakers, & Gremmen, 1999). Finally, the low number of items meeting our stringent criteria unavoidably reduces the statistical power of our tests. Despite these caveats, we nonetheless report item analyses, more because of convention than because of their appropriateness. Readers should bear in mind these caveats in interpreting the significance levels of all item tests reported for the current studies.

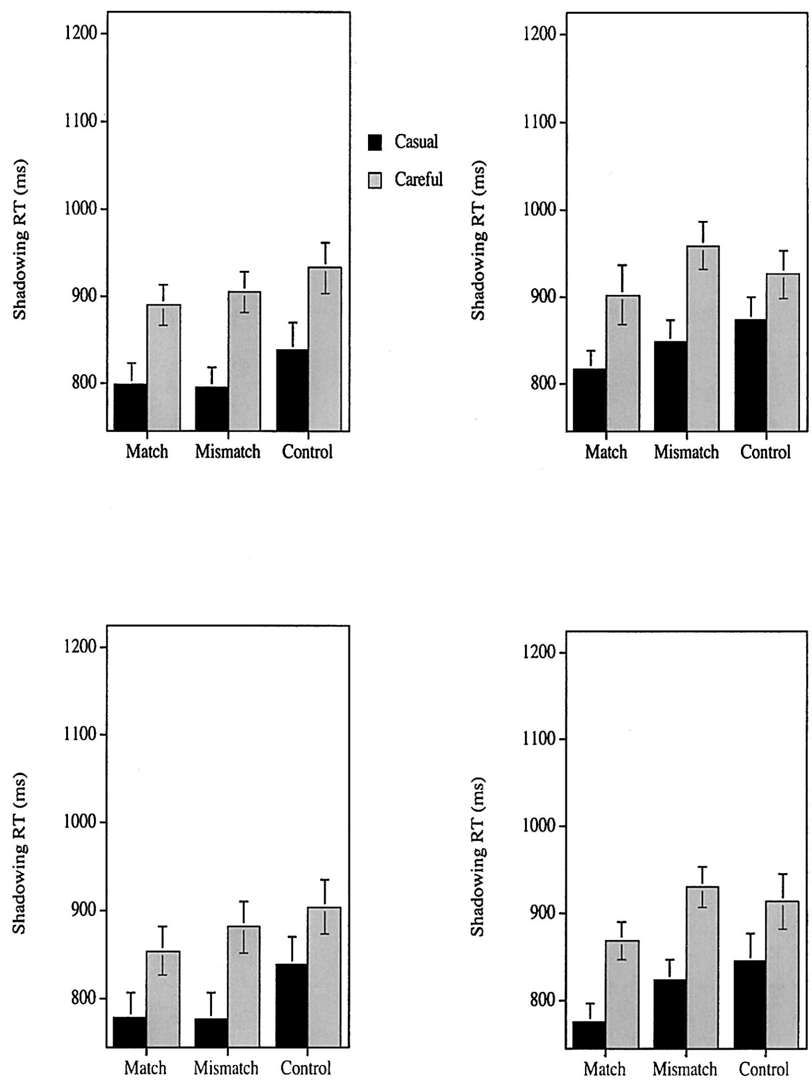

Figure 2. Top: Mean reaction times (RTs) for the alveolar (left) and non-alveolar (right) stimuli in Experiment 1. Bottom: Mean RTs for the alveolar (left) and non-alveolar (right) stimuli in Experiment 2.

Planned comparisons based on the main effect of prime type revealed significant differences between match and control conditions and between match and mismatch conditions, $F_1(1, 44) = 5.25$, $F_2(1, 22) = 8.22$ and $F_1(1, 44) = 6.08$, $F_2(1, 22) = 6.10$, respectively. There was no difference between mismatch and control conditions, F_1 and $F_2 < 1$.

For the non-alveolar stimuli, which did not contain flaps, facilitative priming was observed only when production style (careful and casual) matched. These results are consistent with theories that posit distinct surface representations but contrast with the results obtained for the alveolar stimuli.

Discussion

Experiment 1 revealed two notable findings. First, the shadowing times for the alveolar stimuli demonstrated that casually and carefully articulated words are equally effective primes for both casually and carefully articulated targets. Second, the shadowing times for the non-alveolar stimuli demonstrated that words matching in articulation style were more effective primes for casually and carefully articulated non-alveolar targets than mismatching words.

These results suggest that underlying intermediate representations are activated during processing of phonologically ambiguous flapped stimuli. However, in the absence of ambiguity, surface representations appear to suffice, as evidenced by the non-alveolar stimuli. Thus, the present data provide evidence for the existence of both surface and underlying lexical representations in memory. As a result, these findings join a growing body of evidence in support of lexical representations that preserve surface information (e.g., Church & Schacter, 1994; Goldinger, 1996) while calling into question a purely instance- or exemplar-based model of the mental lexicon (e.g., Goldinger, 1998).

Another possible explanation for the lack of specificity observed for the alveolar stimuli is that lemmas (i.e., semantic–syntactic

Table 2
Reaction Times, Standard Errors, and Magnitudes of Specificity and Priming for Experiments 1–6

		Reaction time (ms)							
		Match		Mismatch		Control			
Experiment	Stimuli	*M*	*SE*	*M*	*SE*	*M*	*SE*	MOS	MOP
1. Shadowing	Alveolar	844	18	850	18	886	22	−6	−42*†
	Non-alveolar	859	21	903	20	900	19	−44*†	−41*†
2. Shadowing	Alveolar	817	20	829	20	871	22	−12	−54*††
	Non-alveolar	822	21	877	20	879	21	−55*†	−57*†
3. EDLD	Alveolar	819	24	934	34	914	26	−115*†	−95*†
4. HDLD	Alveolar	953	25	997	29	1071	33	−44	−118*†
5. EDLD–shadowing	Alveolar	874	15	861	15	904	17	13	−30*††
6. Shadowing–EDLD	Alveolar	873	17	870	21	929	24	3	−56*††

Note. MOS = magnitude of specificity (match − mismatch); MOP = magnitude of priming (match − control). EDLD = easy-discrimination lexical decision; HDLD = hard-discrimination lexical decision.
* $p < .05$, by participants. † $p < .05$, by items. †† $.07 > p > .05$, by items.

representations), and not intermediate form-based representations, may have mediated the priming effect. For example, the ambiguous flapped stimulus [æɾəm] may have activated the lemmas for both *atom* and *Adam*, which in turn may have facilitated processing of the lemmas as targets. If this is the case, there is no need to posit activation of underlying intermediate *form-based* representations corresponding to /t/ or /d/: The facilitative effect of prime on target may have emanated exclusively from the semantic level. The data for the non-alveolar stimuli contradict this hypothesis. Clearly, if the long-term repetition priming effect is lemma based, mismatches in articulation style should have no effect on the magnitude of facilitative priming. However, the results for the non-alveolar stimuli revealed facilitation only when stimuli matched. For example, casually articulated *bacon* failed to prime carefully articulated *bacon*. If the priming effect were lemma based, we would have expected little or no diminution of priming as a function of differences in articulation style.

The literature is also replete with demonstrations that long-term repetition priming is primarily form based. For example, long-term repetition priming is typically modality specific (e.g., Jackson & Morton, 1984). Were the effect lemma based, changes in modality should have no effect on facilitative priming. Moreover, nonwords—which, by definition, have no semantic representations—show long-term repetition priming (e.g., Fisher, Hunt, Chambers, & Church, 2001; Goldinger, 1998). Finally, whether participants' attention is focused on the sound or the meanings of words does not appear to affect long-term repetition priming (e.g., Church & Schacter, 1994). In short, given previous findings—as well as our own results for the non-alveolar stimuli—we can be confident that the locus of the effect is at the form level.

We observed a small numerical trend toward specificity among the carefully articulated alveolar stimuli (match: 890 ms; mismatch: 905 ms; see upper left panel of Figure 2). Therefore, it is possible that the lack of interaction between prime and target type is due to a lack of power. In other words, low power might have been at least partially responsible for the lack of a specificity effect among the carefully articulated alveolar items in Experiment 1. If this is the case, ruling out the possibility that surface representations also dominate processing for carefully articulated alveolar stimuli may be premature. Given the potentially important theoretical implications of these findings, we attempted to replicate Experiment 1. Moreover, combining the data from Experiments 1 and 2 should increase the power of the statistical tests for detecting what may be a weak effect of specificity.

Experiment 2

Method

Participants. A different group of 24 participants were recruited from the University at Buffalo community. They were paid $5 or received partial credit for a course requirement. Participants met the same criteria as those in Experiment 1.

Materials and procedure. The materials and procedure were identical to those used in Experiment 1.

Results

Less than 3% of the RTs and no participants were excluded from the analyses. Accuracy was greater than 90% and produced no significant outcomes.

Alveolar stimuli. RTs for the alveolar stimuli as a function of prime and target type are plotted in the lower left panel of Figure 2. Magnitudes of specificity and priming are shown in Table 2. Casually articulated (i.e., flapped) items were again responded to more quickly than carefully articulated items, $F_1(1, 23) = 47.27$, $MSE = 5{,}053.58$, $F_2(1, 11) = 12.43$, $MSE = 10{,}578.86$. And again, we obtained a main effect of prime type, $F_1(2, 46) = 3.31$, $MSE = 11{,}964.97$, $F_2(2, 22) = 2.14$, $MSE = 8{,}412.95$, $p = .14$. Most important, prime type and target type did not interact, F_1 and $F_2 < 1$.

Planned comparisons revealed a significant difference between the match and control conditions, $F_1(1, 46) = 6.02$, $F_2(1, 22) = 3.96$, $p = .059$; the difference between the mismatch and control conditions was marginally significant by participants but not by items, $F_1(1, 46) = 3.59$, $p = .064$, $F_2(1, 22) = 2.19$, $p = .153$. However, the difference between the match and mismatch conditions was not significant, F_1 and $F_2 < 1$. Aside from the somewhat weaker statistical outcomes, the present results replicate those obtained in Experiment 1.

Non-alveolar stimuli. RTs for the non-alveolar stimuli as a function of prime and target type are plotted in the lower right panel of Figure 2. Magnitudes of specificity and priming are shown in Table 2. The results for the non-alveolar stimuli also replicated Experiment 1: Casually articulated items were again responded to more quickly than carefully articulated items, $F_1(1, 23) = 23.56$, $MSE = 12{,}177.54$, $F_2(1, 11) = 17.69$, $MSE = 8{,}133.70$. There was also a significant effect of prime type, $F_1(2, 46) = 8.89$, $MSE = 5{,}675.61$, $F_2(2, 22) = 8.36$, $MSE = 3{,}391.67$. Prime type and target type did not interact, F_1 and $F_2 < 1$.

Planned comparisons revealed significant differences between match and control conditions and between match and mismatch conditions, $F_1(1, 46) = 13.82$, $F_2(1, 22) = 15.84$ and $F_1(1, 46) = 12.84$, $F_2(1, 22) = 7.83$, respectively. There was no difference between mismatch and control conditions, $F_1 < 1$ and $F_2(1, 22) = 1.40$.

Combined Analyses for Experiments 1 and 2: Alveolar Stimuli

Not surprisingly, the difference between the casually articulated (i.e., flapped) items and the carefully articulated items was significant, $F_1(1, 46) = 103.56$, $MSE = 566{,}738.17$, $F_2(1, 23) = 35.66$, $MSE = 8{,}467.82$. The combined analyses also revealed a significant main effect of prime type, $F_1(2, 92) = 6.89$, $MSE = 9{,}065.53$, $F_2(2, 46) = 5.37$, $MSE = 6{,}676.88$. Crucially, despite the increased power obtained by combining the analyses from Experiments 1 and 2, the interaction between prime type and target type failed to reach significance, F_1 and $F_2 < 1$. These results suggest that the findings of Experiments 1 and 2 are not simply due to lack of statistical power.

Planned comparisons based on the main effect of prime type revealed significant differences between the match and control conditions and between the mismatch and control conditions, $F_1(1, 92) = 12.21$, $F_2(1, 46) = 9.51$ and $F_1(1, 92) = 8.01$, $F_2(1, 46) = 6.24$, respectively. The difference between the match and mismatch conditions for the alveolar stimuli once again failed to reach significance, F_1 and $F_2 < 1$.

Discussion

The data for both Experiments 1 and 2 revealed no statistically significant effects of specificity (articulation style) for alveolar stimuli, whereas differences in articulation style for non-alveolar stimuli completely blocked facilitative priming. In both experiments, we obtained evidence that a flap activates its underlying representations. Specifically, presentation of a flapped item facilitated processing of items containing either /t/ or /d/, and presentation of carefully articulated items containing /t/ or /d/ facilitated processing of flapped stimuli. The priming of flaps by carefully produced items, and vice versa, indicates that shared underlying representations are activated during processing. On the other hand, the pattern of results for the non-alveolar stimuli was markedly different: Primes facilitated their corresponding targets only when articulation style matched. In contrast to the results for the alveolar stimuli, the data for the non-alveolar items indicate that highly specific surface representations are responsible for long-term repetition priming.

Note that casual articulation of the alveolar, but not the non-alveolar, stimuli produces phonological (and lexical) ambiguity. For example, casual production of the word *atom* [æɾəm] (an alveolar stimulus item) is ambiguous between [ætəm] and [ædəm]; however, casual production of the word *bacon* (a non-alveolar stimulus item) is unambiguous. In other words, flaps map onto two possible underlying phonological (and lexical) representations, whereas casual productions of non-alveolar stimuli have only one corresponding representation. This distinction is presumably what led to the pronounced difference between the two sets of stimuli.

The finding that alveolar items activate underlying representations, whereas non-alveolar stimuli appear to contact only highly specific surface representations, can be accounted for within a resonance framework similar to that proposed by Vitevitch and Luce (1999) and based on Grossberg's ARTPHONE model (Grossberg, Boardman, & Cohen, 1997). According to this framework (illustrated in Figure 3), acoustic–phonetic input activates chunks corresponding to sublexical and lexical representations (only lexical representations are illustrated). (A chunk can be

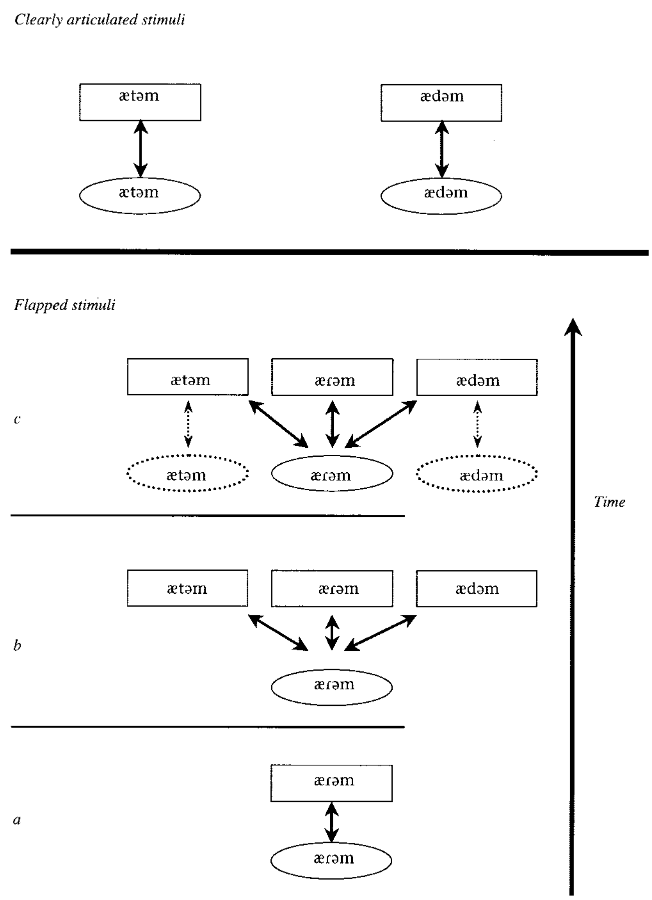

Figure 3. Illustration of the proposed resonances between input and chunks. Ovals correspond to input, rectangles correspond to chunks, and double-sided arrows correspond to resonances (i.e., percepts). For simplicity, only lexical chunks are shown.

thought of as a learned set of associated features that may vary in size from allophone to word.) These chunks then resonate with the input, with the resonance between input and chunk constituting the percept (Grossberg, 1986). We propose that input takes the form of specific and relatively veridical surface representations that preserve articulation style (including allophonic variation). These surface representations resonate with chunks that correspond to both allophonic *and* more abstract phonemic representations (see Figure 3b). That is, when confronted with phonologically and lexically ambiguous flapped stimuli, sublexical and lexical chunks corresponding to a flapped representation, /ɾ/, and both underlying /t/ and /d/ resonate with the surface representation. Underlying /t/ and /d/ chunks resonate to flapped segments because of learned associations between flaps and those lemmas that are also associated with form-based representations having fully specified /t/ and /d/ medial stops.

Once activated, the chunks corresponding to underlying /t/ and /d/ will establish resonances with surface representations. In the case of flapped input, no surface representation will match the activated /t/ or /d/ chunks exactly. We propose that in the absence of an exactly corresponding surface form, the activated chunk will itself instantiate a surface representation with which it will resonate most strongly (see Figure 3c). We envision this process to be much like the one proposed by Grossberg and Meyers (2000, p. 738) to account for phoneme restoration:

> In phonemic restoration experiments, broadband noise may be perceived as different phonemes depending on the context. These percepts may be attributed to a process by which active list chunks use their learned top-down expectations to select the noise components that are consistent with the expected formations and suppress those that are not (Grossberg, 1995, 1999).

Thus, the ambiguous flap is analogous to a noise segment and is perceived in the context of resonating list chunks that correspond to underlying /t/ and /d/. Indeed, we propose that activation of the phonemic chunks by the ambiguous flap results in restoration of a surface representation not actually present in the input. As stated by Grossberg and Stone (1986), "top down signal patterns . . . constitute the read out of optimal templates [e.g., phonemic chunks] in response to ambiguous or novel bottom-up signals [e.g., flaps] . . . to form completed composite patterns that are a mixture of actual and expected information" (p. 58). These completed composite patterns serve as the basis for the long-term priming effect.

In the case of the non-alveolar stimuli—which showed evidence of complete specificity in long-term priming—resonances between surface forms and underlying chunks again serve as the percept. However, given the absence of phonological and lexical ambiguity, the underlying chunks simply resonate with the surface forms to which they match and do not require the restoration of forms not present in the input. Hence, the surface forms that mediate the priming for the non-alveolar stimuli preserve their specific characteristics.

To review, we propose that because of phonological and lexical ambiguity, underlying representations (or chunks) activated by flaps restore surface representations that serve as the basis for long-term repetition priming. In the absence of ambiguity (i.e., for our non-alveolar stimuli), underlying representations resonate with surface forms that preserve detail, hence producing marked effects of specificity in priming.

Before proceeding, we should note that the adaptive resonance account also suggests why we observed a numerical (but not statistically significant) data pattern for the carefully produced alveolar stimuli that is somewhat suggestive of specificity (see Figure 2). That is, RTs for these careful target stimuli tended to be somewhat slower in the mismatching than matching conditions, consistent with some degree of specificity in long-term priming. On the basis of the current framework, attenuation in priming for the carefully produced stimuli in the mismatching condition might be expected given that the prime is actually a restored or instantiated representation based on processing of the flapped stimulus, which may serve as a less effective long-term prime.

According to the resonance framework, instantiation or restoration of the surface form by the underlying chunks should require time. Thus, a task that taps into the recognition process before restoration of the underlying form should show strong effects of specificity in long-term repetition priming, given that the underlying representations may not have had sufficient time to establish resonance with a restored surface form. Although the single-word shadowing task typically produces fairly rapid responses, we expect that the need to contact a representation that drives the production response will allow—indeed encourage—the establishment of resonances between underlying and restored surface forms. We should note that our working assumption is that underlying forms will always instantiate surface forms when there is phonological or lexical ambiguity. However, it should be possible to devise a situation in which we tap the recognition process *before* the restoration of the surface form by the underlying representations.

To test this hypothesis, we conducted two auditory lexical decision experiments in which we manipulated the time required to decide whether a spoken item is a word or nonword. In Experiment 3, we made the word–nonword discrimination task easy by including very un-wordlike nonwords (e.g., *thushthudge*). When presented with nonwords whose sound patterns bear relatively little resemblance to real words, participants in the lexical decision task should be able to base their decisions on overall lexical activity in the system, rather than a near exhaustive analysis of the stimulus itself (Luce & Pisoni, 1998; see also Coltheart, Davelaar, Jonasson, & Besner, 1977). If the sound patterns of the nonwords are quite dissimilar to those of words, they should produce little lexical activity. Thus, only a modicum of lexical activation should signal the presence of a real lexical item, thus allowing for a rapid lexical decision response. In short, easy discrimination in this task should encourage fast processing of the word stimuli.

On the other hand, difficult word–nonword discrimination (Experiment 4) should slow processing. If participants hear very wordlike nonwords (e.g., *bacov,* created from the word *bacon*), processing of the word stimuli should require more than a superficial assessment of lexical activity, given that the nonwords themselves should strongly activate similar lexical items in memory. Note, however, that we expect longer RTs in both Experiments 3 and 4 than in Experiments 1 and 2 because of the additional processing required to make a lexical decision.

By manipulating ease of discrimination, we were able to test the hypothesis that instantiation of surface forms by underlying representations takes time. We predict that in the easy-discrimination lexical decision task, marked effects of specificity should be observed for the alveolar items, because lexical decisions in the target block should be accomplished before instantiation of a surface form corresponding to the underlying representation. However, when discrimination is difficult, we predict that effects of underlying representations should once again be detectable in long-term priming.

The use of the lexical decision task also allowed us to determine the degree to which the activation of underlying representations by flapped items is dependent on the shadowing task itself. It is possible that the underlying representations mediating the priming effects in Experiments 1 and 2 are fundamentally in the service of speech production and have little relevance to perception. Thus, replicating our results in the hard-discrimination lexical decision task would enable us to determine to what extent the speech production mechanism must be involved in the activation of underlying representations.

Experiment 3

Method

Participants. A different group of 24 participants were recruited from the University at Buffalo community. They were paid $5 or received partial credit for a course requirement. Participants met the same criteria as those in Experiment 1.

Materials. The materials were the same as in Experiments 1 and 2 with one exception. To create the lexical decision task, we replaced the non-alveolar stimuli used in Experiments 1 and 2 with low phonotactic probability nonwords (e.g., *thushshug*). However, all of the nonwords used in this experiment were phonotactically legal in English.

Procedure. Except for the task, the procedure was the same as in Experiments 1 and 2. Participants performed a lexical decision task in which they were instructed to decide as quickly and accurately as possible whether the item they heard was a real English word or a nonword. They indicated their decision by pressing one of two appropriately labeled buttons (*word* on the right and *nonword* on the left) on a response box positioned directly in front of them.

Results

RTs less than 500 ms or greater than 2,500 ms were replaced with the appropriate condition mean.[2] Less than 3% of the RTs and 1 participant were excluded from the analyses. Accuracy was greater than 87% and produced only one significant outcome. There was a main effect of prime type, $F_1(2, 44) = 4.07$, $MSE = 271.74$, $F_2(2, 22) = 3.54$, $MSE = 118.17$, that was entirely driven by low accuracy in the control condition. We report data only for the words.

RTs as a function of prime and target type are plotted in the upper left panel of Figure 4. Magnitudes of specificity and priming are shown in Table 2. Casually articulated (i.e., flapped) items were responded to more quickly than carefully articulated items, $F_1(1, 22) = 6.09$, $MSE = 36{,}670.94$, $F_2(1, 11) = 10.49$, $MSE = 10{,}325.40$. There was also a main effect of prime type, $F_1(2, 44) = 6.44$, $MSE = 26{,}620.07$, $F_2(2, 22) = 6.25$, $MSE = 19{,}926.17$. Prime type and target type did not interact, F_1 and $F_2 < 1$.

Planned comparisons based on the main effect of prime type revealed significant differences between the match and control conditions and between the match and mismatch conditions, $F_1(1, 44) = 7.67$, $F_2(1, 22) = 8.56$ and $F_1(1, 44) = 11.30$, $F_2(1, 22) = 10.14$, respectively. However, the difference between the mismatch and control conditions was not significant, F_1 and $F_2 < 1$.

Discussion

Matched primes produced significant facilitative effects on RTs to targets, whereas mismatched primes failed to do so: Facilitative priming was observed only when production style (careful and casual) matched. These results are consistent with surface theories that posit separate representations for casually and carefully articulated stimuli but contrast with the results obtained in Experiments 1 and 2.

As predicted, the lexical decision task with easily discriminated words and nonwords produced specificity effects for the alveolar stimuli, in contrast to Experiments 1 and 2. We hypothesize that the easy-discrimination lexical decision task taps the recognition process before the underlying representations for phonemic /t/ and /d/ have had time to establish resonance with a restored surface form (which, according to our hypothesis, serves as the basis for the long-term repetition priming effect). Hence, repetition effects were observed only for those stimuli matching in articulation style. To garner further evidence for this hypothesis, we conducted another lexical decision experiment in which word–nonword discrimination was made more difficult. We hypothesize that the additional processing required to make the more difficult lexical decision should enable underlying abstract representations sufficient opportunity to establish resonance with the surface form, thus attenuating the specificity effect.

Experiment 4

Method

Participants. A different group of 24 participants were recruited from the University at Buffalo community. They were paid $5 or received partial credit for a course requirement. Participants met the same criteria as those in Experiment 1.

Materials. All materials were the same as in Experiment 3, with one exception: The nonwords were created from the non-alveolar stimuli used in the first two experiments by changing the word endings (e.g., *bacon* → *bacov*), resulting in more wordlike nonwords and presumably more difficult discrimination between words and nonwords.

Procedure. The procedure was the same as in Experiment 3.

Results

RTs less than 500 ms or greater than 2,500 ms were replaced with the appropriate condition mean. Less than 8% of the RTs were excluded from the analyses. In addition, 1 participant was excluded. Accuracy was greater than 81% and produced no significant outcomes. As in Experiment 3, we report data only for the words.

RTs as a function of prime and target type are plotted in the upper right panel of Figure 4. Magnitudes of specificity and

[2] Different upper and lower cutoffs were employed for the two types of tasks (shadowing and lexical decision) because of the overall longer RTs in the lexical decision task.

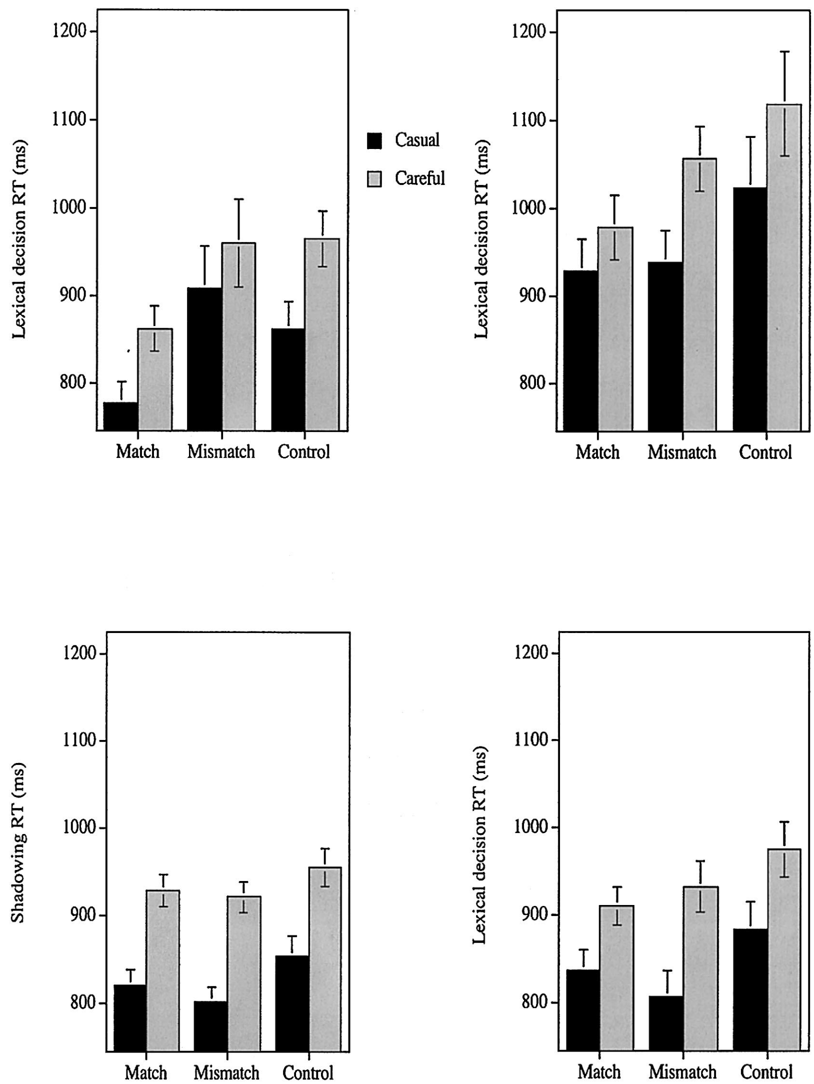

Figure 4. Top: Mean reaction times (RTs) for the stimuli in Experiments 3 (left) and 4 (right). Bottom: Mean RTs for the stimuli in Experiments 5 (left) and 6 (right).

priming are shown in Table 2. Casually articulated (i.e., flapped) items were responded to more quickly than carefully articulated items, $F_1(1, 22) = 8.40$, $MSE = 32{,}019.86$, $F_2(1, 11) = 8.71$, $MSE = 21{,}326.97$. There was also a main effect of prime (significant by participants and marginal by items), $F_1(2, 44) = 6.63$, $MSE = 24{,}764.27$, $F_2(2, 22) = 2.88$, $MSE = 16{,}925.01$, $p = .077$. Prime type and target type did not interact, F_1 and $F_2 < 1$.

Planned comparisons based on the main effect of prime type revealed significant differences between the match and control conditions and between the mismatch and control conditions (although the latter effect was marginal by items), $F_1(1, 44) = 13.01$, $F_2(1, 22) = 5.45$ and $F_1(1, 44) = 5.03$, $F_2(1, 22) = 2.72$, $p = .113$, respectively. However, the difference between the match and mismatch conditions was not significant, $F_1(1, 44) = 1.86$, $F_2 < 1$. Matched and mismatched primes produced significant facilitative effects on target RTs. These results replicated those for the alveolar items in Experiments 1 and 2.

Discussion

As predicted, increasing the difficulty of word–nonword discrimination in the lexical decision task attenuates specificity effects.[3] We propose that more difficult lexical discrimination forces a more exhaustive analysis of the stimulus, resulting in increased opportunities for underlying representations to establish resonance with restored surface representations. Despite a numerical trend

[3] If our manipulation was successful in increasing the difficulty of word–nonword discrimination, RTs to target items should be significantly longer in Experiment 4 (difficult discrimination) than in Experiment 3 (easy discrimination). To confirm the effectiveness of manipulating the wordlikeness of the nonwords, we performed an ANOVA on mean RTs to target items in Experiments 3 ($M = 888.86$ ms) and 4 ($M = 1{,}007.10$ ms). The main effect of experiment was significant, indicating that our manipulation was indeed successful.

toward specificity in Experiment 4, the overall pattern of results supports the hypothesis that increasing the depth of processing should result in stronger resonance of underlying representations with surface forms (see also Goldinger, 1996).[4]

The evidence presented thus far suggests that underlying phonemic representations are contacted during recognition, but only under circumstances in which resonances between underlying and surface forms are encouraged to develop. In particular, we see evidence of underlying representations in long-term repetition priming when a certain degree of depth of processing is required, either by having to generate a production response in the shadowing task or by having to make a difficult word–nonword discrimination in the lexical decision task.[5]

A subtle, but potentially important, alternative hypothesis regarding our depth of processing account deserves consideration. Perhaps underlying forms *always* establish resonances with surface forms, regardless of the circumstances. However, the point at which the response taps the perceptual process may be the determining factor in whether repetition priming effects show effects of specificity. The observed effects of specificity in Experiment 3 may have arisen because lexical decision responses in the second block tapped the recognition process before establishment of resonance between underlying and surface forms, not because the resonances were never established at all.

To further evaluate the hypothesis that underlying representations always resonate with surface forms, despite the fact that such resonances may take time to develop, we conducted two further experiments in which we combined the shadowing and lexical decision tasks. In Experiment 5, participants performed the easy-discrimination lexical decision task in the first block and the shadowing task in the second block. In Experiment 6, the tasks were reversed (shadowing followed by easy-discrimination lexical decision).

Recall that we observed specificity effects in Experiment 3, in which we presented the easy-discrimination lexical decision task in both the first and second blocks. If resonances between underlying and surface forms fail to develop in this task, we should observe only specificity effects in long-term priming, regardless of the task employed in the second block. Very simply, if underlying forms are not contacted in the first block, we would not expect priming for stimuli mismatching in articulation.

However, if underlying forms establish resonances in the first block even in the easy-discrimination lexical decision task, use of the shadowing task in the second block should reveal activation of underlying forms, given that we have already established that the shadowing task affords the opportunity for the underlying representations to resonate with the surface forms. In Experiment 6, we reversed the tasks, presenting the shadowing task followed by the easy-discrimination lexical decision task. In this experiment, we asked whether effects of underlying representations can be found with a task that typically taps into the system before the establishment of the required resonances (the easy-discrimination lexical decision task).

Experiment 5

Method

Participants. A different group of 48 participants were recruited from the University at Buffalo community. They were paid $5 or received partial credit for a course requirement. Participants met the same criteria as those in Experiment 1.

Materials. The materials in Block 1 were the same as those in Experiment 3, and the materials in Block 2 were the same as those in Experiments 1 and 2.

Procedure. The procedure was the same as in the previous experiments with one exception. In the first block participants performed a lexical decision task, and in the second block participants performed a shadowing task.

Results

RTs less than 200 ms or greater than 2,000 ms were replaced with the appropriate condition mean. Less than 2% of the RTs were excluded from the analyses. In addition, 1 participant was excluded. Accuracy was greater than 97% and produced no significant outcomes.

RTs as a function of prime and target type are plotted in the lower left panel of Figure 4. Magnitudes of specificity and priming are shown in Table 2. Casually articulated (i.e., flapped) items were responded to more quickly than carefully articulated items, $F_1(1, 46) = 81.45$, $MSE = 10{,}385.18$, $F_2(1, 23) = 53.82$, $MSE = 8{,}923.44$. There was also a main effect of prime, $F_1(2, 92) = 4.64$, $MSE = 9{,}978.85$, $F_2(2, 46) = 3.61$, $MSE = 7{,}867.89$. Prime type and target type did not interact, F_1 and $F_2 < 1$.

Planned comparisons based on the main effect of prime type revealed significant differences between the match and control conditions and between the mismatch and control conditions, $F_1(1, 92) = 4.31$, $F_2(1, 46) = 3.47$, $p = .069$ and $F_1(1, 92) = 8.82$, $F_2(1, 46) = 6.80$, respectively. However, the difference between the match and mismatch conditions was not significant, F_1 and $F_2 < 1$.

Discussion

Overall, matched and mismatched primes produced significant facilitative effects on target RTs. These results replicated those for the alveolar items in Experiments 1 and 2. Furthermore, these results confirm our earlier assumption that although we are able to tap into the system at a point before the development of resonances between underlying and restored surface forms, processing continues and these resonances are eventually established.

One final question now arises: Can we prime resonances between input and underlying chunks to cause them to develop more

[4] To investigate in more detail the trend toward specificity for the careful items in Experiments 1, 2, and 4, we performed contrasts based on the nonsignificant Prime Type × Target Type interactions. Despite the numerical trends in all three experiments, there was no statistical support for the conclusion that careful items resulted in more specificity than casual items.

[5] Because our argument rests on the hypothesis that depth of processing mediates magnitude of specificity in long-term priming, we conducted additional analyses directly comparing the results from Experiments 1–4. Specifically, we conducted a one-way ANOVA on the magnitude of specificity (MOS; see Table 2) for the alveolar stimuli in Experiments 1–4. The main effect of experiment was significant. As expected, planned contrasts revealed that MOS for Experiment 3 was significantly larger than in Experiments 1, 2, and 4. Moreover, none of the differences in MOS between Experiments 1, 2, and 4 were significant. This analysis confirms that significantly larger specificity effects were obtained only in Experiment 3, as expected.

quickly? In operational terms, can we induce priming of underlying representations in an easy-discrimination lexical decision task by having participants shadow stimuli in the prime block? Despite the fact that in Experiment 3 we found that easy-discrimination lexical decision produces pronounced specificity effects, we predict that, even in this task, resonances between input and underlying chunks can be made to develop more quickly if they have recently been established in a deeper processing task (i.e., shadowing).

Experiment 6

Method

Participants. A different group of 48 participants were recruited from the University at Buffalo community. They were paid $5 or received partial credit for a course requirement. Participants met the same criteria as those in Experiment 1.

Materials. The materials in Block 1 were the same as those in Experiments 1 and 2, and the materials in Block 2 were the same as those in Experiment 3.

Procedure. The procedure was the same as in Experiment 5 except that the order of tasks was reversed. In the first block participants performed a shadowing task, and in the second block participants performed a lexical decision task.

Results

RTs less than 500 ms or greater than 2,500 ms were replaced with the appropriate condition mean. Less than 3% of the RTs were excluded from the analyses. In addition, 1 participant was excluded. Accuracy was greater than 94% and produced no significant outcomes.

RTs as a function of prime and target type are plotted in the lower right panel of Figure 4. Magnitudes of specificity and priming are shown in Table 2. Casually articulated (i.e., flapped) items were responded to more quickly than carefully articulated items, $F_1(1, 46) = 24.82$, $MSE = 26{,}438.17$, $F_2(1, 23) = 20.71$, $MSE = 15{,}106.97$. There was also a main effect of prime by participants, $F_1(2, 92) = 3.85$, $MSE = 27{,}195.83$, although the effect failed to reach significance by items, $F_2(2, 46) = 1.94$, $MSE = 18{,}146.59$, $p = .16$. Prime type and target type did not interact, F_1 and $F_2 < 1$.

Planned comparisons based on the main effect of prime type revealed significant differences by participants between the match and control conditions and between the mismatch and control conditions; the effects by items were statistically somewhat weaker, $F_1(1, 92) = 5.42$, $F_2(1, 46) = 3.78$, $p = .058$ and $F_1(1, 92) = 6.12$, $F_2(1, 46) = 1.53$, $p = .223$, respectively. However, the difference between the match and mismatch conditions was not significant, F_1 and $F_2 < 1$.

Discussion

Overall, matched and mismatched primes produced facilitative effects on target RTs (although statistical support was somewhat weaker than in previous experiments). These results replicated those for the alveolar items in Experiments 1 and 2 but differed from the results of Experiment 3 with the same stimuli and task in Block 2. Indeed, the only difference between Experiments 3 and 6 was the task performed by participants in Block 1. Rather than performing the easy-discrimination lexical decision task in both Blocks 1 and 2, as had been done in Experiment 3, Experiment 6 participants performed the shadowing task in Block 1.

As predicted, preceding a superficial processing task with one that encourages contact with underlying representations results in a lack of specificity effects.[6] Simply put, resonances prime: Contacting an underlying representation makes it easier to establish that same resonance at a slightly later time (see also Grossberg & Meyers, 2000, p. 738).

General Discussion

This investigation began with a simple question: Are flaps mapped onto their underlying phonemic counterparts during perceptual processing? If affirmative, the answer provides evidence against direct access models of recognition, instead supporting, in part, the more traditional mediated models of speech perception and spoken word recognition according to which the recoded speech waveform is mapped onto more abstract, underlying representations.

The six long-term repetition priming experiments reported here provide some evidence for mediated models (broadly construed; see subsequent discussion) while also suggesting the precise circumstances under which underlying representations may be contacted during recognition. In Experiments 1 and 2, in which participants shadowed flapped and carefully produced alveolar stimuli in both the prime and target blocks, flaps primed carefully articulated stimuli and vice versa, a result consistent with the notion that flaps activate their underlying phonemic counterparts. Crucially, however, non-alveolar stimuli produced marked specificity effects, suggesting that the long-term repetition priming effect is not lemma based. Moreover, the finding that non-alveolar stimuli prime only when they match on articulation style, whereas alveolar stimuli need not match to produce facilitative repetition effects, suggests that the phonological and lexical ambiguity inherent in flapped stimuli is a necessary condition for activation of underlying representations.

---

[6] To compare more directly the effects of different study and test tasks (i.e., shadowing vs. lexical decision) on priming, we conducted a series of comparisons across experiments in which we held test task constant while varying study task, and vice versa. For the comparisons in which test task was held constant, RTs in Experiment 5 (lexical decision followed by shadowing) were compared with target RTs in Experiments 1 and 2 (shadowing followed by shadowing). In addition, target RTs in Experiment 6 (shadowing followed by lexical decision) were compared with target RTs in Experiment 3 (lexical decision followed by lexical decision). The Prime Type × Experiment interaction was not significant in the comparison of Experiments 1 and 2 with Experiment 5, indicating an equivalent lack of specificity across these experiments. However, as expected, the Prime Type × Experiment interaction *was* significant in the comparison between Experiments 3 and 6, confirming that shadowing during study attenuates specificity effects when participants make lexical decisions during test. For the analyses in which study task was held constant, the comparison of Experiments 1 and 2 with Experiment 6 resulted in a nonsignificant Prime Type × Experiment interaction, indicating an equivalent lack of specificity across these experiments. However, as expected, the Prime Type × Experiment interaction was significant in the comparison between Experiments 3 and 5, confirming that shadowing attenuates specificity, even when participants are presented with the specificity-inducing easy-lexical-discrimination task during study.

In Experiments 3 and 4, participants made lexical decisions in both the prime and target blocks to the same alveolar stimuli used in Experiments 1 and 2. By manipulating the difficulty of word–nonword discrimination across the two experiments, we tested the hypothesis that "depth" of processing may be crucial in activating underlying representations. In Experiment 3, in which word–nonword discrimination was made easier by the inclusion of *unwordlike* nonwords, there was no evidence that flapped stimuli activated their underlying phonemic counterparts: Flapped words failed to prime carefully articulated words, and vice versa. However, in Experiment 4, in which discrimination was made difficult by the inclusion of nonwords that were very *wordlike*, clear evidence for activation of underlying representations reemerged: Flaps primed carefully articulated words, and vice versa. Experiments 3 and 4 suggest that depth of processing (manipulated through ease of lexical discrimination) mediates the activation of underlying representations.

Finally, in Experiments 5 and 6, we crossed the shadowing and lexical decision tasks, using lexical decision in the prime block and shadowing in the target block in Experiment 5 and the reverse arrangement of tasks in Experiment 6. The particular version of the lexical decision task used in these two experiments was the same one employed in Experiment 3 (i.e., the easy-discrimination task) in which we observed no evidence of activation of underlying representations. Thus, we combined a task that consistently produced activation of underlying forms (i.e., shadowing) with a task that produced no evidence of underlying activation (i.e., easy-discrimination lexical decision).

In Experiment 5, in which the easy-discrimination lexical task occurred in the prime block and the shadowing task occurred in the target block, we obtained evidence for activation of underlying representations. Thus, despite the fact that easy-discrimination lexical decision failed to produce activation of underlying forms in Experiment 3, stimuli in this task still acted as effective primes for underlying representations when the shadowing task was used in the target block. This result suggests that underlying representations are indeed contacted in the easy-discrimination lexical task: hence the priming effect. However, these underlying forms may have little or no effect on processing when a more superficial analysis of the stimulus suffices (as in easy lexical discrimination in the target block of Experiment 3).

In Experiment 6, in which the situation was reversed (shadowing was used in the prime block and easy-discrimination lexical decision in the target block), we again obtained evidence (albeit statistically somewhat weaker) for activation of underlying forms. Note that this latter finding contrasts with Experiment 3, in which easy-discrimination lexical decision in the target block produced only specificity effects. Apparently, once the underlying forms have been activated in the prime block, their effects are sufficiently strong and long-lasting to manifest themselves even in a task that requires only superficial stimulus processing.

Overall, we obtained a data pattern consistent with activation of the underlying phonemic counterparts of flaps during spoken word processing. Indeed, in only two circumstances did we observe specificity effects: (a) for the non-alveolar stimuli in Experiments 1 and 2 and (b) for the alveolar stimuli when the easy-discrimination lexical decision task was used in both prime and target blocks. *Underlying* representations appear to dominate processing when spoken input is phonologically ambiguous (i.e., when flaps are present), when items are processed to a deep level (as in the shadowing task), and when enough time is allowed for the underlying representations to have an effect on recognition (as in the hard-discrimination lexical decision task). Alternatively, *surface* representations appear to dominate processing when spoken input is unambiguous (i.e., when non-alveolar stimuli are used), when items are not processed to a deep level (as in easy-discrimination lexical decision), and when there is insufficient time for the underlying representations to have an effect on recognition (again as in the easy-discrimination lexical decision).

How, then, do we account for the activation of underlying forms and the circumstances under which their effects are manifested in the recognition process? We propose an account of these findings based on Grossberg's ARTPHONE model (Grossberg et al., 1997; see also Vitevitch & Luce, 1999). To review, acoustic–phonetic input composed of relatively veridical surface representations resonates with chunks corresponding to more abstract phonological representations, as well as chunks corresponding to less abstract, allophonic representations. These resonances serve as the basis for long-term repetition priming.

In the absence of ambiguity in the input, the resonances between surface forms and chunks corresponding to underlying representations preserve detail (see Grossberg & Meyers, 2000). However, underlying representations (or chunks) activated by *ambiguous* flaps result in a restoration of surface representations not actually included in the input.

Deep phonological processing associated with shadowing spoken stimuli and making difficult word–nonword discriminations encourages the restoration of surface representations by underlying representations, and instantiation or restoration of the surface form by the underlying chunks requires time. Thus, tasks that tap into the recognition process before restoration of the underlying form show strong effects of specificity in long-term repetition priming, presumably because the underlying representations may not have had sufficient time to establish resonance with a restored surface form. However, the present evidence suggests that even though recognition may occur before establishment of resonances with underlying phonemic representations, these representations are nonetheless contacted, probably obligatorily. Moreover, it may be possible to prime the resonances themselves, such that previous activation of underlying forms makes establishing resonances with these forms easier at a later time.

As an aside, we should note that what we have referred to throughout as *depth of processing* may or may not be coextensive with the *time course* of processing (see also Luce, McLennan, & Charles-Luce, in press). It is not always the case that the fastest responses result in the most specificity, although there is certainly a trend in that direction, especially within a given task. For example, compare the data for two lexical decision experiments in which the faster responses resulted in marked specificity. However, shadowing produced roughly equivalent RTs (in Experiments 1, 2, and 5) to lexical decision (in Experiment 3) for the alveolar stimuli, yet shadowing consistently resulted in diminished specificity effects (relative to Experiment 3). At present, we can only acknowledge that depth of processing may be strongly associated with the time course of processing but may also encompass other variables, such as the need to produce a response.

The present results bear a marked resemblance to recent work reported by Hallé, Chéreau, and Segui (2000), who examined the effects of voice assimilation in French on the perception of underlying phonemic forms. Because of voice assimilation in French, voiced stops followed by voiceless segments, as in words such as /absyrd/, are devoiced, as in [apsyrd]. Hallé et al. examined whether French participants would perceive the first consonant as a /p/, which is actually present in the signal, or as a /b/, which is consistent with both the underlying representation and the orthography. Using a phonemic gating task, they found that /p/ initially dominated participants' responses, with /b/ responses gradually increasing over time and eventually overcoming the initial /p/ responses. In terms of the resonance framework, information present in the signal (in this case /p/) dominates processing until sufficient time has elapsed for the underlying /b/ to instantiate a corresponding surface representation. After the chunk corresponding to the underlying /b/ instantiates the appropriate surface representation (as a result of the learned association between [apsyrd] and underlying /b/), /b/ responses dominate. Although the authors were unable to determine definitively whether their effect had an orthographic or morphophonemic locus, their results are clearly consistent with the framework we are proposing.

The present results may pose some difficulties, although presumably not insurmountable difficulties, to current connectionist models of spoken word recognition. TRACE and Shortlist, for example, both lack an allophonic layer of representation, a minimal requirement dictated by the finding that, under appropriate circumstances, flaps activate their phonemic counterparts. Only PARSYN incorporates an explicit allophonic level. However, PARSYN lacks phonemic representations, which may prove problematic in accounting for the activation of underlying forms (although PARSYN's lexical representations are phonemically coded).

Although in their current forms, TRACE, Shortlist, and PARSYN all may have some difficulties in accounting for the complete set of results, nothing in their architectures prohibits the necessary modifications: TRACE and Shortlist could add an allophonic level, and PARSYN could add a phonemic level. However, even with the appropriate representations, it is unclear how these or similar models could account for the observed effects of depth of processing. One possibility may be to incorporate an attentional focus by manipulating weights at various levels. Overall, however, we believe that the adaptive resonance framework most naturally handles the range of observed effects in the present studies.

We now come full circle to ask how the present results bear on the distinction between mediated and direct access models, the original theoretical focus of our work. Clearly, one aspect of mediated models has been supported, namely the activation of underlying abstract forms in spoken word processing. However, the adaptive resonance framework we have adopted bears a strong resemblance to a direct access model. After all, we have proposed that veridical representations first make contact with fairly specific (i.e., allophonic) representations, only after which do underlying forms come into play. Moreover, the results for the non-alveolar stimuli suggest that fairly specific representations dominate processing in the absence of phonological ambiguity or deeper processing. In short, the adaptive resonance framework requires us to reconceptualize the problem. Indeed, neither of the prototypical models illustrated in Figure 1 can adequately account for the present data, in part because they fail to acknowledge that perception may be better characterized as a resonance between learned expectation and sensory input in which the percept may reside neither in the sensory data nor in the long-term representation but in some mélange of the two.

References

Bloomfield, L. (1933). *Language.* New York: Holt.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English.* New York: Harper & Row.

Church, B. A., & Schacter, D. L. (1994). Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 43–57.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535-555). Hillsdale, NJ: Erlbaum.

Fisher, C., Hunt, C., Chambers, K., & Church, B. (2001). Abstraction and specificity in preschoolers' representations of novel spoken words. *Journal of Memory and Language, 45,* 665–687.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1166–1183.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105,* 251–279.

Grossberg, S. (1986). The adaptive self-organization of serial order in behavior: Speech, language, and motor control. In E. C. Schwab & H. C. Nusbaum (Eds.), *Pattern recognition by humans and machines: Vol. 1. Speech perception* (pp. 187–294). New York: Academic Press.

Grossberg, S. (1995). The attentive brain. *American Scientist, 83,* 438–449.

Grossberg, S. (1999). The link between attention, brain learning, and consciousness. *Consciousness and Cognition, 8,* 1–44.

Grossberg, S., Boardman, I., & Cohen, M. (1997). Neural dynamics of variable-rate speech categorization. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 483–503.

Grossberg, S., & Meyers, C. W. (2000). The resonant dynamics of speech perception: Interword integration and duration-dependent backward effects. *Psychological Review, 107,* 735–767.

Grossberg, S., & Stone, G. O. (1986). Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychological Review, 93,* 46–74.

Hallé, P. A., Chéreau, C., & Segui, J. (2000). Where is the /b/ in "absurde" [apsyrd]? It is in French listeners' minds. *Journal of Memory and Language, 43,* 618–639.

Harris, Z. S. (1955). From phoneme to morpheme. *Language, 31,* 190–222.

Jackson, A., & Morton, J. (1984). Facilitation of auditory word recognition. *Memory & Cognition, 12,* 568–574.

Jusczyk, P., & Luce, P. A. (2002). Speech perception and spoken word recognition: Past and present. *Ear & Hearing, 23,* 1–40.

Kempley, S. T., & Morton, J. (1982). The effects of priming with regularly and irregularly related words in auditory word recognition. *British Journal of Psychology, 73,* 441–454.

Kenstowicz, M., & Kisseberth, C. (1979). *Generative phonology: Description and theory.* New York: Academic Press.

Klatt, D. H. (1989). Review of selected models of speech perception. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 169–226). London: MIT Press.

Kučera, H., & Francis, W. (1967). *Computational analysis of present day American English.* Providence, RI: Brown University Press.

Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition, 38,* 245–294.

Luce, P. A., Goldinger, S. D., Auer, E. T., & Vitevitch, M. S. (2000).

Phonetic priming, neighborhood activation, and PARSYN. *Perception & Psychophysics, 62,* 615–625.
Luce, P. A., & Lyons, E. A. (1998). Specificity of memory representations for spoken words. *Memory & Cognition, 26,* 708–715.
Luce, P. A., McLennan, C. T., & Charles-Luce, J. (in press). Abstractness and specificity in spoken word recognition: Indexical and allophonic variability in long-term repetition priming. In J. Bowers & C. Marsolek (Eds.), *Rethinking implicit memory.* Oxford, England: Oxford University Press.
Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear & Hearing, 19,* 1–36.
Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review, 101,* 653–675.
McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1–86.
McQueen, J. M., Norris, D. A., & Cutler, A. (1999). Lexical influence in phonetic decision making: Evidence from subcategorical mismatches. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1363–1389.
Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52,* 189–234.
Patterson, D., & Connine, C. M. (2001). Variant frequency in flap production: A corpus analysis of variant frequency in American English flap production. *Phonetica, 58,* 254–275.
Pisoni, D. B., & Luce, P. A. (1987). Acoustic-phonetic representations in word recognition. *Cognition, 25,* 1–52.
Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with "the language-as-fixed-effect-fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language, 41,* 416–426.
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.
Studdert-Kennedy, M. (1974). The perception of speech. In T. A. Sebeok (Ed.), *Current trends in linguistics: Vol. XII* (pp. 2349-2385). The Hague: Mouton.
Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40,* 374–408.
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.

Appendix

Stimuli Used in the Present Experiments

Alveolar (E1–E6)	Non-alveolar (E1 and E2)	Nonwords (E3, E5, and E6)	Nonwords (E4)	Unrelated (E1 and E2)	Unrelated (E3–E6)
Adam	bacon	jʌʃðʌʧ	bekəv	luggage	soap
atom	baggage	θʌsjʌʤ	bægənt	jagged	paper
coder	boycott	ʧʌʃθʌʤ	boɪkɔf	nugget	folder
coater	bucket	jʌʧʧʌʤ	bʌkəm	ribbon	globe
grading	bygone	θʌʧʃʌʤ	baɪgəps	rugged	kɪkbæp
grating	bypass	ðaɪðʃaɪð	baɪpæb	topic	mædkʌs
padding	cabbage	ʃaɪðʧaɪð	kæbəv	turban	bɑmʃɛz
patting	cabin	gaɪððaɪz	kɔkəg	weapon	kaɹfæp
paddy	caucus	ðaɪbdʒaɪz	sɜkɚ		
Patty	circuit	ðaɪvʃaɪb	kopɚg		
pedal	circus	ʧaɪzwaɪð	ʤægʌp		
petal	coping	jiʃgiʃ	wɛpʌks		
pudding		ziʃjið			
putting		ziðgið			
raider		ðiθjiθ			
rater		ziθgiθ			
raiding		θeʒθeð			
rating		θeθθeg			
seeding		ʤeʒdʒeð			
seating		ʧeʒʧeð			
tudor		ʤeθʃeʒ			
tutor		θeʤʃeð			
udder		jɜzjɜθ			
utter		ʃɜθjɜg			

Note. E = experiment.