

Representation of Lexical Form: Evidence From Studies of Sublexical Ambiguity

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

The authors examined the role of intermediate, sublexical representations in spoken word perception. In particular, they tested whether flaps, which are neutralized allophones of intervocalic /t/s and /d/s, map onto their underlying phonemic counterparts. In 2 shadowing tasks, the authors found that flaps primed their carefully articulated counterparts, and vice versa. Because none of the flapped stimuli were lexically ambiguous (e.g., between *rater* and *raider*), these results provide evidence that such priming is sublexically mediated. Therefore, the current study provides further insights into when underlying form-based representations are activated during spoken word processing. In particular, the authors argue that phonological ambiguity, inherent in their flapped stimuli, is one of the conditions leading to the activation of underlying representations.

Keywords: lexical, sublexical, phonological ambiguity, allophonic specificity, spoken word recognition

Allophonic variability can, in certain contexts, result in the neutralization of segmental distinctions, which in turn can lead to sublexical and/or lexical ambiguity. For example, in casually produced American English, when a /t/ or a /d/ is produced between stressed and unstressed vowels, as in *water* or *greedy*, it is often realized as a flap, /ɾ/, a segment that is neither exactly a /t/ nor exactly a /d/ (Fox & Terbeek, 1977; see also Charles-Luce, 1997; Patterson & Connine, 2001).¹ The intermediate status of the flap presents a unique opportunity for exploring the nature of the representations that underlie spoken word recognition. In particular, the phenomenon of flapping in American English may provide a means for determining the potential abstractness of phonological representations and may provide a means for evaluating mediated and direct access theories of spoken word recognition, which make opposing claims regarding the levels of representation that may intervene between sensory recoding and lexical representation. *Mediated access theories* posit some form of intermediate representations (e.g., allophones, phonemes, and/or syllables) between

initial recoding and lexical representation. In contrast, *direct access theories* propose that after the initial recoding of sensory data, information is mapped directly onto form-based lexical representations.

In the current study, we attempted to determine whether flaps map onto their underlying, abstract phonemic counterparts, /t/ and /d/, at some point during processing (see also Connine, 2004; McLennan, Luce, & Charles-Luce, 2003). 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). 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.

In both previous and current research, we have used the repetition-priming paradigm, a standard experimental technique used to examine the nature of representations in memory. In this paradigm, stimuli are presented in two blocks. The first block consists of *primes*, and the second consists of *targets*. The depen-

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 a meeting of the International Congress of Phonetic Sciences, Barcelona, Spain, August 2003. 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 Theofanis Pantazis, Melissa Pfenig, and Danielle Richardson for their help preparing stimuli and conducting experiments.

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, NY 14260. E-mail: mclennan@buffalo.edu or luce@buffalo.edu

¹ We use the term *allophone* in keeping with Ladefoged's (2001) definition:

The variants of the phonemes that occur in detailed phonetic transcriptions are known as **allophones**. . . . For example, we know that in most varieties of American English, /t/ has a voiced allophone when it occurs between a stressed and unstressed vowel. (p. 37)

Although it could be argued that a flap is more like a /d/ than like a /t/ (e.g., because it is voiced), we screened all flaps in our current experiments to ensure that they were indeed ambiguous between /t/ and /d/. Moreover, although flapping is common in American English casual speech, it is not our intention to restrict flapping to a casual speech phenomenon. The crucial property of American English flaps for the purposes of the present study is contrast (/t/-/d/) neutralization.

dent variable of interest is a change in the participant's response to the target as a function of prior presentation of the prime. Facilitation of target processing by a related prime indicates that the prime and target share a common representation. Facilitative repetition priming can be used to determine whether two nominally different stimuli activate the same mental representation. Although similar representations may be sufficient to produce some degree of priming, equivalent degrees of priming by exact and nominally different stimuli presumably indicate that the two map onto a common mental representation.

In a previous study, we presented *casually* articulated (flapped) and *carefully* articulated versions of bisyllabic words (e.g., *rater* and *raider*) as primes and targets (McLennan et al., 2003). Participants shadowed, or repeated, the targets as quickly as possible, with shadowing times serving as indices of processing speed. We attempted to determine whether flapped primes facilitate processing of carefully articulated targets and vice versa. Failure to observe equivalent priming between flaps and careful stimuli would indicate that the two do not share a common representation. However, equal facilitation of target processing by flaps and careful primes would provide evidence for a shared, presumably abstract representation.

The results demonstrated that casually articulated flapped words (e.g., *reɪrə*) primed careful words (e.g., *reitə* or *reɪdə*) as much as they primed themselves (and vice versa), providing evidence for shared underlying representations.² Although these results support mediated access theories, in which the shared representation is intermediate between input and lexical form-based representations, they are not entirely inconsistent with direct access theories: Allophonically specific form-based lexical representations might indeed have been accessed directly, with subsequent activation of associated underlying forms, an issue to which we return below.

We also obtained evidence for the activation of more specific form-based representations. Using nonalveolar stimuli (e.g., *bacon*), which did not contain flaps and thus were not ambiguous when articulated carefully or casually, we found that only those stimuli matching in articulation style (e.g., careful *bacon* and careful *bacon*, casual *bacon* and casual *bacon*) primed one another: Mismatches in articulation style (e.g., careful *bacon* and casual *bacon*) attenuated or eliminated long-term repetition priming. On the basis of these data, we argued that one of the circumstances leading to the activation of underlying form-based representations is *sublexical* ambiguity. Note that when casually articulated, the flapped stimuli (*reɪrə*) map onto two possible phonological representations (*reitə*, *reɪdə*), whereas casual productions of the nonalveolar stimuli (*beɪkən*) have only one corresponding representation (*beɪkən*). Thus, we argued that underlying intermediate representations are activated during processing of phonologically ambiguous stimuli.

On the basis of the tenets of adaptive resonance theory (ART; Grossberg, 1986; Grossberg & Myers, 2000), we proposed that acoustic-phonetic input activates and resonates with chunks of well-learned associated features (corresponding, potentially, to segments, syllables, and words) and that this resonance between input and chunk constitutes the percept (see Luce & McLennan, 2005; McLennan, in press). When phonologically ambiguous flapped stimuli serve as input, lexical and/or segmental chunks consistent with both flaps and underlying /t/ and /d/ are activated,

on the basis of (a) similarity between chunk and input and, crucially, (b) expectations based on prior experience that a flap may be consistent with either a /t/ or a /d/ (or consistent with lexical items containing a /t/ or /d/). Note that those chunks corresponding to /t/ and /d/ (or lexical items containing /t/ and /d/) do not match the input directly. If resonances are to be established between these underlying chunks and the input, the resonances must restore, or fill in, information not actually presented. Thus, we proposed that chunks corresponding to underlying representations that are not actually present in the input resonate in such a way as to create a percept that combines both top-down and bottom-up information. This composite percept then serves as the basis for long-term priming. Again, however, this account predicts such restoration only in the case in which two possible lexical or sublexical interpretations have been regularly associated with the input, as in the case with flaps, which may map onto either /t/ or /d/.

Within this framework, then, the crucial question regarding mediated and direct access theories concerns the nature of the chunks: Are they solely lexical, consistent with the direct access theories, or do they represent various grains of analysis, consistent with both ART and mediated access theories? To evaluate this question, we use the same design as in our earlier work. However, rather than using lexically ambiguous stimuli, we use (a) alveolar stimuli without lexical pairs, such as *pretty* and *greedy*, and (b) alveolar nonwords, such as *oytem* and *pawding*.

The use of words without lexical pairs avoids lexical ambiguity while maintaining sublexical ambiguity. According to our implementation of the adaptive resonance framework, if lexical chunks are solely responsible for activation of underlying forms in the presence of flapped input, there should be no learned expectation that a flap in a lexical item such as *pretty* should also map onto the nonword *preddy* and, hence, no restoration of distinct phonetic forms corresponding to the underlying /t/ and /d/. The flapped and carefully articulated versions of a word such as *pretty* should be treated simply as surface variants with no lexical consequences and, like the careful and casual nonalveolar *bacon* in our previous research, should be treated as distinct phonetic forms that do not serve to effectively prime one another. Conversely, if the locus of the restoration effect is at the level of the sublexical chunk, repeated experience with flapped segments should lead to the expectation that the segments themselves regularly map onto underlying /t/ and /d/, in which case lexical status may be irrelevant.

The use of nonwords also avoids lexical ambiguity and provides an even more stringent test of the sublexical ambiguity hypothesis. In the case of nonword stimuli, no lexical chunks should be strongly activated. Thus, priming between flaps and carefully articulated stimuli should reflect the operation of similarity relations and expectations at a sublexical level. Conversely, failure to observe priming with nonword stimuli would suggest that lexical representations themselves must be strongly activated to produce the kind of restoration indicative of priming among flaps, /t/s, and /d/s.

² However, we have also obtained evidence for the more specific flap representation under other experimental conditions (see McLennan et al., 2003; see also Connine, 2004).

Experiment 1: Words Without Lexical Pairs

Method

Participants. Seventy-two members of the University at Buffalo, The State University of New York, community participated in the experiment. 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. Twelve minimal pairs of bisyllabic English stimuli containing intervocalic /t/ and /d/ were selected in which one member of the pair was a word (e.g., *pretty*, *greedy*) and one was a nonword (e.g., *preddy*, *greety*). Each pair was recorded twice by a phonetically sophisticated male speaker (i.e., this individual has had extensive phonetic training), once in a carefully articulated manner that preserved the /t/-/d/ contrast and once in a casually articulated manner in which the /t/-/d/ contrast was neutralized to produce a flap.

The mean duration for the carefully articulated stimuli was 585 ms. The mean duration for the flapped stimuli was 369 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, low-pass filtered at 10 kHz, and digitized at a sampling rate of 20 kHz using a 16-bit analog-to-digital converter. All words were edited and stored on computer disk.

The stimuli were screened by 10 listeners who indicated whether each word contained a /t/ or /d/. To avoid any bias in favor of a real word, we altered the stimuli or truncated them to create nonwords (e.g., *greedy* was truncated so that it became /idi/). For the carefully articulated stimuli, at least 9 listeners correctly identified the intended segment. For the flaps, no more than 6 participants labeled any of the flapped items as containing a /t/ or /d/. In other words, for each of the flapped items, one half of the participants responded “/t/,” and the other half responded “/d/.” Thus, the flaps were perceived to be ambiguous. See the Appendix for a complete list of the stimuli used in both of the current experiments.

Design. The intact original stimuli were presented in two successive blocks consisting of 24 stimuli each. Block 1 consisted of primes and Block 2 of targets. Orthogonal combination of three prime types (match, mismatch, control) and two target types (flap, careful) resulted in six conditions, illustrated in Table 1. For example, in Condition 1, a carefully

articulated prime in Block 1 was followed by the same carefully articulated item in Block 2. Across participants, each careful and casual item participated 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 *pretty* in one of the blocks, he or she did not hear any version of that word again in the same block. All remaining stimuli were fillers.

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 the target blocks, the stimuli were presented binaurally over headphones. The headphones had an attached microphone that was placed approximately 1 in. (2.54 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 the onset of the presentation of the stimulus word to the 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 excluded from the analyses, which resulted in the elimination of fewer than 1% of RTs. Any participant whose overall mean RT fell 2 standard deviations beyond the grand mean was excluded, which resulted in the elimination of 3 participants.

Prime Type (match, mismatch, control) × Target Type (careful, casual) participant (F_1) and item (F_2) analyses of variance (ANOVAs) were performed on RTs for correct responses and

Table 1
Experimental Conditions and Examples

Condition	Experiment	Example	
		Block 1	Block 2
Match			
Careful prime → careful target	1. Words without pairs	prti	prti
	2. Nonwords	oit3m	oit3m
Casual prime → casual target	1. Words without pairs	prri	prri
	2. Nonwords	oir3m	oir3m
Mismatch			
Casual prime → careful target	1. Words without pairs	prri	prti
	2. Nonwords	oir3m	oit3m
Careful prime → casual target	1. Words without pairs	prti	prri
	2. Nonwords	oit3m	oir3m
Control			
Unrelated prime → careful target	1. Words without pairs	libn	prti
	2. Nonwords	kikbæp	oit3m
Unrelated prime → casual target	1. Words without pairs	libn	prri
	2. Nonwords	kikbæp	oir3m

percentages correct for the experimental stimuli in Block 2.³ Effects are significant at the .05 level unless otherwise indicated. Accuracy was greater than 95% and produced no significant effects.⁴

RTs for the experimental stimuli as a function of prime type are listed in Table 2. Casually articulated (i.e., flapped) items were responded to more quickly than carefully articulated items, $F_1(1, 68) = 287.65$, $MSE = 6,662.72$; $F_2(1, 10) = 18.27$, $MSE = 13,832.90$, presumably because of the differences in duration. Also, the main effect of prime type was significant, $F_1(2, 136) = 3.25$, $MSE = 7,845.18$; $F_2(2, 20) = 6.33$, $MSE = 1,076.04$, indicating that items in the match and mismatch conditions were responded to more quickly than items in the control condition. Prime type and target type did not interact, $F_1(2, 136) = 1.01$, $MSE = 8,065.96$, $p = .37$; $F_2 < 1$.

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, $F_1(1, 136) = 5.82$; $F_2(1, 20) = 11.85$; and $F_1(1, 136) = 3.67$, $p = .0574$; $F_2(1, 20) = 6.27$, respectively. There was no difference between the match and mismatch conditions (F_1 and $F_2 < 1$).

Discussion

Flapped and carefully articulated words were equally effective primes for flapped and carefully articulated targets, consistent with the notion that flaps are mapped onto their underlying phonemic counterparts (i.e., /t/ and /d/). This pattern of results replicates our previous work, with one notable exception: Whereas in our earlier research flapped items were lexically ambiguous (e.g., *rater-raider*), the current stimuli (e.g., *pretty*) had no real word counterparts. Thus, lexical ambiguity is not a requisite criterion for the activation of underlying representations for flaps.

In short, the available evidence suggests that the underlying representations activated in the perception of flaps are segmental, in keeping with mediated access models. Moreover, these results are consistent with the ART framework, in which sublexical chunks corresponding to the underlying forms of /t/ and /d/ resonate with the input, creating a restoration effect in which resonances for sublexical /t/ and /d/ serve as the basis for long-term facilitative priming. Although this account is reasonable, it is also still possible that because other, similar lexical items containing flaps are regularly activated with their lexical counterparts (e.g., *rater-raider*), flapped words without counterparts may also weakly activate nonword counterparts, in which case the observed effects may still be, at least in part, lexical.

Therefore, a more stringent test of the segmental hypothesis would involve minimizing lexical activation as much as possible. To this end, we conducted Experiment 2, in which we presented minimal pairs of stimuli consisting solely of nonwords.

Experiment 2: Nonwords

Method

Participants. Seventy-two different members of the University at Buffalo community participated in the experiment. 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. Twelve pairs of bisyllabic nonwords containing intervocalic /t/s and /d/s were created. Each pair was recorded twice by the same phonetically sophisticated male speaker as in Experiment 1, once in a carefully articulated manner that preserved the /t-/d/ contrast and once in a casually articulated manner in which the /t-/d/ contrast was neutralized to produce a flap.

The mean duration for the carefully articulated stimuli was 604 ms. The mean duration for the flapped stimuli was 373 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, low-pass filtered at 10 kHz, and digitized at a sampling rate of 20 kHz using a 16-bit analog-to-digital converter. All words were edited and stored on computer disk.

The stimuli were screened by 10 listeners who indicated whether each nonword contained a /t/ or /d/. For the carefully articulated stimuli, at least 9 listeners correctly identified the intended segment. For the flaps, no more than 6 participants labeled any of the flapped items as containing a /t/ or /d/. Thus, the flaps were perceived to be ambiguous. See the Appendix for a complete list of the stimuli used in both of the current experiments.

Moreover, we conducted a second round of screening to ensure that none of the nonwords would result in the consistent activation of particular real words. A separate group of 10 participants was presented with the nonword stimuli and asked to generate the first real word that came to mind as quickly as possible. None of the nonwords resulted in the consistent generation of the same real word.

Design and procedure. The design and procedure were the same as those used in Experiment 1.

Results

RTs less than 200 ms or greater than 2,000 ms were excluded from the analyses, which resulted in the elimination of fewer than 1% of RTs. Any participant whose overall mean RT fell two

³ The Prime Type (match, mismatch, control) \times Target Type (careful, casual) interaction did not approach significance in either experiment. Thus, in all analyses of prime type, we collapsed across target type and refer to the collapsed conditions simply as *match*, *mismatch*, and *control*. Also, in the current experiment, one item (*saddle-sattle*) produced overly long RTs and thus was excluded from the item analyses. The pattern of results was the same with this item included.

Furthermore, item analyses may not be appropriate for the current experiments. The stimuli used nearly exhaust the (small) universe of available items that both met our specific criteria and successfully passed the screening phase, reducing the need for generalization beyond the present set of stimuli. The low number of items meeting our criteria also 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. The reader should bear in mind these caveats in interpreting the significance levels of all item tests reported for the current studies.

⁴ Correct production of all segments making up the stimulus was required for the response to be scored as correct. Because the flap is an acceptable allophone of /t/ and /d/, flaps produced in response to clear /t/s and /d/s were not counted as errors (and vice versa; however, it is much less common for participants to produce a clear /t/ or /d/ in response to a flap. The overwhelming tendency for American English speakers is to produce a flap, regardless of whether they are responding to a stimulus that contains a flap or a clear /t/ or /d/).

Table 2
Reaction Times and Standard Errors for Experiments 1 and 2

Experiment	Match		Mismatch		Control	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1. Words without lexical pairs	839	12	844	13	865	13
2. Nonwords	904	12	911	12	934	13

standard deviations beyond the grand mean was excluded, which resulted in the elimination of 2 participants.

Prime Type (match, mismatch, control) \times Target Type (careful, casual) participant (F_1) and item (F_2) ANOVAs were performed on RTs for correct responses and percentages correct for the experimental stimuli in Block 2.⁵ Effects are significant at the .05 level unless otherwise indicated. Accuracy was greater than 90% and produced no significant effects.

RTs for the experimental stimuli as a function of prime type are listed in Table 2. Casually articulated (i.e., flapped) items were responded to more quickly than carefully articulated items, $F_1(1, 69) = 229.52$, $MSE = 10,161.17$; $F_2(1, 11) = 45.61$, $MSE = 8,424.77$. Also, the main effect of prime type was significant by participants, $F_1(2, 138) = 3.10$, $MSE = 11,035.54$; $F_2(2, 22) = 1.52$, $MSE = 4,199.56$, $p = .2420$, indicating that items in the match and mismatch conditions were responded to more quickly than items in the control condition. Prime type and target type did not interact (F_1 and $F_2 < 1$).

Planned contrasts based on the main effect of prime type revealed a significant difference (by participants) between the match and control conditions and a marginal difference between the mismatch and control conditions, $F_1(1, 138) = 5.77$; $F_2(1, 22) = 2.97$, $p = .0987$; and $F_1(1, 138) = 3.15$, $p = .0780$; $F_2(1, 22) = 1.14$, $p = .2966$, respectively. There was no difference between the match and mismatch conditions (F_1 and $F_2 < 1$).

Discussion

Flapped and carefully articulated stimuli were equally effective primes for flapped and carefully articulated targets, consistent with the notion that flaps are mapped onto their underlying phonemic counterparts (i.e., /t/ and /d/). This pattern of results replicates Experiment 1, with one notable exception: Whereas in Experiment 1 flaps were based on lexical items (e.g., *pretty*), the current stimuli (e.g., *oytem*) were nonwords. Thus, lexical activation caused by the presence of real words is not a requisite criterion for the activation of underlying segmental representations for flaps.

In short, the available evidence suggests that the underlying representations activated in the perception of flaps are sublexical, in keeping with mediated-access models and ART. This argument is supported in spite of the more stringent test of the sublexical hypothesis, which involved minimizing lexical activation by presenting minimal pairs of stimuli consisting solely of nonwords. In short, the results of the present experiment point strongly to the resonance of sublexical chunks and input in the observed facilitative priming, demonstrating that significant lexical activation is not required for activation of underlying forms.

General Discussion

Our investigation revealed two notable findings: First, evidence from shadowing times demonstrated that matching and mismatching items are equally effective primes for flaps and careful targets. Both types of primes produced equivalent levels of facilitation for both types of targets, supporting the notion of underlying abstract representations in spoken word recognition. Second, these experiments support the sublexical ambiguity hypothesis. Experiment 1 confirms that lexical ambiguity is not a necessary condition for the activation of underlying /t/ and /d/ segmental representations. Flapped words that were only ambiguous at the segmental level produced the abstract data pattern observed in our earlier work. Moreover, Experiment 2 confirms that significant lexical activation is also unnecessary for the activation of mediating /t/ and /d/ representations. Flapped nonwords that were ambiguous at the segmental level produced the same pattern observed both in our earlier work and in Experiment 1.

The results of Experiments 1 and 2 are consistent with our interpretation of the ART framework, in which spoken word recognition may be subserved by the activation of variously sized grains or chunks. These results strongly suggest the operation of learned expectations about flaps as segmental or sublexical entities associated with the underlying forms of /t/ and /d/. However, these results cannot entirely or unequivocally rule out all possible accounts in which there is no independent sublexical level of representation. For example, the present data are also consistent with a distributed framework, in which only certain parts of lexical representations can be selectively activated. Although such accounts do not require a separate sublexical level of representation, they require the ability of sublexical representations to become acti-

⁵ Both of the experiments reported here were designed with the intent of collapsing across /t/ and /d/. That is, potential differential effects as a function of underlying voicing were not a focus of this investigation. Indeed, the design of these experiments precluded us from conducting the full complement of analyses. Moreover, the data at the lowest level of analysis were somewhat unstable given the low number of observations per cell. Nevertheless, we performed exploratory analyses for both experiments, both for the full design and for the carefully articulated targets alone, and the patterns accord with the main analyses that collapse across /t/ and /d/, with one exception: The carefully articulated /d/ stimuli that were primed by flaps produced longer than expected RTs in the nonword experiment only. Despite this one anomaly across the two experiments, the overall pattern of results is consistent with the claim that flaps activate underlying representations. Research is underway with an appropriately powerful design to examine potential differences in the priming of /d/ and /t/.

vated in the absence of full lexical activation—that is, without all aspects of the distributed representation making up the lexical items becoming activated (or at least not to the same degree)—which may in fact be tantamount to positing a separate sublexical level of representation.

The evidence for abstract underlying representations also should not be taken as evidence against the existence of more specific form-based representations. First, considerable evidence demonstrates that specific representations of spoken words are stored in memory (Church & Schacter, 1994; Goldinger, 1996) and that representational specificity has consequences for perceptual processing (Mullennix, Pisoni, & Martin, 1989), suggesting that models of spoken word recognition must incorporate both abstract and specific representations. Second, in other circumstances (e.g., different items, tasks), we too observed evidence for the activation of more specific representations. Thus, the evidence overall suggests that models of spoken word recognition must incorporate both abstract and more specific representations. Indeed, the current challenge is to determine the circumstances under which each type of representation is likely to dominate processing. Recent work in our laboratory has suggested that processing of abstract and specific information follows a predictable time course that is, at least in part, a function of frequency of experience (McLennan & Luce, 2005).

The implications of our results for mediated and direct access theories are not entirely straightforward. Clearly, our results indicate a role for sublexical representations, in keeping with mediated theories positing intermediate representations. Yet one cannot say with certainty that intermediate representations must be activated prior to contact with lexical form-based representations. Indeed, a principal tenet of the ART framework is that the chunk that is most predictive of the input dominates processing, be it lexical or sublexical. Although sublexical chunks may play a role in processing ambiguous segments, such as flaps, activation of intermediate representations need not be a constant and unbending requirement. Thus, whereas the present results clearly implicate a role for sublexical representations, their obligatory activation on the way to the lexicon is still in question. Given the current state of our knowledge regarding this issue, theoretical relativism is perhaps the most prudent stance. In fact, pitting mediated and direct access theories against one another may be inappropriate and may simply stem from the architecture of most traditional models of spoken word recognition (e.g., PARSYN; Luce, Goldinger, Auer, & Vitevitch, 2000; TRACE; McClelland & Elman, 1986), which explicitly posit distinct levels of representation for sublexical and lexical units. Alternatively, the ART framework posits that acoustic-phonetic input activates chunks corresponding to various sized representations, including sublexical and lexical representations, without necessarily requiring distinct levels or tiers of representations. Thus, when the matter is viewed from the perspective of the ART framework, rather than deciding between

mediated and direct access theories, the more fundamental issue may be to determine the precise circumstances under which different types of representation, regardless of lexical status, dominate processing.

References

- Charles-Luce, J. (1997). Cognitive factors involved in preserving a phonemic contrast. *Language and Speech*, *40*, 229–248.
- 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.
- 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. *Psychonomic Bulletin & Review*, *11*, 1084–1089.
- Fox, R. A., & Terbeek, D. (1977). Dental flaps, vowel duration and rule ordering in American English. *Journal of Phonetics*, *5*, 27–34.
- 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.
- Grossberg, S. D. (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., & Myers, C. W. (2000). The resonant dynamics of speech perception: Interword integration and duration-dependent backward effects. *Psychological Review*, *107*, 735–767.
- Ladefoged, P. (2001). *A course in phonetics* (4th ed.). New York: Harcourt College.
- 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., & McLennan, C. T. (2005). Spoken word recognition: The challenge of variation. In D. B. Pisoni & R. E. Remez (Eds.), *Handbook of speech perception* (pp. 591–609). Malden, MA: Blackwell.
- McLennan, C. T. (in press). The time course of variability effects in the perception of spoken language: Changes across the lifespan. *Language and Speech*.
- McLennan, C. T., & Luce, P. A. (2005). Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 306–321.
- McLennan, C. T., Luce, P. A., & Charles-Luce, J. (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 539–553.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86.
- Mullennix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*, *85*, 365–378.
- 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.

Appendix

Experimental and Filler Stimuli Used in Experiments 1 and 2

Experiment 1		Experiment 2	
Experimental items	Filler items	Experimental items ^a	Filler items
/d/ real words	bacon	ʃaɪdə	bəkəv
greedy / greeety	baggage	ɡɪɹdə	bægənt
feeding / feeting	boycott	pɑɪdɪl	bɔɪkɔʔ
teddy / tetty	bucket	ɕʒaɪdɪŋ	bɹkəm
reading / reating	bygone	ʃɪɹdɪŋ	bɑɪgəps
riddle / rittle	bypass	tɹdə	bɑɪrəb
saddle / sattle	cabbage	udə	kæbəv
	cabin	ɡɪɹdɪŋ	kəkəɡ
	caucus	vɔdɪŋ	səkəkə
/t/ real words	circuit	pɔdɪŋ	kɔpəɡ
	circus	ɔɪdəm	ɕʒægɹp
	coping	pɔɪdɪ	wɛrɹks
pretty / preddy	luggage		kɪkbəp
meeting / meeding	jagged		mædkɹs
butter / budder	nugget		bəmʃɛz
eating / eading	ribbon		kɑrfəp
letter / ledder	rugged		ɹɹɪŋ
water / wader	topic		tʃɑrə
	turban		vɪdə
	weapon		vutə

^aAll experimental nonword items are transcribed here with a medial /d/. However, as explained in the text, these stimuli were also presented with clear /t/s and ambiguous flaps (/r/s).

Received June 17, 2004
 Revision received December 22, 2004
 Accepted May 3, 2005 ■