Representational specificity of lexical form in the perception of spoken words

M[c]Lennan, C. T., Luce, P. A., and Charles-Luce, J.
University at Buffalo, United States of America
E-mail: mclennan@buffalo.edu, luce@buffalo.edu, cdscluce@buffalo.edu

ABSTRACT

Information-processing theories of spoken word recognition posit that perception consists of a series of stages, with representations becoming successively more abstract. These mediated-access models have been challenged by direct-access models, which assume input maps directly onto lexical representations. These models differ as to whether processing levels intervene between recoding and lexical representation. We previously examined the status of intermediate representations by exploring allophonic variation. We tested whether flaps map onto their underlying phonemic counterparts, consistent with mediated-access. We found that flaps primed their carefully-articulated counterparts (e.g., casual *rater* primed careful *rater* and *raider*) and vice versa. We argued that phonological ambiguity results in the activation of underlying representations. However, our flaps were also lexically ambiguous. The present research examines whether such priming is lexically- or sublexically-mediated by examining phonological ambiguity in the absence of lexical ambiguity. The results provide insights into when abstract representations are activated during spoken word processing.

1. INTRODUCTION

Information processing theories have typically characterized spoken word perception as being comprised of a series of linguistic stages of analysis, with form-based representations becoming successively more abstract at each stage of processing. Studdert-Kennedy [10] provides one of the earliest explicit articulations of this kind of *mediated lexical access model*.

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, Marslen-Wilson and Warren's [5] direct access featural model proposes that lexical representations are accessed directly from binary phonetic features.

Although both theories assume sensory information is initially recoded, they differ as to whether additional levels of representation intervene between sensory recoding and lexical representation.

We recently examined the status of intermediate representations by exploring the perceptual consequences of flapping [6]. According to mediated access models, flaps map onto intermediate phonemic forms /t/ and /d/. Alternatively, direct access models assume flaps map directly onto lexical representations.

2. PREVIOUS RESEARCH

In our previous work on the representation of flaps in American English, we employed 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 dependent variable of interest is a change in the participant's response to the target caused by 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 if two nominally different stimuli activate the same mental representation.

We presented carefully articulated and flapped versions of bisyllabic words (e.g., *atom, Adam*) as primes and targets, as shown in Table 1. Participants shadowed, or repeated, the targets as quickly as possible. Shadowing times served as the indices of processing speed. We attempted to determine if 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 intermediate representation, consistent with direct access models. However, equal facilitation of target processing by flaps and careful primes would indicate the existence of a shared, presumably abstract, mediating representation, consistent with mediated access models.

We obtained evidence for intermediate representations with our flapped stimuli. As shown in Figure 1, targets mismatching in articulation style were responded to more quickly than controls (i.e., words not presented as primes) and as quickly as targets matching in articulation style. Based on these data (reported in [6]), we argued that one of the circumstances leading to the activation of underlying form-based representations is sublexical ambiguity. Our flapped stimuli were ambiguous at a

ISBN 1-876346-48-5

Conditions	Experiment	Block 1: Prime	Block 2: Target
1.) Match: Careful prime → Careful target	Previous Research	ætəm	ætəm
	Words without pairs	prɪti	prɪti
2.) Casual prime → Casual target	Previous Research	ærəm	ærəm
	Words without pairs	prɪɾi	prɪɾi
3.) Mismatch: Casual prime → Careful target	Previous Research	ærəm	ætəm
	Words without pairs	prɪɾi	prɪti
4.) Careful prime → Casual target	Previous Research	ætəm	ærəm
	Words without pairs	prɪti	prɪɾi
5.) Control: Unrelated prime → Careful target	Previous Research	pepəɾ	ætəm
	Words without pairs	glob	prɪti
6.) Unrelated prime → Casual target	Previous Research	pepəɾ	ærəm
	Words without pairs	glob	prɪɾi

Table 1. Stimulus conditions and examples from our previous research and current experiment.

segmental level between /t/ and /d/. However, our stimuli were also lexically ambiguous (e.g., between *atom* and *Adam)*. Consequently, the aim of the current study is to test our claim that sublexical ambiguity, and not lexical ambiguity, is responsible for producing an abstract data pattern.

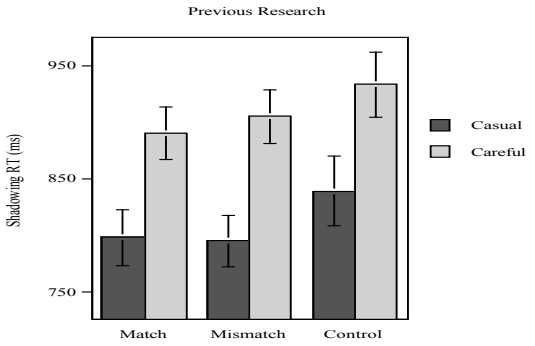

Figure 1. Shadowing times for casually articulated (i.e., flapped) and carefully articulated targets in Block 2 as a function of prime type in Block 1 for the words with lexical pairs from our previous research.

In the current experiment, we followed the same design as in our earlier work. However, rather than employing lexically ambiguous stimuli, we used words without lexical pairs, such as *pretty* and *greedy*. If our sublexical ambiguity hypothesis is correct, we should observe the same pattern of results with the present stimuli that we obtained in our previous research.

3. EXPERIMENT: WORDS WIHOUT PAIRS

3.1. Method.

3.1.1. Participants. Forty-eight members of the University at Buffalo community participated in the experiment. All participants were native American English speakers with no reported speech or hearing disorders.

3.1.2. 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*) and one was a nonword (e.g., *preddy*). Each pair was recorded twice by a phonetically sophisticated male speaker, 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 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 ten listeners who indicated whether each word contained a /t/ or /d/. In order to avoid any bias in favor of a real word, the stimuli were digitally

truncated to create nonwords (e.g., the *gr* was removed from *greedy* so that it became /idi/). For the carefully articulated stimuli, at least nine listeners correctly identified the intended segment. For the flaps, no more than six participants labeled any of the flapped items as containing a /t/ or /d/. Thus, the flaps were perceived to be ambiguous.

3.1.3. 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, control) 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. In Condition 2, a casually articulated prime (i.e., a flap) in Block 1 was followed by the same casually articulated item in Block 2, and so on. All remaining stimuli were fillers. No version of any word was presented more than once in a given block. Across all 48 participants, each word appeared in each condition.

3.1.4. Procedure. Participants were tested individually and performed a single-word shadowing task in both blocks of trials. At the beginning of the experiment, participants were *not* instructed that they would be presented with two blocks. A Macintosh Centris 650 controlled stimulus presentation and response collection. The stimuli were presented over headphones at a comfortable listening level. Participants responded by repeating the items as quickly and accurately as possible into a microphone attached to their headphones. Shadowing times were recorded from the beginning of the stimulus to the onset of the participant's verbal response.

A typical trial proceeded as follows: A cue light at the top of a response box illuminated to indicate the beginning of the trial. The participant was then presented with a stimulus word binaurally over the headphones. After the participant responded, the next trial was initiated. Each participant received a different random order of stimuli in each block.

3.2. Results.

A 3 (prime: match, mismatch, control) X 2 (target: flap, careful) repeated measures analysis of variance was performed on the shadowing times for target stimuli in Block 2. We obtained significant effects of target (F (1, 45) = 228.35, $p < .05$) and prime (F (2, 90) = 3.06, $p < .05$). The interaction between prime and target was not significant ($F < 1.0$). The main effect of target was due to faster shadowing times for flapped compared to carefully articulated stimuli. More important, the main effect of prime revealed that participants' shadowing times to targets varied as a function of the prime.

As shown in Figure 2, targets primed by flapped and careful stimuli were shadowed more quickly than targets in the control condition (i.e., targets preceded by an unrelated prime). Planned contrasts revealed significant differences between the match and control conditions (F (1, 90) = 4.77, $p < .05$) and between the mismatch and control conditions (F (1, 90) = 4.40, $p < .05$). Finally, the difference between the match and mismatch prime conditions was not significant ($F < 1.0$).

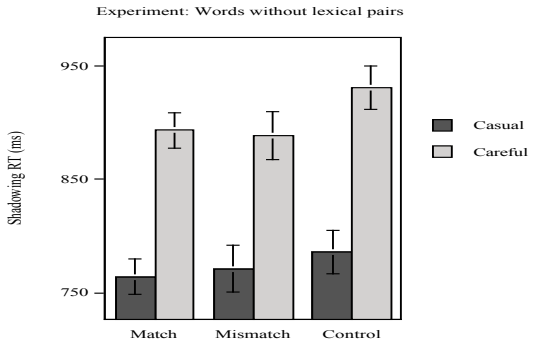

Figure 2. Shadowing times for casually articulated (i.e., flapped) and carefully articulated targets in Block 2 as a function of prime type in Block 1 for the words without lexical pairs from our current experiment.

Flapped and carefully articulated words were equally effective primes for flapped and carefully articulated targets, consistent with 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.

4. 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 mediation of underlying abstract representations in spoken word recognition. Second, the current experiment supports our sublexical ambiguity argument. The present results confirm that lexical

ISBN 1-876346-48-5

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.

The evidence for underlying mediating representations should not be taken as evidence against the existence of more specific form-based representations. Considerable evidence demonstrates that specific representations of spoken words are stored in memory [1, 2] and have consequences for perceptual processing [7], suggesting that models of spoken word recognition must incorporate both abstract and specific representations.

The current challenge is to determine the circumstances under which specific and abstract representations are likely to dominate processing. In the case of flapping, it may that segmental frequency is one such determining factor. Because flaps are more common intervocalically in American English than underlying /t/ or /d/ [8], we predict – and our results have confirmed [6] – that the effects of underlying segments on perceptual processing should be relatively late in emerging.

In short, our current results suggest coexisting specific and abstract form-based representations. Moreover, our results place the activation of underlying representations squarely at the segmental level, providing further support for mediated-access models in which input is mapped onto successively larger units of processing.

ACKNOWLEDGEMENTS

This research was supported (in part) by research grant number R01 DC 0265801 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

REFERENCES

[1] 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, & Cognition, 18*, 43-57.

[2] Goldinger, S.D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 22, 5*, 1166-1183.

[3] Grossberg, S. & Myers, C.W. (2000). The resonant dynamics of speech perception: Interword integration and duration-dependent backward effects. *Psychological Review 107, 4*, 735-767.

[4] Ladefoged, P. (2000). *A course in phonetics (5th ed.)*. San Diego: Harcourt, Brace, and Jovanovich.

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

[6] McLennan, C.T., Luce, P.A., & Charles-Luce, J. (in press). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, & Cognition.*

[7] 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, 1,* 365-378.

[8] 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.

[9] Pisoni, D. B. (1997). Some thoughts on "normalizaton" in speech perception. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 9-32). San Diego, CA: Academic Press.

[10] Studdert-Kennedy, M. (1974). The perception of speech. In. T. A. Sebeok (Ed.), *Current Trends in Linguistics*. The Hague: Mouton.