Phonological variation in spoken word recognition: Episodes and abstractions[1]

CYNTHIA M. CONNINE and ELENI PINNOW

Abstract

Phonological variation in spoken words is a ubiquitous aspect of spontaneous speech and presents a challenge for recognition of spoken words. We discuss two classes of models, abstract and episodic, that have been proposed for spoken word recognition. Abstract theories rely on inference processes and/or underspecified representations to account for spoken word recognition. Episodic theories assume a lexical representation that encodes each spoken word event with exposure frequency linked to strength of a lexical entry. A model is proposed that posits a frequency-driven phonological variant lexical representation. The model assumes that a word may have more than one variant representation and that exposure to phonological variant form influences the strength of a given variant representation. Evidence for the proposed model is reviewed for a number of variants (nasal flaps, schwa deletion and medial flaps).

1. Introduction

Spoken language comprehension is among the most complex cognitive abilities of humans – a transient physical signal that is noisy, incomplete and potentially ambiguous must be comprehended, typically under non-ideal listening conditions. One process along this route to comprehension is the mapping of acoustic-phonetic features extracted from the signal onto a lexical representation. The process of word recognition is complicated by the highly variable nature of spoken language. Of particular interest here is that the realization of

1. Preparation of the manuscript was supported by NIH grant R01DC02134. Additional support was provided by the Center for Cognitive and Psycholinguistic Science.

The Linguistic Review 23 (2006), 235–245
DOI 10.1515/TLR.2006.009
0167–6318/06/023-0235

segments can vary as a function of the surrounding segments via such phonological processes as assimilation and deletion. The variability that occurs as a result of phonological processes presents a particular challenge for listeners in the process of word recognition because phonological processes can result in large differences among alternative pronunciations. The research reviewed here focuses on the mechanisms that underlie a listener's ability to accommodate phonological variation.

There are two general classes of spoken word recognition theories: abstract and episodic. In abstract models, the speech signal is coded into abstract features which in turn serve to activate abstract lexical representations (see, e.g., Marslen-Wilson and Warren 1994; McClelland and Elman 1986; Norris, McQueen and Cutler 2000; Connine et al. 1997). Of particular importance to the present discussion is that the abstractness of the lexical representation is presumed to be such that surface details are lost during encoding. This central assumption distinguishes abstract models from episodic ones. Central to episodic models of lexical representation is the idea that the lexicon consists of traces of each heard production of a word (Goldinger 1998). The encoded traces retain detailed surface information (cf. Goldinger 1998; Johnson 2005; Pierrehumbert 2001, 2003). In the sections to follow, we review evidence that, we argue, supports a hybrid model, that is, a model in which a degree of normalization takes place, but in which multiple phonological variants of a given word may be stored.

Two strands of research motivate this work: First, research into the effects of lexical frequency on language processing (e.g., Luce 1986; Marslen-Wilson 1993; Balota and Chumbley 1984, 1990; Connine et al. 1990) and pronunciation variation (e.g., Bybee 2002; Pierrehumbert 2001; Jurafsky et al. 2001; Greenberg 1999; Fosler-Lussier and Morgan 1999; Raymond, Dautricourt and Hume 2006; Patterson, LoCasto and Connine 2003; Patterson and Connine 2001; Crystal and House 2001); and second, research suggesting that hearers encode fine details of utterances they hear, including subphonemic detail (e.g., Mullennix, Pisoni and Martin 1989; Nygaard, Sommers and Pisoni 1994; Mullennix et al. 1995; see also Johnson and Mullennix 1997 and chapters therein). Taken together, these lines of investigation suggest that word recognition is likely to be affected by phonological variation. A number of recent studies have demonstrated that this is so and have proposed models of word recognition in the face of phonological variation (e.g., MacLennan, Luce, and Charles-Luce 2003, and Luce, MacLennan, and Charles-Luce 2003 for the recognition of flapped and non-flapped variants of /t/ and /d/; Johnson 2005; Connine 2004). The research reviewed here (primarily from the authors of this review) utilizes a conversational corpus in order to establish occurrence frequencies for phonological variants. Subsequent work examines the extent to which variant frequency predicts performance in behavioral studies. Limitations of variant frequency accounts for processing phonological variation are noted along with suggestions for further research.

2. Representation of phonological variant frequency

In the following section, the current status of the role of variant frequency as an explanatory mechanism for recognition of phonological variants is reviewed. For each variant that is discussed, the relevant corpus statistics are described and are used to predict effects in processing phonological variants.

2.1. Word-final /t, d/-deletion

Word-final stops can be articulated with or without a release burst in American English. In an analysis of released/no-release patterning conducted by Crystal and House (1988) a detailed acoustic analysis of a 600 word corpus of read speech found that 59 % of the stops were complete (an identifiable hold followed by a release). A second important finding was that the occurrence of the release patterned with voicing characteristics of the stop: voiced stops in word final position include a release less often (18 %) than voiceless stops (42 %). The statistical trends for the released/no-release patterning were investigated in a phoneme monitoring experiment in which frequency-matched words ending in voiced or voiceless stops were presented with or without their final release (Deelman and Connine 2001). Deelman and Connine found that reaction time was faster for the release-bearing than for the no-release tokens. This overall finding is consistent with the relative frequency of the release type – more frequent forms (release-bearing) are responded to more quickly than less frequent forms (no-release). Of particular interest was whether the weakly probabilistic released/no-release pattern influenced word recognition in a more subtle way suggested by differences in voicing. The results revealed a complex interplay of voicing class and the integrity of the lexically matching information in the stimulus – word stimuli showed a comparable advantage for a release-bearing stimulus in phoneme monitoring reaction time compared to no-release words. However, the introduction of a mismatching segment at a word's onset to create a pseudo-word revealed a pattern of results that are consistent with the statistical likelihood of a release. Pseudo-words ending in voiceless segments showed faster phoneme monitoring reaction times for the release tokens. However, pseudo-words ending in voiced segments showed no advantage for release-bearing tokens: tokens with and without a release showed comparable priming. The complex interaction of voicing class and lexical activation clearly argues against a simple inference account for the apparent ease of processing no-release words. Rather, the results strongly suggest that the distributional facts concerning the completeness of the release across voiced and voiceless word sets is related to ease of word recognition.

2.2. *Word medial flap*

In contrast to released/no-release variants, some classes of variation show strongly dominant variants. One such variant is the American English flap (i.e., the flapped variant of coronal stops in post-tonic position as in [ˈpʰɹɪɾɪ] “pretty”). In their corpus analysis, Patterson and Connine (2001) extracted productions of potentially flapped words from the Switchboard database of American English (Godfrey, McDaniel and Holliman 1992), a collection of elicited telephone conversations between strangers on assigned topics (approximately 2,400 two-sided telephone conversations among 543 speakers). Patterson and Connine (2001) found that 96 % of the tokens (N = 2172) consisted of unambiguous, flapped productions with the remaining 4 % distributed across [t] and glottal stop productions. The predominance of the flapped production held for all words and for two different dialect regions. The hypothesis that representation of lexical form includes the highly frequent flap was examined using a phoneme identification task (Connine 2004). Uncontroversially, the specification of the flap as highly frequent was based on its predominance in the corpus analysis. Listeners identified the initial segment (*b* or *p*) in word-nonword speech continua (e.g., [ˈpʰɹɪɾɪ] – [ˈbɹɪɾɪ]). Of critical importance was that the to-be-identified segment was embedded in either a flap or a [t] bearing carrier word (e.g., [ˈpʰɹɪɾɪ]-[ˈbɹɪɾɪ] or [ˈpʰɹɪtɪ]-[ˈbɹɪtɪ]). The results showed more identification responses forming a real word (e.g., more *p* responses) when the to-be-identified segment occurred in the more frequently experienced flap carrier compared to when the to-be-identified segment occurred in the less frequently experienced [t] carrier.

2.3. *Schwa deletion*

The data just described demonstrate that a highly frequent variant, the flap, is represented but leaves open the question of representation for the very infrequent form. A more direct examination of the question of multiply represented forms was addressed for a phonological variant with no single dominant production, schwa deletion in post-tonic position (e.g., [ˈhɪstəɹɪ] ‘history’ → [ˈhɪstɹɪ]). A corpus analysis of schwa deletion frequency (Patterson, LoCasto and Connine 2003) revealed that words vary dramatically in deletion rates for three syllable words with a potential medial schwa deletion. The corpus statistics were based on the Switchboard conversational database and replicated with a smaller sample of speech using elicited productions from pairs of speakers re-telling stories (see Patterson et al. 2003 for details). The range of deletion rates permitted selection of words with either high deletion rates (greater than 50 %) and low deletion rates (less than 50 %) in an experiment conducted by Connine, Ranbom and Patterson (2005). Connine et al. presented listeners tokens from speech continua in which the duration of a schwa was manipulated. For example, the medial schwa of [ˈhɪstəɹɪ] 'history' was systematically shortened in small steps to create a token of [ˈhɪstɹɪ]. A control condition was created by replacing the initial and final syllables to create a nonword (e.g., *foshtoro*). Note that in the nonword control condition, the medial schwa and its surrounds were acoustically identical to the word condition. Subjects in the experiment were asked to indicate whether a schwa was present or absent for each token. An influence of deletion rate frequency was revealed in three aspects of the data. First, low deletion rate words showed more schwa-present judgments as compared with high deletion rate words. Second, control nonword carriers with the same physical schwa information (and surrounding segments) as their high and low deletion rate word counterparts did not differ. This indicates that the deletion rate effect for words was not a consequence of idiosyncratic properties of the schwa or its environs. Third, both high and low deletion rate words showed more vowel-present judgments relative to their nonword counterparts, but the difference was larger for the low deletion rate words. Thus, the tendency to detect a schwa in a word context depended upon the frequency of that variant form in speech. This is consistent with the findings for the medial flap (described above) and provides a more subtle influence of variant frequency – the lexical effect varied with the frequency of the variant form and is consistent with frequency-based variant representations.

2.4. Nasal flap

A second word-internal variant that has been investigated is the nasal flap (henceforth NT) where a /t/ in a ˈVn_V environment can be realized as a nasal flap (e.g., the surface form [ˈtʰwɛntɪ] 'twenty' can alternate with a surface form [ˈtʰwɛɾ̃ɪ]). Picard (1984) identified the process of nasal flapping as optional, resulting from /t/-deletion followed by flapping of the intervocalic nasal segment. Other linguistic analyses (Vaux 2000) suggest that the coronal stop is the more active of the two segments in creating nasal flapping. Ranbom and Connine (2004) showed in a corpus analysis of the Switchboard database that the nasal flap production is dominant, occurring in nearly 82 % of productions. Similar to the schwa deletion statistics, the predominance of the nasal flap form varied across words with some items more frequently produced with an enunciated NT. Similar to the previous experiments, the corpus statistics permitted the selection of stimuli that varied variant frequency (>50 % or <50 %

nasal flap occurrence). The high and low variant frequency stimuli were subsequently presented to listeners in their enunciated NT and nasal flap forms in a series of experiments. In the first of this series, a lexical decision experiment, the results showed that overall, lexical decisions were faster and more accurate for the enunciated NT variant – a finding clearly inconsistent with the overall frequency statistics. However, a more subtle influence of variant frequency emerged when responses were considered in terms of presented form (nasal flap or enunciated NT). Responses to nasal flap forms showed a variant frequency effect – lexical decision reaction times were faster and more accurate for words that had nasal flaps highly frequently (words with greater than 50% occurrence as nasal flaps) than for words that had nasal flaps infrequently. No such variant frequency effect was found for the NT productions. These results suggest that experienced frequency influences lexical representation for the nasal flap form but not the NT form. The overall advantage for the NT production (along with the variant frequency effect for the nasal flap) was also found in a repetition cross modal priming experiment. This experiment used an additional set of control words, medial consonant cluster words (e.g., *whisper*), that were presented in their intact form or as a mispronunciation without the second medial consonant (e.g., *whisser*). If activation of a lexical representation is tolerant of a missing segment, then similar priming effects for the nasal flap and the mispronounced consonant-cluster stimuli should be found. However, a significant priming effect was found for the nasal flap productions and no priming for the consonant-deleted productions. These results rule out the possibility that the nasal flap is simply recognized as a best-match for a represented NT form (Connine et al. 1997).

Ranbom and Connine (2004) suggest that the results support the presence of two phonological representations: a gradient representation for the nasal flap based on frequency of occurrence, along with a representation for the enunciated NT for all words. The advantage for processing the NT form relative to the nasal flap (in defiance of frequency) and the insensitivity of the NT form to variant frequency is clearly inconsistent with the notion that more frequent variant forms are more strongly represented. In considering possible explanations for the lack of a variant frequency effect for NT pronunciations, Ranbom and Connine (2004) suggest that there may be more than one route for the lexicalization of phonological forms for spoken word recognition. Specifically, orthographically consistent forms may achieve a strongly lexicalized representation (irrespective of its spoken frequency) via reading. For nasal flaps, the suggestion is that an enunciated NT representation is established via an orthographic route while a gradient representation for the nasal flap becomes more entrenched as the nasal flap production is encountered more often. An influence of experience with visual language on spoken word recognition is not entirely ad hoc as a similar proposal has been made by researchers in visual word recognition. Cross-talk among orthographic and phonological representations has been formalized in some models that assume explicit connections between orthographic and phonological representations. One such model, the bimodal interactive activation model, assumes that orthographic representations are automatically activated during auditory word recognition (and vice versa, see Grainger and Ferrand 1996) and support for a facilitatory relationship between modalities has been demonstrated in experiments showing an orthographic neighbor influence on processing spoken words (Ziegler, Muneaux and Grainger 2003). Some evidence for cross talk between orthographic and phonological representation that is of particular relevance is some recent research examining the nature of the speech code for words with an infrequent but orthographically consistent phonological variant. Using a novel paradigm, Inhoff, Connine, and Radach (2002; see also Inhoff et al. 2004) had participants read sentences in which fixating a target word (e.g., PRETTY) triggered auditory presentation of the frequent flap, the infrequent [t] form or a similar sounding word (e.g., GRITTY). The similar word resulted in increased post-target reading fixations, but the flapped and the [t] versions showed shorter but equivalent reading times. The equivalent effects for flapped and [t] variants suggests that the orthographic lexicon reflects both the experienced spoken word environment and orthography-to-phonology mappings. A parallel influence of orthography in processing spoken words was suggested by Connine (2004) for the hyperarticulated 't' version of words such as *pretty* where processing of this form may be facilitated by a phonological representation that developed as a result of exposure to the written form of a word. It is important to note that development of a phonological representation via reading does not rule out a variant frequency effect, but it requires a wider consideration of frequency in that input from both domains may dilute a spoken word variant frequency contribution. The implication is that predictions based solely on spoken variant frequency may be partially disconfirmed if an infrequent phonological variant is also consistent with the orthography.

3. Learning phonological variants

So far, the research reviewed here has supported the claim that lexical organization capitalizes on systematicity and frequency in representing phonological variation form. That representation of spoken form utilizes these dimensions is perhaps not surprising, since systematicity and frequency are properties inherent in other aspects of lexical knowledge. The explicit representation of phonological variants is also consistent with evidence suggesting that indexical characteristics of voices as well as within-category information about speech sounds are encoded by listeners. Listeners can recognize frequent variant forms more effectively than infrequent forms. These results suggest that lexical representations effectively evolve to accommodate the experienced language environment. However, what flexibility do listeners have in representing low frequency forms? Is simple exposure to a low frequency form sufficient to render that form more easily recognizable? A recent set of experiments investigated precisely this question for infrequently occurring schwa-deleted variants for bisyllabic schwa-deleted words (e.g., [bɪˈliːv] 'believe' → [ˈbliːv]). Previous research demonstrated that disyllabic schwa-deleted words occur infrequently (less than 11 %) compared to their schwa bearing counterparts (Patterson, LoCasto and Connine 2003) and that recognition of these forms was slow and inaccurate (LoCasto and Connine 2002). Pinnow and Connine (2005) examined the influence of exposure on these difficult to recognize forms by providing prior exposure to the stimuli in a training session in which participants were presented an auditory version of the schwa-deleted word form along with its correctly spelled orthographic form. In the training phase, participants were simply asked to listen carefully to each vowel-deleted token and read the orthographic form. The influence of exposure was assessed in a subsequent lexical decision experiment – the results showed increased accuracy (correct word identifications) for the single exposure group compared with a control group that received no training. An identical pattern of results was found for a training/test set of conditions in which the voice used in training differed from the voice used in test. Moreover, accuracy rates in the same voice condition did not differ from the different voice condition. A second experiment demonstrated that repeated exposure (four repetitions) of a variant during training increased the percentage of 'word' responses (relative to the control and similar to the single repeat group) but also served to speed the lexical decision response. Clearly, exposure to a low frequency variant is sufficient to strengthen that form's lexical representation. The lack of a talker effect is consistent with a relatively abstract representation that is immune to idiosyncratic differences among speakers. The relatively abstract nature of the information encoded in the training phase is supported by an additional experiment in which the training and test stimulus sets differed (that is participants heard one set of schwa-deleted stimuli in training and performed a lexical decision task on a new set of schwa-deleted stimuli in test). Surprisingly, the variant learning effect was also found for this second group of listeners, that is, exposure to one set of schwa-deleted words served to improve recognition of a new set of words. Even more surprising, the transfer and repetition group showed a statistically equivalent increase in 'word' responses (relative to a control group who performed the lexical decision without any training). These findings suggest that listeners are able to generalize across the set of learned stimuli and extend this knowledge to new words in a given variant class.

4. Conclusions

The preceding review shows that experience with phonological variants serves to predict performance in spoken word recognition tasks and that manipulations of experienced frequency can increase accuracy in recognizing a given form. Further, the nature of the information learned from exposure to a set of experienced variants can be extracted and applied to novel instances of that variant class. The influence of pre-existing variant frequency along with the effects of manipulated exposure frequency support a powerful role for experiential factors in lexical representation. The ability of listeners to apply learned patterns of words to a new set of words also suggest that the learning mechanism underlying the encoding of word forms is powerful enough to generalize across new patterns. The entire set of findings is consistent with the notion that exposure to a variant form underlies the development of lexical representations. The focus on the nature of the representation shifts the burden for processing phonological variants to an experience-based lexicon. However, an adequate model of word recognition may require both abstract and episodic representations. Representations along these lines have been suggested in linguistic theory that encode surface details to distinguish among speakers. Pierrehumbert (2003) discusses a distinction between parametric phonetics and abstractions across phonetic space. Parametric phonetics encodes the acoustic/articulatory space that represents individual occurrences of a segment. Abstractions across phonetic space represent word forms in the lexicon. There are many questions concerning the ways in which surface detail comes to be represented by listeners and in what form. Despite these many remaining issues, a focus on lexicalization of variant forms based on exposure frequency moves the question of what special processes or representations underlie recognition of phonological variants to the question of how do lexical representations develop and in what form. From this perspective, learning and experience take a front seat in considering the structure of the mental lexicon and its role in spoken language processing.

Binghamton University
State University of New York

References

Balota, David and James Chumbley (1984). Are lexical decisions a good measure of lexical access? The role of word-frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance* 10: 340–357.

— (1990). Where are the effects of frequency in visual word recognition tasks? Right where we said they were! Comment on Monsell, Doyle and Haggard (1989). *Journal of Experimental Psychology: General* 119: 231–237.

Bybee, Joan (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14 (3): 261–290.

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

Connine, Cynthia, John Mullennix, Eva Shernoff, and Jennifer Yelen (1990). Word familiarity and frequency in auditory and visual word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition* 16 (6): 1084–1096.

Connine, Cynthia, Larissa Ranbom and David Patterson (2005). On the representation of phonological variant frequency in spoken word recognition. Ms.

Connine, Cynthia, Titone, Debra, Deelman, Thomas, and Blasko, Dawn (1997). Similarity mapping in spoken word recognition. *Journal of Memory and Language* 37: 463–480.

Crystal, Thomas and Arthur House (1988). The duration of American-English stop consonants: An overview. *Journal of Phonetics* 16: 285–294.

Deelman, Thomas and Cynthia Connine (2001). Missing information in spoken word recognition: Non-released stop consonants. *Journal of Experimental Psychology: Human Perception and Performance* 27: 656–663.

Fosler-Lussier, Eric and Nelson Morgan (1999). Effects of speaking rate and word predictability on word pronunciation in conversational speech. *Speech Communication* 29: 137–158.

Godfrey, John, Jane McDaniel and Edward Holliman (1992). SWITCHBOARD: A telephone speech corpus for research and development. In *ICASSP Proceedings*, 517–520. San Francisco.

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

Grainger, Jonathan and Ludovic Ferrand (1996). Masked orthographic and phonological priming in visual word recognition and naming: Cross-task comparisons. *Journal of Memory and Language* 35: 623–647.

Greenberg, Steven (1999). Speaking in shorthand – a syllable-centric perspective for understanding pronunciation variation. *Speech Communication* 29: 159–176.

Inhoff, Albrecht, Cynthia Connine and Ralph Radach (2002). A contingent speech technique in eye movement research on reading. *Behavior Research Methods, Instruments, and Computers* 34: 471–480.

Inhoff, Albrecht, Cynthia Connine, Brianna Eiter, Ralph Radach, and Dieter Heller (2004). Phonological representation of words in working memory during sentence reading. *Psychonomic Bulletin and Review* 11: 320–325.

Johnson, Keith (2005). Massive reduction in conversational American English. Manuscript.

Johnson, Keith and John Mullennix (eds.) (1997). *Talker Variability in Speech Processing*. San Diego: Academic Press.

Jurafsky, Daniel, Alan Bell, Michelle Gregory and William Raymond (2001). Probablistic relations between words: Evidence from reduction in lexical production. In *Frequency and the Emergence of Linguistic Structure*, Joan Bybee and Paul Hopper (eds.), 229–254. Amsterdam: John Benjamins.

LoCasto, Paul and Cynthia Connine (2002). Rule-governed missing information in spoken word recognition: Schwa vowel deletion. *Perception and Psychophysics* 64: 208–219.

Luce, Paul (1986). Neighborhoods of words in the mental lexicon. PhD dissertation. Indiana University, Bloomington, IN.

Luce, Paul, Connor MacLennan, and Jan Charles-Luce (2003). Abstractness and specificity in spoken word recognition: Indexical and allophonic variability in long-term repetition priming. In *Rethinking Implicit Memory*, J. Bowers and C. Marsolek (eds.), 197–214. Oxford: Oxford University Press.

MacLennan, Connor, Paul Luce and Jan Charles-Luce (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory and Cognition* 29: 539–553.

Marslen-Wilson, William (1993). Issues of process and representation in spoken language understanding. In *Cognitive Models of Speech Processing*, G. Altmann and R. Shillcock (eds.), 187–209. East Sussex, UK: Erlbaum.

Marslen-Wilson, William and Paul Warren (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review* 101: 653–675.

McClelland, James and Jeffrey Elman (1986). The TRACE model of speech perception. *Cognitive Psychology* 18: 1–86.

Mullennix, John, Keith Johnson, Meral Topcu-Durgun and Lynn Farnsworth (1995). The perceptual representation of voice gender. *Journal of the Acoustical Society of America* 98: 3080–3095.

Mullennix, John, David Pisoni and Christopher Martin (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America* 85: 365–378.

Norris, Dennis, James McQueen and Anne Cutler (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences* 23: 299–370.

Nygaard, Lynne, Mitchell Sommers and David Pisoni (1994). Speech perception as a talker-contingent process. *Psychological Science* 5: 42–46.

Patterson, David and Cynthia Connine (2001). Variant frequency in flap production: A corpus analysis of variant frequency in American English flap production. *Phonetica* 58: 254–275.

Patterson, David, Paul LoCasto and Cynthia Connine (2003). Corpora analyses of frequency of schwa deletion in conversational American English. *Phonetica* 60: 45–69.

Picard, Marc (1984). English aspiration and flapping revisited. *Canadian Journal of Linguistics* 29: 42–57.

Pierrehumbert, Janet (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In *Frequency Effects and Emergent Grammar*, Joan Bybee and Paul Hopper (eds.), 137–157. Amsterdam: John Benjamins.

— (2003). Probabilistic phonology: Discrimination and robustness. In *Probabilistic Linguistics*, Rens Bod, Jennifer Hay and Stefanie Jannedy (eds.), 177–228. Cambridge, MA: The MIT Press.

Pinnow, Eleni and Cynthia Connine (2005). Learning new phonological variants in spoken word recognition: Episodes and abstraction. Paper presented at the Annual Meeting of the Psychonomic Society. Toronto, Canada.

Ranbom, Larissa and Cynthia Connine (2004). On the role of phonological variant frequency in spoken word recognition. Meeting of the Acoustical Society of America, New York City.

Raymond, William, Robin Dautricourt and Elizabeth Hume (2006). Word-medial /t, d/ deletion in spontaneous speech: Modeling the effects of extra-linguistic, lexical, and phonological factors. *Language Variation and Change* 18: 55–97.

Vaux, Bert (2000). Flapping in English. Paper presented at the Linguistic Society of America, Chicago, IL.

Ziegler, Johannes, Mathilde Muneaux and Jonathan Grainger (2003). Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory and Language* 48: 779–793.