Cognition 119 (2011) 131–136

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

The role of variation in the perception of accented speech

Meghan Sumner

Department of Linguistics, Stanford University, United States

ARTICLE INFO

Article history:
Received 4 November 2009
Revised 22 October 2010
Accepted 23 October 2010
Available online 8 December 2010

Keywords:
Phonetic variation
Phonetic categories
Speech perception
Perceptual learning
Accented speech
VOT

ABSTRACT

Phonetic variation has been considered a barrier that listeners must overcome in speech perception, but has been proved beneficial in category learning. In this paper, I show that listeners use within-speaker variation to accommodate gross categorical variation. Within the perceptual learning paradigm, listeners are exposed to p-initial words in English produced by a native speaker of French. Critically, listeners are trained on these words with either invariant or highly-variable VOTs. While a gross boundary shift is made for participants exposed to the variable VOTs, no such shift is observed after exposure to the invariant stimuli. These data suggest that increasing variation improves the mapping of perceptually mismatched stimuli.

1. Introduction

Variation in speech abounds. A single word is produced differently each time it is uttered. How do listeners make the many-to-one mapping necessary to understand a single word produced in any number of ways as an instance of one particular word, and not another? This is particularly tricky since minimal differences between words are generally meaningful. Adding non-native accents into the mix makes this task even more daunting. Consider the sentence in (1).

(1) All these pears are slipping.

This sentence, uttered by a non-native speaker of English with a variety of non-native qualities, may sound like (2) to a native speaker of English.

(2) All deece bears are sleeping.

While systematic in nature, non-native speech that includes extreme cases of phonetic variation as in (2) must nonetheless be understood by native speakers.

Listeners are sensitive to subtle phonetic cues during speech perception, and these cues are used during spoken word recognition (Clayards, Tanenhaus, Aslin, & Jacobs, 2008; Deelman & Connine, 2001; Gow, 2001, 2003; Green, Tomiak, & Kuhl, 1997; McMurray & Aslin, 2005; McMurray, Tanenhaus, & Aslin, 2009; Mitterer & McQueen, 2009; Nygaard & Pisoni, 1998; Nygaard, Sommers, & Pisoni, 1994; Sumner & Samuel, 2005, 2009). In addition, phonetic variation is beneficial under certain conditions. For example, high variability training conditions yield improved category learning (Lively, Logan, & Pisoni, 1993). Exposure to more variable speech (typically in the form of multiple speakers) also results in improved generalization (Bradlow & Bent, 2008) and word recognition (Goldinger, 1996, 1998; Nygaard & Pisoni, 1998). While these results have been critical to our understanding of category learning and the nature of phonological representations, the mechanisms that underlie these effects are not well established. Listeners must navigate extreme cases of variation without explicit, multi-session training and without the benefit of multiple speakers. In order to understand the sentence in (2) to mean (1), native English listeners must quickly adjust to gross categorical mismatches (e.g., unaspirated voiceless stops fall into the English voiced category rather than the intended voiceless one (Flege, 1984, 1993;

E-mail address: sumner@stanford.edu

0010-0277/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2010.10.018

Flege & Hillenbrand, 1987; Studdert-Kennedy, Liberman, Harris, & Cooper, 1970)).

Currently, theories accounting specifically for variation are representation-based and include theories of abstract representations (Archangeli, 1988; Cutler, McQueen, & Norris, 2006; Hallé, Chéreau, & Segui, 2000; Lahiri & Marslen-Wilson, 1991; Taft, 2006) or specific ones (Bybee, 2001; Goldinger, 1998; Johnson, 1997; Pierrehumbert, 2001, 2002). Generally, exemplar models have fared well in this area, but McQueen, Cutler, and Norris (2006) have shown that, while representations are more detailed than previously thought, relying solely on lexical specificity does not account for the wide array of data that exists. Minimal exposure to a particular speaker or an accent does not provide sufficient tokens to be helpful under these extreme conditions via exemplar dynamics.

Missing from many accounts of variation are both a focus on the process – the mechanisms that underlie accommodation that may not be based solely on detailed representations – and a controlled examination of variation itself. In this paper, the effect of variation on gross categorical mismatches is examined. Variation is manipulated within-speaker to understand the robust benefits of variation found in areas such as high variability learning and the factors that drive these effects. I propose that the traditional idea of a lack of invariance problem (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967) might be true in some contexts (e.g., an atypical sound when produced within-accent), but not in others (e.g., the same sound produced by an accented speaker). I suggest that when there is more variation, listeners have more opportunities to find a "close enough" token to latch onto and help adjust to a speaker's different system. Two predictions fall out from this idea. First, in gross categorical mismatches as in (2), more variation should yield better accommodation than less variation. And, second, this is a low-level mechanism relating similarities in the acoustic signal that may be sensitive to order effects.

In the perceptual learning paradigm, exposure to an atypical variant results in a perceptual shift or adjustment enabling an organism to adjust to its surroundings (see Goldstone (1998) for a review). Examining perceptual learning in speech, Norris, McQueen, and Cutler (2003) have shown that listeners adjust their phonetic category judgments after relatively minimal exposure to incoming speech. For example, hearing an ambiguous sound embedded in words (e.g., some sound along a [s] – [ʃ] continuum embedded in words like *ob**sc**ene* or *ambi**ti**ous*) prompts listeners to subsequently alter the way they categorize relevant speech sounds – resulting in a perceptual shift dependent on the context in which the ambiguous sound was presented. This effect is robust, found for different sounds and speakers (Clarke-Davidson, Luce, & Sawusch, 2008; Kraljic & Samuel, 2005, 2006; Eisner & McQueen, 2006; Norris et al., 2003) and has been offered as an account for listeners' accommodation of non-native accents.

Perceptual learning in speech typically examines sounds ambiguous between two phonetic categories. In reality, though, variation across accents does not always fall into an ambiguous niche between two categories, but may include much grosser differences – many times falling *unambiguously* into an *unintended* category (e.g., American English speakers consistently perceive unaspirated voiceless stops as voiced); referred to here as *bad maps*.

An example of a bad map is conflicting voice onset times (VOT) for voiced and voiceless stops across languages. In French, voiced labial stops have an average VOT of about −60 ms and voiceless stops have an average VOT of about 10 ms, while the averages in English are approximately 10 ms and 60 ms, respectively (Caramazza & Yeni-Komshian, 1974; Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973; Ladefoged, 1975; Lisker & Abramson, 1964).[1] Fig. 1 illustrates the stops for the two languages along a VOT continuum. Critically, VOT for French voiceless stops falls into the VOT range for voiced stops in English – not in the ambiguous region between categories.

The boundary shift needed to adjust to a bad map is much greater than the one needed to adjust to an ambiguous speech sound. While a flexible category boundary is most certainly beneficial under normal listening conditions that tend to involve fast speech and noisy backgrounds, it is not clear whether this subtle adjustment extends to more extreme cases of variation like bad maps.

While Fig. 1 is an accurate representation of the average VOTs, this by no means implies the presence of invariance. In fact, while VOT in French may average 10 ms, when speaking English, a speaker may produce a more English-like VOT or a more French-like one. Given that speech is naturally variable, and listeners are never exposed to invariant speech, the question examined here is whether the presence of this variation results in a greater boundary shift than continuous exposure to an invariant, but poorly mapped sound. If phonetic variation is relied upon to adjust to bad maps, we might expect exposure to a sound with a wide range of variation to result in a boundary shift greater than numerous tokens of an invariant bad map, in addition to order effects.

2. Methods

2.1. Participants

One hundred and twenty Stanford University undergraduates participated for pay or for credit. All participants were monolingual English speakers with no background (family, friends, prior residence, instruction) in French or Spanish. No hearing-related issues were reported.

2.2. Stimuli

Stimuli were produced by a 28-year-old male native speaker of French who is a late-learner of English, has a late onset of schooling in English, is new to the US, and had never previously lived in an English-speaking country.

The speaker was chosen based on the work by Pineda and Sumner (2010) that examined the distributional

[1] The VOT averages here are summed across a number of studies and are subject to great variance depending on speaking rate (Kessinger & Blumstein, 1998) and other individual differences (Allen, Miller, & DeSteno, 2003).

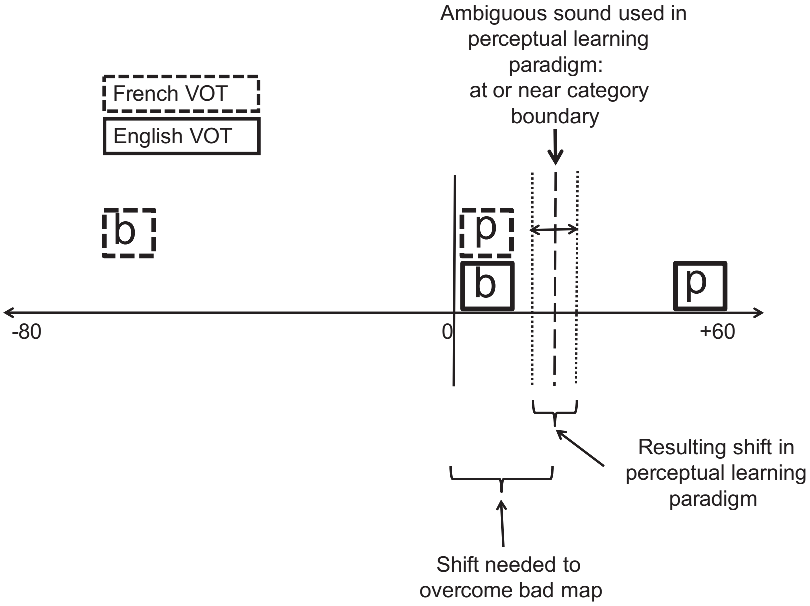

Fig. 1. VOT range for French and English, English category boundaries, and example shifts for perceptual learning effects and those needed to accommodate bad maps.

properties of French-accented speakers of English and showed that speakers with late onset of English education produced the most narrow, peaked distribution of [p] in English – with little to no overlap with the native English /p/ category. The speaker recorded lists of randomized words in isolation. Included in the list were 56 p-initial words (e.g., *poach*, *paint*), 56 b-initial words (e.g., *beef*, *barn*), and 224 words that did not include initial stops. Critically, none of these words have a minimal pair differing from the initial consonant in voicing (e.g., *poach*/**boach*; *beef*/**peef*). Three repetitions were recorded. The average VOT for voiceless labial stops was 1.28 (min = 0; max = 13.8; median = 0.5; stdev = 4.84) and the average VOT for voiced labial stops was −58.6. These numbers indicate that the speaker does indeed transfer his native temporal pattern into English (though this is clearly not the case for all speakers).

For each p-initial word, a continuum ranging from 0 ms VOT to +70 ms VOT was created, and for b-initial words, a continuum ranging from −70 ms to 0 ms was created in 5 ms steps using the PSOLA duration manipulation in Praat (Boersma, 2001). For p-initial words, the token with the most aspiration was used for the manipulation. For items with no aspiration in any of the three productions (e.g., *pork*), the onset from a similar word (e.g., *port*) was spliced onto the rime.

Seven blocks of stimuli corresponding to 10 ms VOT ranges were created, each with eight p-initial words, eight b-initial words, and 32 fillers (Table 1). For each VOT range, two words had the shortest VOT, four words had the average VOT, and two words had the longest VOT. In Block 7, for example, two of the voiceless words had 0 ms VOT, four had 5.5 ms VOT, and two had 10 ms VOT. This design avoided any obvious VOT shifts across blocks. Four quasi-randomized lists were created, ensuring that each word was presented with at least two VOTs across listeners to avoid item-specific effects. These lists were used during the exposure phase of the experiment. In addition to these items, the speaker also recorded the syllables *ba* and *pa* in order to create the *ba*–*pa* continuum for the categorization task. The continuum was created using the same methods as used for the critical words.

The blocks were used to create five experimental conditions: Control, DownShift, Random, Stable, and UpShift. In the Control condition, participants only completed the categorization task (no exposure phase). The four remaining experimental conditions included an exposure phase in addition to the categorization task. In the DownShift condition, the order of presentation of the stimuli was manipulated; participants received the stimuli ordered from Block 1 to Block 7 – gradually shifting down the VOT continuum. In the Random condition, participants received all stimuli in the exposure phase unblocked in random order. In the Stable condition, participants received invariant [p] an [b]. The VOT for p-initial words was 1.3 ms, and that for b-initial words was −58 ms, modeled

Table 1
Range of VOT for b-initial and p-initial words for Blocks 1–7.

Block	1	2	3	4	5	6	7
b VOT	−10 to 0	−20 to −10	−30 to −20	−40 to −30	−50 to −40	−60 to −50	−70 to −60
p VOT	+60 to +70	+50 to +60	+40 to +50	+30 to +40	+20 to +30	+10 to +20	0 to +10

after the natural averages of the speaker. Critically, the [p] used is an acoustic match to English [b], and is not ambiguous between the two categories. Finally, in the UpShift condition, participants received the stimuli ordered from Block 7 to Block 1.

2.3. Procedure

In the exposure phase, listeners were presented with a word printed on a monitor, followed 50 ms later by the auditory presentation of that word and were asked to rate the pronunciation of each word on a scale of 1–5, with five being native-English-like and one being not English-like. Clarke-Davidson et al. (2008) found that perceptual learning effects rely on items being perceived as words by listeners. Inherently in this task, listeners were presented with voiceless stops with VOTs near zero. Therefore, a method that forced items to be understood as words (e.g., paint) and not pseudowords (e.g., baint) was used. Listeners received one of the four experimental conditions during the exposure phase. The exposure phase contained 336 items and lasted no more than 30 min.

In the categorization phase following the exposure phase, listeners were presented with randomly ordered ba–pa syllables from the continuum and labeled the initial sounds as B or P. The VOT continuum ranged from −20 ms to +60 ms in 5 ms steps resulting in a 20 ms midpoint; close to the natural category boundary for native English speakers. Each participant responded to six repetitions of the continuum.

If variation is advantageous when perceiving bad maps, the mean proportion of B responses in the categorization task should be lower in the Random than in the Stable condition, but similar for the Stable and Control conditions. If, in addition to the variation, the relationship between variants matters, we might expect to find a greater shift (in terms of lower proportion B responses) in the DownShift condition than in the UpShift or Random conditions.

3. Results and discussion

Consistent with Norris et al. (2003), a by-subject ANOVA was run on the proportion of *B responses* for each condition, as a by-item analysis with a single continuum used in all conditions is impossible. The ANOVA showed a main effect of Condition ($F1(4, 115) = 4.29$, $p < .01$, $MSe = 4.172$). Table 2 provides the average proportion B responses for each condition, along with 95% confidence intervals. The mean percent B responses for each VOT step are provided in Fig. 2.

Table 2
Mean proportion B responses for each condition.

Condition	Proportion B	CI
Control	0.452	0.026
DownShift	0.364	0.028
Random	0.423	0.031
Stable	0.491	0.025
UpShift	0.447	0.028

Planned comparisons were used to compare the five training conditions in greater detail. Compared to the Control, a lower proportion of B responses was found for the DownShift and Random conditions (DownShift–Control ($F1(1, 46) = 11.28$, $p < .001$, $MSe = 2.152$); Random–Control ($F1(1, 46) = 7.43$, $p < .01$, $MSe = 2.826$)), a higher proportion of B responses was found for the Stable condition ($F(1, 46) = 8.17$, $p < .01$, $MSe = 1.642$), and no difference was found for the UpShift condition ($F < 1$).

Exposure to variable VOTs (Random, DownShift) resulted in a gross boundary shift, but exposure to an invariant VOT (Stable) did not.[2] Here, the *lack* of variance is costly. One might argue that this is a matter of recency, a view that is supported by the participants' behavior in the DownShift condition. This cannot be entirely ruled out, and may indeed be a factor in the large shift in the DownShift condition, but if recency were the only factor involved, UpShift responses should diverge from the Control condition (a shift which is possible, evidenced by the Stable condition). The fact that the Random condition is different from both the Control condition and the DownShift condition suggests that recency is not the only cause of the resulting gross category shift.

Consider the difference between the DownShift and the Random conditions ($F1(1, 46) = 6.24$, $p < .05$, $MSe = 2.696$). In one case (DownShift), listeners start with something native-like and end up with something non-native like. In the other case (Random), there are instances of each VOT randomized throughout the exposure phase. The Random condition is more representative of actual discourse. There is a range of variation to which a listener is exposed, and some mechanism enabling listeners to map that variable signal on to the intended meaning. These data shed light on the mechanism behind the shift – as discussed earlier, if multiple variants are related during perception, we might expect order effects to arise. There is no guarantee in the Random condition that an immediate sequence enabling a listener to map the non-native sounds to a similar one will be available. On this view, a larger shift in the DownShift condition than in the Random condition is expected. These results suggest that timing and/or sequencing is important in immediate perceptual adjustment to accents, in addition to variation.

One issue in speech perception is whether lexical effects drive accommodation, or whether there is a lower-level mapping component involved in the accommodation process. I propose that adjustments listeners make are low-level changes in perception – independent of the effect of higher order information (see also McQueen et al., 2006) – as the lexical support for the invariant bad maps is insufficient to prompt a boundary shift. This is not to say that there are no lexical effects in speech perception, as such effects are well established in the literature (e.g., Ganong, 1980; McClelland, Mirman, & Holt, 2006; Samuel, 1981, 2001). However, under certain conditions

[2] The reverse effect of the Stable condition is somewhat surprising. Without further investigation, one potential explanation is that the category expands after consistent exposure to a prototype /b/, but at this point, this is conjecture.

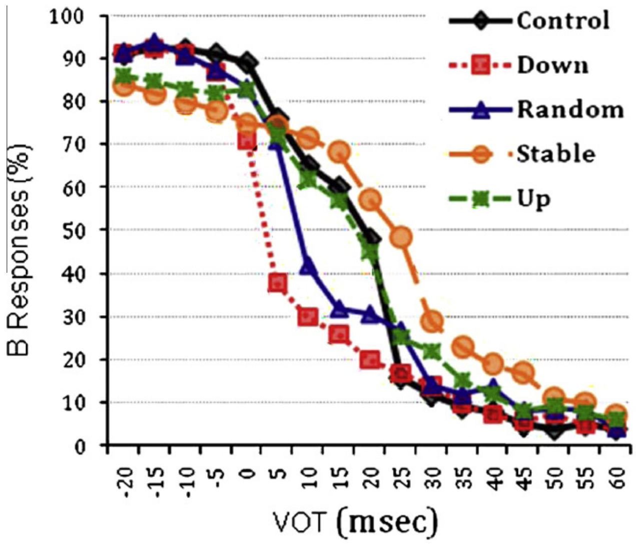

Fig. 2. Percent B responses along each VOT step for each condition.

such as adjusting to accents, some pre-lexical mapping mechanism may be used.

From this perspective, a number of predictions can be made. First, results found in high-variability learning literature have an elegant explanation – the number of speakers is not critical, but the range of variation is. Each speaker produces a range of variation for any given sound. Adding speakers increases the range of variation, the resolution and the number of opportunities for listeners to latch onto similar variants of a particular sound. Second, lexical access should improve as variation increases (in the case of bad maps). Finally, this idea may shed light on the dissociation in the literature between accent strength and comprehension (Munro & Derwing, 1999), where a strong accent rating does not imply incomprehensibility. Degree of accent may be indicative of a range of variation that is not consistent with our own, whereas comprehensibility may depend on the amount of overlap of a particular sound with our own categories.

This paper examined the role of variation in the perception of accented English. In general, consistent exposure to an invariant, bad map does not result in a boundary adjustment, while exposure to variable stimuli resulted in a gross boundary shift, dependent on order of presentation. Overall, these results suggest that variation is both necessary and beneficial in overcoming gross categorical mismatches in speech.

Acknowledgments

I would like to thank Jonathan Pelsis and Nicole Fernandez for experimental assistance. This material is based upon the work supported by the National Science Foundation Grant No. 0720054 to Meghan Sumner. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

Allen, J. S., Miller, J., & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *Journal of the Acoustical Society of America, 113*, 544–552.

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology, 5*, 183–207.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5*, 341–345.

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition, 106*, 707–729.

Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Caramazza, A., & Yeni-Komshian, G. H. (1974). Voice onset time in two French dialects. *Journal of Phonetics, 2*, 239–245.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French–English bilinguals. *Journal of the Acoustical Society of America, 54*, 421–428.

Clarke-Davidson, C., Luce, P. A., & Sawusch, J. R. (2008). Does perceptual learning in speech reflect changes in phonetic category representation or decision bias? *Attention, Perception & Psychophysics, 70*, 604–618.

Clayards, M., Tanenhaus, M. K., Aslin, R. N., & Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition*, 804–809.

Cutler, A., McQueen, J., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science, 30*, 1113–1126.

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

Eisner, F., & McQueen, J. M. (2006). Perceptual learning in speech: Stability over time. *Journal of the Acoustical Society of America, 119*(4), 1950–1953.

Flege, J. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America, 76*, 692–707.

Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. *Journal of the Acoustical Society of America, 93*, 1589–1608.

Flege, J. E., & Hillenbrand, J. M. (1987). A differential effect of release bursts on the stop voicing judgments of native French and English listeners. *Journal of Phonetics, 15*, 68–73.

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 110–125.

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.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*, 585–612.

Gow, D. W. (2001). Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language, 45*, 133–159.

Gow, D. W. (2003). Feature parsing: Feature cue mapping in spoken word recognition. *Perception and Psychophysics, 65*, 575–590.

Green, K. P., Tomiak, G. R., & Kuhl, P. K. (1997). The encoding of rate and talker information during phonetic perception. *Perception and Psychophysics, 59*, 675–692.

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

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145–165). San Diego, CA: Academic.

Kessinger, R. H., & Blumstein, S. E. (1998). Effects of speaking rate on voice-onset time and vowel production: Some implications for perception studies. *Journal of Phonetics, 26*, 117–128.

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology, 51*, 141–178.

Kraljic, T., & Samuel, A. G. (2006). How general is perceptual learning for speech. *Psychonomic Bulletin & Review, 13*, 262–268.

Ladefoged, P. (1975). *A course in phonetics*. New York: Hartcourt Brace Janovich, Inc.

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

Liberman, A., Cooper, F., Shankweiler, D., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*, 431–461.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384–422.

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/: II. The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America, 94*, 1242–1255.

McClelland, J. L., Mirman, D., & Holt, L. L. (2006). Are there interactive processes in speech perception? *Trends in Cognitive Sciences, 10*(8), 363–369.

McMurray, B., & Aslin, R. N. (2005). Infants are sensitive to within-category variation in speech perception. *Cognition, 95*, B15–B26.

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2009). Within-category VOT affects recovery from "lexical" garden paths: Evidence against phoneme-level inhibition. *Journal of Memory and Language, 60*, 65–91.

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science, 30*(6), 1113–1126.

Mitterer, H., & McQueen, J. M. (2009). Processing reduced word-forms in speech perception using probabilistic knowledge about speech production. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 244–263.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 49*, 285–310.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47*, 204–238.

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics, 60*, 355–376.

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker contigent processes. *Psychological Science, 5*, 42–46.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency effects and the emergence of linguistic structure* (pp. 137–157). Amsterdam: John Benjamins.

Pierrehumbert, J. (2002). Word-specific phonetics. *Laboratory Phonology, 7*, 101–139.

Pineda, M., & Sumner, M. (2010). Non-native long-lag VOT: Insights from a distributional analysis. In *159th Meeting of the acoustical society of America, Baltimore, MD*.

Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General, 110*, 474–494.

Samuel, A. G. (2001). Knowing a word affects the fundamental perception of the sounds within it. *Psychological Science, 12*, 348–351.

Studdert-Kennedy, M., Liberman, A. M., Harris, K. S., & Cooper, F. S. (1970). Motor theory of speech perception: A reply to Lane's critical review. *Psychological Review, 77*, 234–249.

Sumner, M., & Samuel, A. G. (2005). Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language, 52*, 322–338.

Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language, 60*, 487–501.

Taft, M. (2006). Orthographically influenced abstract phonological representation: Evidence from non-rhotic speakers. *Journal of Psycholinguistic Research, 35*, 67–78.