

Positional effects in the lexical retuning of speech perception

Alexandra Jesse · James M. McQueen

Published online: 7 July 2011
© Psychonomic Society, Inc. 2011

Abstract Listeners use lexical knowledge to adjust to speakers' idiosyncratic pronunciations. Dutch listeners learn to interpret an ambiguous sound between /s/ and /f/ as /f/ if they hear it word-finally in Dutch words normally ending in /f/, but as /s/ if they hear it in normally /s/-final words. Here, we examined two positional effects in lexically guided retuning. In **Experiment 1**, ambiguous sounds during exposure always appeared in word-initial position (replacing the first sounds of /f/- or /s/-initial words). No retuning was found. In **Experiment 2**, the same ambiguous sounds always appeared word-finally during exposure. Here, retuning was found. Lexically guided perceptual learning thus appears to emerge reliably only when lexical knowledge is available as the to-be-tuned segment is initially being processed. Under these conditions, however, lexically guided retuning was position independent: It generalized across syllabic positions. Lexical retuning can thus benefit future recognition of particular sounds wherever they appear in words.

Keywords Speech perception · Perceptual learning · Talker variability

Electronic supplementary material The online version of this article (doi:10.3758/s13423-011-0129-2) contains supplementary material, which is available to authorized users.

A. Jesse (✉) · J. M. McQueen
Max Planck Institute for Psycholinguistics,
Nijmegen, The Netherlands
e-mail: ajesse@psych.umass.edu

J. M. McQueen
Behavioural Science Institute and Donders Institute for Brain,
Cognition and Behaviour, Centre for Cognition,
Radboud University Nijmegen,
Nijmegen, The Netherlands

One challenge for listeners in understanding speech is pronunciation variability across talkers, even when talkers share language and dialect backgrounds. Listeners must thus tune in to talkers' idiosyncratic pronunciations (Eisner & McQueen, 2005; Norris, McQueen, & Cutler, 2003; Nygaard, Sommers, & Pisoni, 1994). We test here the positional generality of lexically guided perceptual learning about idiosyncratic speech sounds. First, we ask whether the position of the to-be-tuned sound within a word matters: Is there lexically guided learning wherever those sounds appear? Second, we ask whether retuning transfers across positions: Is learning about a sound, once acquired, applied to that sound wherever it appears?

Pronunciation variation can arise because of physiological (Fant, 1973; Peterson & Barney, 1952), psychological, and sociological differences among talkers (Foulkes & Docherty, 2006). Talkers thus vary in the realization of and overlap between their phonetic categories (e.g., Allen, Miller, & deSteno, 2003; Newman, Clouse, & Burnham, 2001). Listeners cope seemingly effortlessly with variation but are sensitive to it (Allen & Miller, 2004; Craik & Kirsner, 1974; Creelman, 1957; Johnson & Mullennix, 1997; Newman et al., 2001). Listeners use multiple sources of information in the signal, including spectral and durational information (Ladefoged & Broadbent, 1957; Miller & Liberman, 1979), to interpret it talker specifically. Listeners also adjust to talker idiosyncrasies through the use of knowledge about how words ought to sound (Norris et al., 2003). We examine this lexically guided retuning process here.

Lexically guided retuning was demonstrated in the Norris et al. (2003) study with an exposure test design. During lexical decision exposure, one group of Dutch listeners heard words such as *octaaf* ("octave"), where a sound ambiguous between /f/ and /s/ replaced the word-final /f/, and words

with an unambiguous final /s/, such as *radijs* (“radish”). Another group heard the same ambiguous sound in the /s/-final words and unambiguous /f/-final words. During test, /f/-trained listeners categorized more sounds from an /ɛf/-/ɛs/ continuum as /f/ than did /s/-trained listeners. Listeners therefore adjusted to speaker idiosyncrasies in line with their exposure. These adjustments generalize to novel words (McQueen, Cutler, & Norris, 2006), suggesting that prelexical representations are retuned. Knowledge about the speaker, hence, can be readily applied to all words containing the retuned sound in the same position (see also Hervais-Adelman, Davis, Johnsrude, & Carlyon, 2008; Maye, Aslin, & Tanenhaus, 2008).

Lexically guided retuning is thus beneficial for listeners (McQueen et al., 2006; Norris et al., 2003): It helps them recognize talker idiosyncrasies. Hence, one might expect it to arise in all situations. In previous studies, during exposure, the to-be-learned sound was presented always in word-final position (McQueen et al., 2006; Norris et al., 2003), always in word-medial position (Kraljic & Samuel, 2005; Stevens, 2007), or in variable positions (Eisner & McQueen, 2006). We asked here whether lexically guided learning induced by individual target words occurs if to-be-tuned sounds are always word initial. If so, this would suggest that the speech–perception system can benefit from all learning opportunities by being able to use word-specific knowledge to learn about talker-specific pronunciations from any position within words. It could be the case, however, that disambiguating information (i.e., lexical knowledge) must be available as the ambiguous sound is being heard for retuning to be possible. If learning is constrained in this way, retuning should not arise if the idiosyncratic sound is always in word-initial position, since, as a word begins, the lexical context does not yet uniquely identify what the first sound must be.

We also examined the generality of the knowledge acquired in lexical retuning. Previous studies have not shown that learning is fully generalized across the lexicon, since critical sounds were presented in the same syllabic positions during exposure and test (e.g., word-finally, McQueen et al., 2006; syllable-finally, Norris et al., 2003; or in syllable-initial word-medial position, Kraljic & Samuel, 2005, 2006). Here, we tested whether perceptual learning allows for positional transfer. Talker idiosyncrasies can be context dependent (e.g., they can depend on speaking rate; Theodore, Miller, & DeSteno, 2009). It may, therefore, be advantageous to listeners not to transfer learning across positions, unless evidence is provided for the position independence of the idiosyncrasy. Alternatively, retuning could be applied when critical phonemes are encountered in any position and then, indeed, fully generalize across the lexicon. But acoustic similarity between training and test sounds can modulate whether learning is applied (Kraljic & Samuel, 2005).

Furthermore, the realization of phonemes varies as a function of their position in a word. Word-initial phonemes tend to be longer, louder, and acoustically more distinct than their word-final counterparts (Keating, Wright, & Zhang, 1999). These allophones are perceived as instances of the same phoneme, but listeners are sensitive to allophonic variation (using it, e.g., to detect word boundaries; Quené, 1992). In this initial test of whether there is positional transfer of knowledge in lexically guided retuning, we therefore controlled for acoustic similarity, using identical fricatives across positions. Fricatives normally show less allophonic variation than do stops, for example. If there is no transfer for acoustically identical fricatives, there is unlikely to be transfer for acoustically dissimilar allophones.

In summary, we tested whether lexical knowledge obtained from target words can retune prelexical categories when critical sounds are all presented word-initially. We also tested whether retuning is applied position-specifically or position-independently. Using the Norris et al. (2003) design, one group heard, during the lexical decision exposure phase, a sound ambiguous between /f/ and /s/ replacing /f/ in words; another group heard the same ambiguous sound replacing /s/ in words. This critical sound was presented word-initially in **Experiment 1** and word-finally in **Experiment 2** (see Table 1). Both experiments contained two test conditions: Half of the participants from each training group categorized an onset /f/-/s/ continuum; half categorized a coda continuum. If knowledge about the target word can retune categories on the basis of word-initial exposure, learning should be found in the onset-to-onset condition of **Experiment 1**. If learning transfers across syllabic positions, it should also be evident in the onset-to-coda condition.

Experiment 1

Method

Participants Ninety-eight paid right-handed native-Dutch university students with no reported hearing problems were tested (26 in each onset-to-onset group; 24 in each onset-to-coda group). One participant was excluded from each onset-to-onset group due to equipment failure, and another from each group due to <50% acceptance of ambiguous targets as words. There were 7 pretest participants.

Test materials and pretest The nonsense syllables /sø/, /fø/, /xø/, /øx/, /øf/, and /øx/ were digitally recorded at 44 kHz spoken by a female Dutch speaker and redigitized at 16 kHz. Fricatives excised from /sø/ and /fø/ were set to the mean fricative duration in the target-initial exposure

Table 1 Experimental design

Expt.	Condition	Exposure: Auditory Lexical Decision	Test: Categorization
1	Onset-to-onset	Word-initial critical sounds /f/-training group: ? <i>antoom</i> & <i>siroop</i> /s/-training group: <i>fantoom</i> & ? <i>iroop</i>	5-step syllable-initial feu-seu continuum
	Onset-to-coda	Word-initial critical sounds /f/-training group: ? <i>antoom</i> & <i>siroop</i> /s/-training group: <i>fantoom</i> & ? <i>iroop</i>	5-step syllable-final euf-eus continuum
2	Coda-to-coda	Word-final critical sounds /f/-training group: <i>octaa?</i> & <i>albatros</i> /s/-training group : <i>octaaf</i> & <i>albatro?</i>	5-step syllable-final euf-eus continuum
	Coda-to-onset	Word-final critical sounds /f/-training group: <i>octaa?</i> & <i>albatros</i> /s/-training group: <i>octaaf</i> & <i>albatro?</i>	5-step syllable-initial feu-seu continuum

words (129 ms) with Praat's PSOLA algorithm (Boersma & Weenink, 2005). The excised fricatives were mixed to create an equally spaced 21-step continuum; frication noises' amplitudes were added in different proportions across the continuum (cf. Norris et al., 2003). A 413-ms vowel excised from a /xø/ syllable was linearly ramped down in amplitude to 0% over its final 75 ms and concatenated onto the fricative steps as onsets. Frication noises were set to the average root-mean square intensity (.0133 Pascal) of those fricatives in the word-initial exposure words. The same intensity-adjusted fricative steps were then concatenated as codas onto the same vowel, except that instead of the vowel, the steps were linearly ramped down to 0% over their last 75 ms.

Pretest participants heard over headphones ten repetition blocks of 15 /fø/–/sø/ continuum steps (steps 3, 5–15, and 17–19); each repetition was presented in a newly randomized order. Stimuli were presented 150 ms after trial onset. After stimulus onset, participants had 2,000 ms to indicate by button press (“S” or “F”) as quickly and as accurately as possible whether they had heard /s/ or /f/. Figure 1 shows the average percentage of [f] responses to each continuum step. Step 9 (49% [f] responses) was selected as the

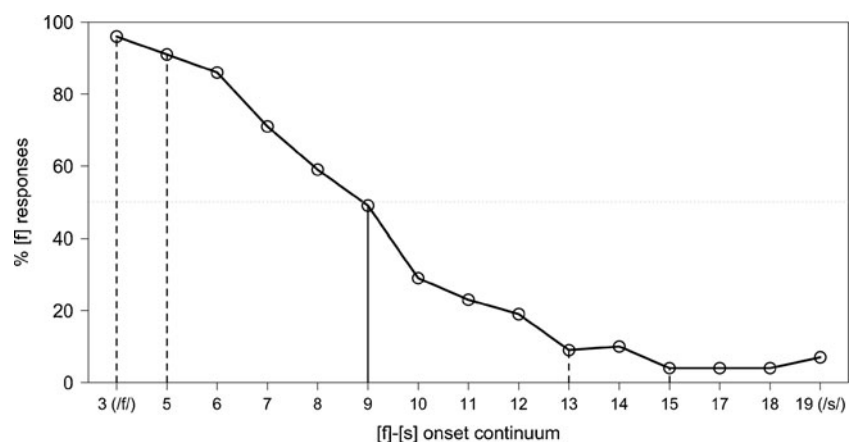
ambiguous training stimulus. Steps 3, 5, 9, 13, and 15 were selected for the test continua.

Exposure materials Twenty /f/-initial and 20 /s/-initial Dutch target words were selected (see the [Supplementary Materials](#)). None of these forms a word when their onset is replaced with the other fricative. Word sets were equated on word frequency (Baayen, Piepenbrock, & Gulikers, 1995), syllable length, lexical stress, and fricative-adjacent vowel. Sixty filler words and 100 phonotactically legal filler nonwords were also selected. No stimuli contained /s/, /f/, /z/, or /v/ (other than the target-initial sounds). Half of all items were bisyllabic; half were trisyllabic.

All items were recorded with the pretest stimuli by the same speaker. Recordings were redigitized from 44 to 16 kHz. Target words were also recorded with an unvoiced velar fricative /x/ replacing /s/ and /f/. The /x/ portion of these tokens was replaced with the ambiguous training sound to create the critical training words.

Procedure Participants were randomly assigned to one of two exposure groups. The /f/-training group was presented with 20 natural /s/-initial and 20 ambiguous /f/-initial

Fig. 1 Mean proportions of [f] responses as a function of step across the [f]–[s] onset continuum in the pretest. The step marked with the solid line was selected as ambiguous sound for the training phase in Experiments 1 and 2. This ambiguous step and the steps marked with the dashed lines were also used as steps of the test continua in Experiments 1 and 2



words. The /s/-training group received 20 natural /f/-initial and 20 ambiguous /s/-initial words. Both groups also heard all 160 fillers. Participants completed a lexical decision task individually in a sound-attenuated booth. Trials began by showing the labels “JA/NEE” (“yes/no”) on a computer screen. These labels corresponded to the buttons of a response box, and their assignment to sides was counterbalanced across participants. A word was presented at a comfortable fixed level over headphones 400 ms after label onset. The next trial was presented 500 ms after a response or, otherwise, 2,000 ms after stimulus onset. Participants had to indicate as quickly and accurately as possible whether they had heard Dutch words. Trial order was pseudorandomized in that 12 fillers were presented first and targets were separated minimally by 3 fillers. The same lists were used for all groups, with the natural and ambiguous versions of the targets appropriately exchanged.

Participants were informed about the test phase only after the exposure phase. In this test phase, six repetitions of five steps of the onset or coda continuum were presented for categorization. Half of the participants in each exposure group categorized the onset continuum; half categorized the coda continuum. The procedure was as in the pretest.

Results and discussion

Results were analyzed with linear mixed-effect models, with a logistic linking function for categorical dependent variables. Models were evaluated by systematic stepwise comparisons using likelihood-ratio tests until the simplest, best-fitting model was obtained. Starting with the full model, we removed factors that did not contribute to a better fit. Each time, the factor with the largest p value was tested first for possible removal. Main effects were never removed before interactions involving the same factor. Best-fitting models for lexical decision included partici-

pants and items as random factors. Ambiguity (natural vs. ambiguous) and fricative (/f/ vs. /s/) were contrast-coded fixed factors. Best-fitting models for phonetic categorization included participants as random factor. Training group (/f/ vs. /s/, contrast coded) and continuum step (as numerical factor, centered on the middle step) were fixed factors. P -values associated with the beta-weights of the models are reported. P -values for the latency models were estimated using Markov chain Monte Carlo simulations ($n = 10,000$).

Lexical decision Correct yes responses were less likely to words containing ambiguous fricatives than to those containing natural fricatives (onset-to-onset, $\beta = 1.25$, $SE = 0.46$, $p = .007$; onset-to-coda, $\beta = 1.57$, $SE = 0.48$, $p = .001$; see Table 2). Correct responses were slower to words with ambiguous fricatives (onset-to-onset, $\beta = -0.22$, $SE = 0.10$, $p = .03$; onset-to-coda, $\beta = -.131$, $SE = 0.04$, $p = .001$).

Phonetic categorization There was no difference across training groups in number of [f] responses (see Fig. 2; onset-to-onset, not a predictor, $\chi^2(1) = 2.04$, $p = .15$; onset-to-coda, $\beta = 0.21$, $SE = 0.57$, $p = .71$). That is, no learning appeared to occur. The more /f/-like a sound, the more likely an [f] response was (onset-to-onset, $\beta = -0.56$, $SE = 0.03$, $p < .00001$; onset-to-coda, $\beta = -0.35$, $SE = 0.02$, $p < .00001$). For the onset-to-coda condition, the response curve across the continuum was steeper for the /s/-training than for the /f/-training group ($\beta = 0.12$, $SE = 0.05$, $p = .01$). Listeners were thus sensitive to the continuum manipulation but did not appear to learn about the ambiguous sound.

Experiment 2

Experiment 1 failed to show lexical retuning from word-initial sounds. Although this null effect is difficult to interpret, it means that we could not assess whether

Table 2 Mean percentages of correct responses and mean response latencies (in milliseconds from target word offset; RT) for ambiguous and natural /f/ and /s/ target words during the auditory lexical decision

Exposure	Test Group	Natural				Ambiguous			
		/f/-Words		/s/-Words		/f/-Words		/s/-Words	
		% Correct	RT	% Correct	RT	% Correct	RT	% Correct	RT
Onset [Expt. 1]	Onset	90	242	91.5	256	79.3	306	84.9	317
	Coda	91	198	91.4	226	78.7	309	80	271
Coda [Expt. 2]	Onset	95.9	136	93.9	219	86.7	250	94.5	254
	Coda	96.1	170	96.2	247	88.1	270	91.5	252

exposure phase in Experiment 1 (onset training) and Experiment 2 (coda training), for onset and coda test groups, respectively

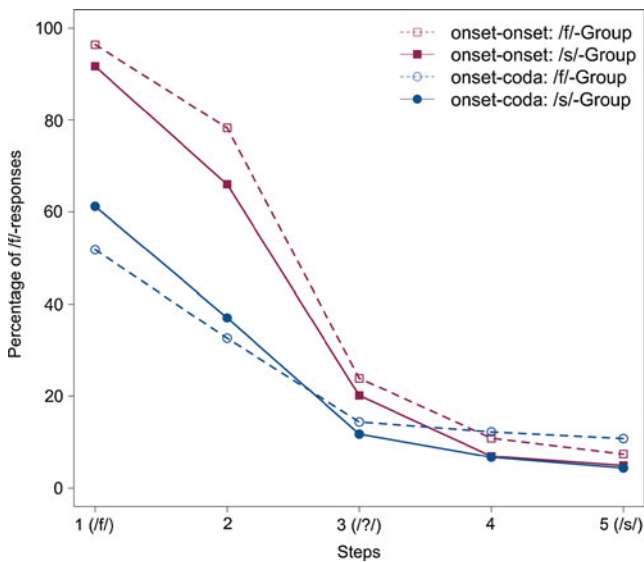


Fig. 2 Mean proportions of [f] responses as a function of step across the [f]–[s] continuum in Experiment 1. Rectangles show the test data for the onset continuum; dots for the coda continuum. Solid lines show data for the /s/-training group, and dashed lines data for the /f/-training group

learning transfers across syllabic positions. Experiment 2 therefore induced lexical retuning from word-final sounds (as in Norris et al., 2003) to test for possible transfer of learning to syllable-initial sounds. The experiments were otherwise identical.

Method

Participants Ninety-two new paid participants from the Experiment 1 population were tested (22 in each coda-to-coda group; 24 in each coda-to-onset group). One participant was excluded in the coda-to-onset condition due to an experimenter error, and one in the coda-to-coda condition due to a <50% ambiguous-target acceptance rate.

Materials Twenty /f/-final and 20 /s/-final Dutch target words were selected following the same criteria as in Experiment 1 (see the Supplementary Materials). These items were recorded along with the Experiment 1 materials. Targets recorded with /x/ replacing their respective fricative were used to create ambiguous versions (with the Experiment 1 ambiguous exposure sound). Ambiguous and unambiguous fricative-final words replaced the fricative-initial exposure words. Materials (exposure fillers and test continua) were otherwise unchanged.

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

Results and discussion

Lexical decision Correct yes responses were less likely to words containing ambiguous fricatives than to those with natural fricatives in the coda-to-coda condition (see Table 2; coda-to-coda, $\beta = 1.27$, $SE = 0.50$, $p = .01$; coda-to-onset, $\beta = 0.68$, $SE = 0.44$, $p = .13$). Correct responses were slower to words with ambiguous fricatives (coda-to-coda, $\beta = -0.11$, $SE = 0.05$, $p = .02$; coda-to-onset, $\beta = -0.17$, $SE = 0.06$, $p < .003$). For the coda-to-onset condition, responses were slower for /s/-words than for /f/-words ($\beta = -.011$, $SE = 0.06$, $p < .045$).

Phonetic categorization More [f] responses were given by the /f/-training groups than by the /s/-training groups (see Fig. 3; coda-to-coda, $\beta = 1.39$, $SE = 0.49$, $p = .005$; coda-to-onset, $\beta = 1.33$, $SE = 0.64$, $p = .04$). This indicates lexical retuning and its transfer across positions. For the coda-to-onset condition, the training effect was larger the more /f/-like stimuli became (coda-to-coda, not a predictor, $\chi^2(1) = 2.39$, $p = .12$; coda-to-onset, $\beta = 0.16$, $SE = 0.07$, $p = .01$). Participants in both conditions gave more [f] responses, the more /f/-like stimuli became (coda-to-coda, $\beta = -0.34$, $SE = 0.02$, $p < .00001$; coda-to-onset, $\beta = -0.59$, $SE = 0.03$, $p < .00001$).

Follow-up analyses across test position conditions confirmed that learning transfers across positions. Training condition affected categorization ($\beta = 1.45$, $SE = 0.42$, $p = .0005$). This learning effect was not modulated by

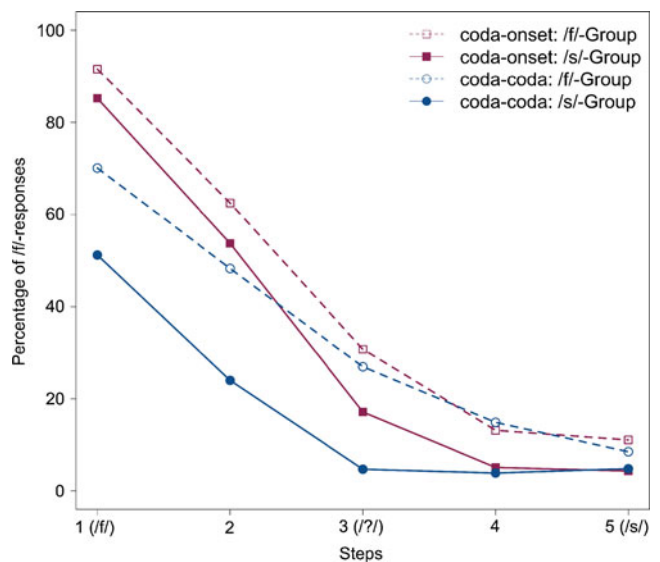


Fig. 3 Mean proportions of [f] responses as a function of step across the [f]–[s] continuum in Experiment 2. Dots show the test data for the coda continuum; rectangles for the onset continuum. Solid lines show data for the /s/-training group, and dashed lines data for the /f/-training group

test position (not a predictor, $\chi^2(1) = 0.52, p = .47$), and also not for steps 1 and 2 only (not a predictor, $\chi^2(1) = 0.58, p = .45$).

A final analysis compared effects for each test position across experiments. Training effects (/f/- vs. /s/-training group) were found for the coda ($\beta = 0.92, SE = 0.39, p = .018$) and onset test continua ($\beta = 0.73, SE = 0.35, p = .035$). Exposure position (initial vs. final) modulated the coda-test training effect ($\beta = 1.52, SE = 0.75, p = .043$) but not the onset-test training effect (not a predictor, $\chi^2(1) = 0.36, p = .55$). Thus, while the coda test showed evidence of lexical retuning only after word-final exposure, the onset test showed that retuning may not be entirely absent after word-initial exposure: Although there was no statistically significant retuning effect in the onset-to-onset condition, the trend found there was not significantly different from the effect found in the coda-to-onset condition.

General discussion

Lexical knowledge obtained from individual words resolves pronunciation ambiguities and retunes prelexical categories (McQueen et al., 2006; Norris et al., 2003). The present study demonstrates the generalizability of this acquired speaker knowledge over phonemic positions within syllables and words. Lexically guided retuning affects the perception of critical phonemes independently of their position. There are, however, positional restrictions on whether retuning arises. When idiosyncratic pronunciations appeared in the initial position in individual target words, there was no clear evidence of lexically guided retuning. Although there was an indication that the retuning effect is not entirely absent after initial-only exposure (the lack of a significant interaction of exposure condition and retuning effect in the onset-test data), the retuning effect after initial-only exposure was itself not significant.

One possible explanation for the lack of a robust retuning effect from word-initial sounds is that, on some trials, lexical knowledge about the words containing those ambiguous sounds never became sufficiently available for listeners to be able to recognize those items as words. If so, no retuning would be expected. Acceptance rates for the onset-exposure words were indeed lower than those for the coda-exposure words (see Table 2). But acceptance rates around 80% are not too low to expect lexically guided retuning: Eisner and McQueen (2005) obtained acceptance rates as low as 72% and, nevertheless, observed learning from coda-exposure words.

A more plausible explanation is that lexical knowledge may guide retuning only when it is available quickly and reliably enough to disambiguate the identity of the

ambiguous sounds as those sounds are being processed. Previous research has shown that listeners can use other information sources to adjust to onset-only idiosyncrasies. Phonotactic information (Cutler, McQueen, Butterfield, & Norris, 2008) and visual speech cues (Bertelson, Vroomen, & de Gelder, 2003) in word-initial position—for example, can induce retuning. In these cases, the disambiguating information is available as (or shortly after) the critical sound is heard. Timing (when the disambiguating information is available relative to the ambiguous sound) could thus be critical in determining whether perceptual learning about speech can occur. Our results are in line with this timing hypothesis. Reliable lexically guided retuning was found only in conditions where lexical knowledge was available before the ambiguous sounds were heard—that is, when the sounds were in word-final position. The timing hypothesis nonetheless predicts that lexically guided retuning could arise in response to sounds in any position of the word (including word-initial position), provided lexical knowledge is available rapidly enough to disambiguate the ambiguous sounds. This could explain the weak indication of an effect on the onset test continuum after onset exposure.

Lexically guided retuning can be explained in at least two ways. On one account, lexical feedback continuously modulates prelexical processing and results in lexical influences on phonemic decision making and on retuning of prelexical representations (see, e.g., Mirman, McClelland, & Holt, 2006). On another, dual-mechanism account, there is continuous feedforward merging of prelexical and lexical information for postlexical phonemic decision making (Norris, McQueen, & Cutler, 2000), but lexical feedback operating over time for prelexical retuning (Norris et al., 2003). The present demonstration that retuning depends on availability of lexical knowledge as the ambiguous sounds are being processed is consistent with both accounts. In particular, this demonstration does not rule out a separate, potentially later learning process (e.g., cumulation of retuning over exposures), as proposed by the dual-mechanism account. This account is supported by recent findings: Lexically guided retuning appears to be prelexical (McQueen et al., 2006), but lexical effects on decisions can occur without prelexical processing consequences and, thus, appear to be postlexical (McQueen, Jesse, & Norris, 2009).

Position-general application of talker-idiosyncratic knowledge allows for full transfer of learning across the lexicon. Listeners can treat idiosyncrasies as position independent, even when exposure is limited to one position. They can thus apply this knowledge to facilitate comprehension whenever unusual sounds are heard. But this finding cannot be taken as evidence that prelexical representations are phonemic rather than allophonic. One

problem with phonemic representations is that there could be overapplication of learning when talker idiosyncrasies are indeed position specific or otherwise context dependent. The system would nevertheless be protected from such unwarranted overgeneralizations if there were a reduced degree of similarity across positions and contexts in such cases. Acoustic similarity between training and test sounds can modulate whether knowledge about a speaker is transferred to another speaker (Kraljic & Samuel, 2005). It is possible that in the present study, therefore, that learning was fully transferred across positions precisely because acoustic similarity across positions was controlled. The same ambiguous training token was used for word-initial and word-final exposure, and the two test continua contained the same frication noises. In natural materials, however, some degree of positional specificity of learning is expected, determined by the degree of acoustic similarity across positions. The less position-variable fricatives will likely result in less positional specificity in learning than will the more position-variable stop consonants (Redford & Diehl, 1999). The present findings therefore do not show whether prelexical representations are allophonic or phonemic. But they do show that, when sounds across positions are acoustically matched, perceptual learning can generalize over position.

Conclusion

These experiments examined two positional effects in lexically guided retuning. First, we found no convincing evidence that lexical retuning arises in response to ambiguous speech sounds presented only in word-initial position. Lexical knowledge may not have been sufficiently available during exposure to guide retuning in this condition. Second, listeners can transfer knowledge about idiosyncratic pronunciation variation across positions within words. Although acoustic similarity may modulate positional transfer to prevent overgeneralization of learning, it is beneficial that retuning, given sufficient acoustic overlap, does apply over positions. Retuning thus helps listeners recognize the words of a talker speaking in an unfamiliar way.

Author Note We thank the student assistants of the MPI Comprehension group for their help with the experiments. We also thank Anne Cutler, Arthur Samuel, and two anonymous reviews for constructive comments on previous drafts. Parts of this work were presented at Interspeech 2007 (Jesse & McQueen, 2007) and at the 49th Annual Meeting of the Psychonomic Society, Chicago (McQueen & Jesse, 2008). Please address correspondence to A. Jesse, Department of Psychology, University of Massachusetts, Tobin Hall, Amherst, MA 01003, or to ajesse@psych.umass.edu.

References

- Allen, J. S., & Miller, J. L. (2004). Listener sensitivity to individual talker differences in voice-onset-time. *Journal of the Acoustical Society of America*, *115*, 3171–3183.
- Allen, J. S., Miller, J. L., & deSteno, D. (2003). Individual talker differences in voice-onset-time. *Journal of the Acoustical Society of America*, *13*, 544–552.
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (CD-ROM)*. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.
- Bertelson, P., Vroomen, J., & de Gelder, B. (2003). Visual recalibration of auditory speech identification: A McGurk aftereffect. *Psychological Science*, *14*, 592–597.
- Boersma, P., & Weenink, D. (2005). *Praat: Doing phonetics by computer* (Version 4.3.14). Retrieved from <http://www.praat.org/>.
- Craik, F. I. M., & Kirsner, K. (1974). The effect of speaker's voice on word recognition. *Quarterly Journal of Experimental Psychology*, *26*, 274–284.
- Creelman, C. D. (1957). Case of the unknown talker. *Journal of the Acoustical Society of America*, *29*, 655.
- Cutler, A., McQueen, J. M., Butterfield, S., & Norris, D. (2008). Prelexically-driven perceptual retuning of phoneme boundaries. In *Proceedings of Interspeech 2008* (p. 2056). Brisbane, Australia.
- Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics*, *67*, 224–238.
- Eisner, F., & McQueen, J. (2006). Perceptual learning in speech: Stability over time (L). *Journal of the Acoustical Society of America*, *119*, 1950–1953.
- Fant, G. (1973). *Speech sounds and features (Vol. XI)*. MIT Press: Cambridge.
- Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, *34*, 406–438.
- Hervais-Adelman, A., Davis, M. H., Johnsrude, I. S., & Carlyon, R. P. (2008). Perceptual learning of noise vocoded words: Effects of feedback and lexicality. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 460–474.
- Jesse, A., & McQueen, J. M. (2007). Prelexical adjustments to speaker idiosyncrasies: Are they position-specific? In H. van Hamme & R. van Son (Eds.), *Proceedings of Interspeech 2007* (pp. 1597–1600). Adelaide: Causal Productions.
- Johnson, K., & Mullenix, J. W. (Eds.). (1997). *Talker variability in speech processing*. San Diego: Academic Press.
- Keating, P. A., Wright, R., & Zhang, J. (1999). Word-level asymmetries in consonant articulation. *UCLA Working Papers in Phonetics*, *97*, 157–173.
- 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). Generalization in perceptual learning for speech. *Psychonomic Bulletin & Review*, *13*, 262–268.
- Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America*, *29*, 98–104.
- Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science*, *32*, 543–562.
- McQueen, J. M., & Jesse, A. (2008). Positional specificity in lexical retuning of speech perception. *Abstracts of the Psychonomic Society*, *13*, 7.
- McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science*, *30*, 1113–1126.

- McQueen, J. M., Jesse, A., & Norris, D. (2009). No lexical-prelexical feedback during speech perception. Or: Is it time to stop playing those Christmas tapes? *Journal of Memory and Language*, *61*, 1–18.
- Miller, J. L., & Liberman, A. M. (1979). Some effects of later-occurring information on the perception of stop consonant and semivowel. *Perception & Psychophysics*, *25*, 457–465.
- Mirman, D., McClelland, J. L., & Holt, L. L. (2006). An interactive Hebbian account of lexically guided tuning of speech perception. *Psychonomic Bulletin & Review*, *13*, 958–965.
- Newman, R. S., Clouse, S. A., & Burnham, J. L. (2001). The perceptual consequences of within-talker variability in fricative production. *Journal of the Acoustical Society of America*, *109*, 1181–1196.
- Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *The Behavioral and Brain Sciences*, *23*, 299–370.
- Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*, 204–238.
- Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, *5*, 42–46.
- Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, *24*, 175–184.
- Quené, H. (1992). Durational cues for word segmentation in Dutch. *Journal of Phonetics*, *20*, 331–350.
- Redford, M. A., & Diehl, R. L. (1999). The relative perceptual distinctiveness of initial and final consonants in CVC syllables. *Journal of the Acoustical Society of America*, *106*, 1555–1565.
- Stevens, M. (2007). *Perceptual adaptation to phonological differences between language varieties*. Unpublished dissertation. Ghent: University Ghent.
- Theodore, R. M., Miller, J. L., & DeSteno, D. (2009). Individual talker differences in voice-onset-time: Contextual influences. *Journal of the Acoustical Society of America*, *125*, 3974–3982.