American Psychological Association
Journal of Experimental Psychology: Learning, Memory, and Cognition

ISSN: 0278-7393
2021, Vol. 47, No. 1, 184–194
http://dx.doi.org/10.1037/xlm0000805

Sentence Context Guides Phonetic Retuning to Speaker Idiosyncrasies

Alexandra Jesse
University of Massachusetts Amherst

Speakers vary in their pronunciations of the sounds in their native language. Listeners use lexical knowledge to adjust their phonetic categories to speakers' idiosyncratic pronunciations. Lexical information can, however, be inconclusive or become available too late to guide this phonetic retuning. Sentence context is known to affect lexical processing, and listeners are typically more likely to categorize steps of a phonetic continuum in line with the semantic content of a sentence. In a series of experiments, we tested whether preceding sentence context can guide phonetic retuning. During a passive-listening exposure phase, English listeners heard a sound ambiguous between /s/ and /f/ spliced into the onset position of minimal word pairs (e.g., *sin* vs. *fin*). Sentence context disambiguated these minimal pairs as /s/-initial for 1 group of listeners and as /f/-initial for another group. At subsequent test, listeners categorized more steps on a /sa/-/fa/ continuum in line with their prior exposure; that is, when sentence context had disambiguated the ambiguous sound during exposure as /s/, listeners gave more /s/ responses than /f/ responses at test. These aftereffects occurred independently of whether contrastive phonemes from the respective other category were provided. No phonetic retuning was found when the disambiguating sentence contexts were replaced with neutral ones. Overall, these results provide evidence that sentence context can guide phonetic retuning, therefore expanding the usefulness of phonetic retuning as a tool for listeners to accommodate speakers.

Keywords: speech perception, perceptual learning, talker variability, sentence context effects

The recognition of speech requires listeners to map incoming acoustic information onto their mental representations of basic sound units. Speakers vary, however, in their production of speech sounds (e.g., Allen, Miller, & DeSteno, 2003; Newman, Clouse, & Burnham, 2001; Theodore, Miller, & DeSteno, 2009) due to physiological (Laver & Trudgill, 1979; Peterson & Barney, 1952) and psychological and sociological (e.g., Clopper & Pisoni, 2004; Foulkes & Docherty, 2006) factors. Although some idiosyncrasies are unique to a talker, others, such as accents, are shared by a group of talkers. Listeners are sensitive to talker variation (Heald & Nusbaum, 2014; Yakel, Rosenblum, & Fortier, 2000) and quickly accommodate to speakers (e.g., Clarke & Garrett, 2004; Vroomen, van Linden, de Gelder, & Bertelson, 2007). In the present study, we tested whether listeners could use sentence context to guide the retuning of their phonetic categories to accommodate idiosyncratic pronunciations of speech sounds.

To adapt to idiosyncratic realizations of a speech sound, listeners alter the boundaries of their phonetic categories to incorporate the idiosyncrasy into the category intended by the speaker. Lexical knowledge can guide this phonetic retuning by disambiguating the idiosyncrasy, as first documented by Norris, McQueen, and Cutler (2003). During exposure, for one group of Dutch listeners, a sound ambiguous between /s/ and /f/ replaced the /f/ at the end of critical words like *octaaf* ("octave"), where /s/ would not form a word. This /f/-training group also received words ending in a clear /s/ (e.g., *radijs* ["radish"]). For another group of participants, the same ambiguous sound replaced the word-final /s/ in words like *radijs* where /f/ would not form a word. This /s/-training group also received words ending in a clear /f/ (e.g., *octaaf*). Both groups heard the critical words one-by-one in an auditory lexical-decision task, intermixed with filler words and nonwords that did not include /s/ or /f/. At a subsequent test, the /s/-training group categorized more steps of an /ɛs/-/ɛf/ continuum as [s] than did the /f/-training group. The critical point is that lexical knowledge was available to disambiguate the speech sounds only during exposure but not at test. The observed shift in categorization at test thus reflects perceptual learning, rather than an immediate effect of lexical knowledge. Listeners had adjusted the boundaries of their phonetic categories to include the ambiguous sound in the category intended by the speaker. At test, listeners therefore perceived more sounds as belonging to that category. Over the past decade, the finding that lexical knowledge can retune phonetic boundaries has been replicated extensively (for an overview see Samuel & Kraljic, 2009), and lexical knowledge has also been observed to guide the retuning of visual phoneme categories to talker idiosyncrasies (van der Zande, Jesse, & Cutler, 2013). Lexically guided retuning is one powerful mechanism that listeners can use to adjust to individual talkers but also to the idiosyncrasies shared by a group of talkers (e.g., Eisner, Melinger, & Weber, 2013).

Lexical knowledge needs, however, to become sufficiently available in time to guide phonetic retuning. The vast majority of studies have thus focused on retuning to idiosyncrasies in late (e.g., Clarke-Davidson, Luce, & Sawusch, 2008; Drouin, Theodore, &

This article was published Online First December 19, 2019.
The author thanks Charles Clifton for constructive comments on a draft of the article.
Correspondence concerning this article should be addressed to Alexandra Jesse, Department of Psychological and Brain Sciences, University of Massachusetts Amherst, 135 Hicks Way, Amherst, MA 01003. E-mail: ajesse@psych.umass.edu

Myers, 2016; Kraljic & Samuel, 2005, 2006) or even final (e.g., Eisner & McQueen, 2005; McQueen, Cutler, & Norris, 2006; McQueen, Norris, & Cutler, 2006; Norris et al., 2003; Sjerps & McQueen, 2010; van der Zande et al., 2013; van Linden & Vroomen, 2007) positions within words. In these cases, lexical knowledge is sufficiently available to disambiguate the sound when it is presented or shortly thereafter. For idiosyncrasies occurring earlier in a word, however, lexical knowledge can become available too late to guide retuning. Jesse and McQueen (2011) failed to observe retuning to idiosyncrasies in onsets of longer words that did not become lexically unique until later (see also McAuliffe & Babel, 2016). When the same ambiguous sound was spliced into word-final position, retuning occurred, showing that the failure to learn about the idiosyncrasy in onset position was not a general failure to learn about the idiosyncrasy.

The failure of lexical knowledge to guide phonetic retuning to idiosyncrasies in word onsets is problematic for listeners given that the production of many sounds is position-specific. An idiosyncratic pronunciation that occurs only in word onsets could be difficult to learn about, because lexical knowledge may become available too late to guide retuning for that position. Learning about that sound realization in other positions would not help with the acoustically different versions of the sound in word onsets. However, listeners can learn about word-initial idiosyncrasies in other ways. Listeners can adjust their phonetic categories when disambiguating phonotactic information is available while the idiosyncrasy is encountered (Cutler, McQueen, Butterfield, & Norris, 2008). Likewise, being explicitly told beforehand about a word-initial idiosyncrasy can induce a shift in categorization (McAuliffe & Babel, 2016), though it is unclear whether this shift has a perceptual nature.

Based on their findings, Jesse and McQueen (2011) posited a timing hypothesis of phonetic retuning, stating that for phonetic retuning to be possible, disambiguating information has to become available within a certain time window relative to the occurrence of the ambiguous sound. A corollary of the timing hypothesis is the prediction that lexical information should also guide retuning to word-initial idiosyncrasies, if it becomes sufficiently available in time. The timing hypothesis was also formulated in acknowledgment of prior findings of phoneme perception showing that acoustic information is preserved for at least a few hundred milliseconds (McMurray, Tanenhaus, & Aslin, 2009) and that the sound's and the word's interpretation remain malleable by context for a about a second (Connine, Blasko, & Hall, 1991; Samuel, 2016; Szostak & Pitt, 2013). Lexical knowledge could hence guide retuning to word-onset idiosyncrasies when it becomes available quickly enough, which should be the case in words that have an early uniqueness point or when the preceding sentence facilitates lexical processing. That sentence context influences lexical processing has been well documented, though how and when its effects arise is still the focus of long-standing debate (for a review see, e.g., Kuperberg & Jaeger, 2016). Because semantic context, but also other contextual information, can modulate the early moments of phonological lexical competition (e.g., Brock & Nation, 2014; Dahan & Tanenhaus, 2004; Fox & Blumstein, 2016; Weber & Crocker, 2012), it could induce phonetic retuning by boosting lexical support for the intended word. Furthermore, semantic sentence context affects the processing of speech sounds. For instance, listeners are more likely to categorize steps of a phonetic continuum in line with the semantic content of the sentence (Abada, Baum, & Titone, 2008; Borsky, Tuller, & Shapiro, 1998; Connine, 1995; Connine et al., 1991; Garnes & Bond, 1976; Gow & Olson, 2016; Guediche, Salvata, & Blumstein, 2013; Miller, Green, & Schermer, 1984) and are more likely to miss that an onset of a word has been replaced by noise if that word is predictable from sentence context (Samuel, 1981). Together, these results suggest that sentence content could also guide phonetic retuning.

The present study tested whether the semantic content of a preceding sentence is sufficient to induce phonetic retuning. During an exposure phase, participants listened passively to sentences. On critical trials, a sound ambiguous between /s/ and /f/ replaced the fricative in minimal pairs (e.g., *sin* vs. *fin*). Sentence content disambiguated these minimal pairs as /s/ for one group of listeners (/s/-group) and as /f/ for another group (/f/-group). We predicted that the disambiguation of the minimal pair, and hence of the ambiguous sound, by the semantic content of the sentence would lead to retuning. Participants in the /s/-group should categorize steps more often as /s/ than should participants in the /f/-group. Control experiments subsequently tested whether retuning was indeed elicited by sentence context.

Experiment 1

Method

Participants. Forty monolingual American English native speakers with no reported hearing, language, or attention deficit participated (mean age = 19.8 years; nine men). Half of the participants were randomly assigned to each condition. Data from one additional participant were excluded from all analyses for not showing a continuum; that is, the person gave 0% (or 100%) [s] responses at both most extreme steps or provided more [s] responses at the most /f/-like step than at the most /s/-like step. Thirty-six additional participants from the same population contributed to pilot studies.

Materials. Sentences with highly constraining content that allowed predicting the word in the final position (see the Appendix for all critical sentence materials) were created. Critical sentences ended in each member of the /s/-/f/-initial minimal pairs (e.g., *Tim could not locate the car key, though the whole time it had been in plain sight* vs. *Laura had been arguing all day, and Tom did not want the argument to turn into a real fight*). Except for the onsets of these critical words, none of these sentences contained /s/ or /f/. Furthermore, they also did not contain /z/ or /v/. Predictability was achieved in variety of ways, such as with idiom completion (e.g., *her only claim to fame*) and semantic priming (e.g., *Brenda had a daughter and a son*). Syntactic information was allowed to contribute. A printed version of each sentence, with an underscore replacing the final word to indicate a gap, was presented to a total of 26 pilot participants, who completed the sentence with the first word coming to mind. Cloze probability was defined as the percentage of times participants continued the sentence with the intended word. The sentences of 20 critical minimal pairs with the highest cloze probability (/s/ versions: $M = 93.75\%$, $SD = 9.72$; /f/ versions: $M = 92.25\%$, $SD = 7.69$), $t(19) = 0.47$, $p = .64$, Cohen's $d = 0.17$, were selected, as well as 60 filler sentences ($M = 95.08\%$, $SD = 8.46$) that matched the critical sentences in cloze probability, $t(98) = 1.2$, $p = .24$, Cohen's $d = 0.24$. Ten percent of the critical words (i.e., two pairs) and of the filler words (i.e., six fillers) were bisyllabic; all others were monosyllabic. The spoken lexical frequency (Davies, 2008) of /s/ words (M = 42,817) and /f/ words (M = 48,969) were matched to each other, $t(19) = 0.31$, $p = .76$, Cohen's $d = 0.1$, and to that of the filler words (M = 41,791), $t(98) = 0.31$, $p = .76$, Cohen's $d = 0.062$.

A female American English speaker was recorded at 48 kHz with a 16-bit sampling rate producing the syllables /sɑ/ and /fɑ/ and the selected sentences. Sentences with critical words were also recorded with /h/ replacing the fricative (e.g., *hight* instead of *sight*/*fight*). To minimize coarticulatory information, these versions served as carriers, into which the ambiguous sound between /s/ and /f/ was later spliced. A continuum between the syllables /sa/ and /fa/ was created. The selected tokens had fricatives with a typical duration (/s/: 205 ms, /f/: 194 ms; see Jongman, Wayland, & Wong, 2000). The vowels of these two tokens were matched in duration (312 ms vs. 318 ms). Both tokens were resynthesized using the PSOLA algorithm in Praat (Boersma & Weenink, 2016) such that their vowels were set to their mean peak amplitude (77 dB), before the fricative portions were excised to be mixed together in complementary intensity ratios to create a 42-step continuum. To create consonant−vowel syllables, the steps were concatenated with a 318-ms-long vowel of a recording of /hɑ/. Formant transitions were minimal in this vowel. This vowel was rescaled to 77 dB and ramped over its initial 8 ms and its final 10 ms.

The two endpoints (0, 41) and six intermediate steps (10−15) were chosen based on the results of a pilot study, in which 10 participants categorized 17 of the 41 continuum steps 12 times. Step 14 also served as the ambiguous exposure sound, because it was the step closest to eliciting 50% (i.e., 40.28%) of [s] responses in the pilot study. The same ambiguous Step 14 was spliced to replace /h/ in all critical carrier sentences, after these had been rescaled such that the mean amplitude ratio between /h/ and the rest of the critical words was the same as the ratio between the fricative in Step 14 and its vowel. The critical ambiguous phoneme was therefore acoustically identical in all target stimuli in both exposure conditions. All critical and filler sentences were then normalized in their intensity.

Procedure. The procedures for all experiments reported here followed a protocol previously approved by the UMass Institutional Review Board. Participants were tested individually in sound-attenuated booths. Stimuli were played at a comfortable, fixed listening level over Sennheiser HD 280 PRO headphones. Participants received instructions for both the exposure and test phases of the experiment at the beginning, and no specific sound examples were mentioned. More detailed written instructions about the tasks were then given right before each phase. The experiment lasted for approximately 20 min.

During the exposure phase, participants listened passively to one sentence per trial, before pressing any button to continue with the next trial. Half of the participants heard 20 critical sentences disambiguating the final word as /s/-initial and did not hear any sentences with /f/ items. The other half heard 20 sentences disambiguating the critical words as /f/-initial and did not hear any sentences ending in /s/ items. Both groups listened to the same 60 filler trials that did not contain /s/, /f/, /z/, or /v/. Presentation order of all 80 sentences was randomized for each participant.

Upon completing the exposure phase, participants pressed a button to start the test phase on their own. During the test phase, they categorized the seven steps of the /sɑ/-/fɑ/ continuum six times. Participants first saw a fixation cross for 250 ms, followed by a black screen for 50 ms before stimulus presentation. Once response labels (<suh>, <fuh>) appeared with the offset of stimulus presentation, participants had 3 s to respond as accurately and as quickly as possible by button press. The next trial started 500 ms after the response deadline or the button press. Presentation order was blocked by continuum repetition, but steps were presented in a random order within each repetition.

Results and Discussion

Figure 1 shows the results of the phonetic categorization task by exposure condition. The phonetic categorization data was analyzed with a generalized mixed-effects regression model with a binomial linking function, implemented in R (R Core Team, 2014) using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015). The p values were estimated using Satterthwaite approximation for degrees of freedoms. Exposure condition was included as a contrast-coded fixed factor (/s/-exposure = −0.5, /f/-exposure = 0.5) and step as a scaled continuous numerical factor. Models included subjects as a random factor and by-step slope adjustments.

Participants showed a significant effect of exposure (Exposure: $\beta = -2.26$, $SE = 0.76$, $z = -2.96$, $p < .01$), in that participants who had heard the ambiguous sound in sentences disambiguated as /s/ gave more [s] responses at test than did participants who had heard the same sound before in sentences disambiguated as /f/. This result shows that the disambiguation of a sound through preceding sentence context can induce phonetic retuning of the involved phonetic categories. The size of this retuning effect was

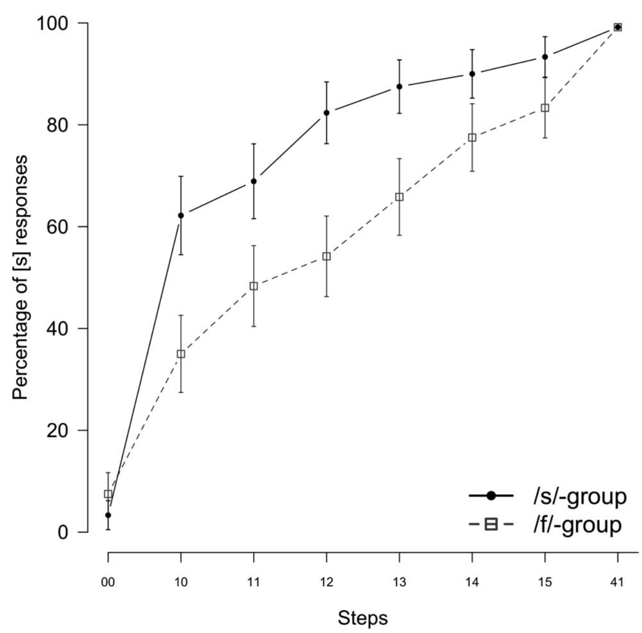

Figure 1. Percentage of [s] responses at test as a function of continuum step and exposure condition in Experiment 1. Error bars show the standard error of the means.

modulated by step (Exposure × Step: $\beta = -1.12$, $SE = 0.49$, $z = -2.29$, $p < .05$). Overall, participants were sensitive to the continuum and gave more [s] responses for more /s/-like steps of the continuum (Step: $\beta = 2.70$, $SE = 0.27$, $z = 10.2$, $p < .00001$).

Experiment 2

Experiment 1 provides evidence that sentence context can guide phonetic retuning. An acoustically ambiguous phoneme placed in a lexical context where both of the phonemes' interpretations could form words (i.e., an ambiguous sound replacing the fricative in, e.g., *sin/fin*) was disambiguated only by preceding sentence context. The lexical context itself was not informative. Only sentence context therefore informed listeners about how the ambiguous speech sound should be interpreted. Listeners adjusted their phonetic categories accordingly, such that, at test, when disambiguating information was no longer made available, listeners categorized the ambiguous phoneme, and its acoustically similar variants, more often as intended by the speaker, indicating phonetic retuning. Preceding sentence context can thus guide the retuning of phonetic categories.

In the retuning paradigm used in Experiment 1, listeners never heard a clear version of the other possible interpretation of the ambiguous phoneme during exposure. Participants who heard the ambiguous sound replace /s/ did not hear an unambiguous /f/ and vice versa. This retuning paradigm hence differs from the lexically guided phonetic retuning paradigm, where participants traditionally hear both the ambiguous sound and a clear unambiguous version of its other interpretation (e.g., listeners hear the ambiguous sound replace /s/ and also hear a clear /f/). Because listeners heard only one of the critical fricatives and never the contrastive other one during exposure in Experiment 1, it is possible that they increased the probability of expecting the respective critical fricative. The aftereffects found in Experiment 1 could thus reflect the perceptual learning of probabilities rather than phonetic retuning guided by sentence context. To rule out this possibility, participants in Experiment 2 heard equivalent numbers of /s/ and /f/ items. That is, participants who heard the sentence context disambiguate the ambiguous sound replacing the fricative in a minimal pair as /s/ also received sentences ending in words with a clear /f/. Participants who heard the sentence context disambiguate the ambiguous sound as /f/ also received sentences ending in words with a clear /s/. If the aftereffects in Experiment 1 reflect probabilistic learning instead of phonetic retuning, then aftereffects should disappear in Experiment 2 when exposure to /s/- and /f/-items has been equated. In contrast, if the aftereffects in Experiment 1 indeed reflect phonetic retuning, then they should persist in Experiment 2.

Method

Participants. Forty monolingual American English native speakers with no reported hearing, language, or attention deficit participated (mean age = 20.05 years; nine men). Half of them were randomly assigned to each exposure condition. None of them had participated in Experiment 1.

Materials. The same materials as in Experiment 1 were used, except for 20 s/f contrast filler sentences replacing 20 of the previous filler sentences. These contrast filler sentences contained 20 /s/-initial or 20 /f/-initial words that had been recorded together with the other materials by the speaker used in Experiment 1. Half of each type of words formed a minimal pair if the fricative were to be replaced with the respective other fricative (e.g., *fake, slaw*); the other half did not (e.g., *form, silk*). However, the other word in a minimal pair was never presented. None of these sentences contained /z/, /v/, or /s/ and /f/ anywhere except for the onsets of these contrast filler words. Just as was the case for the critical items and the other fillers, 10% of each contrast filler type were bisyllabic words. The cloze probability (piloted along with the sentences for Experiment 1) was matched across contrast filler sentence types (/s/-sentences: $M = 92.17\%$, $SD = 10.87$; /f/-sentences: $M = 94.16\%$, $SD = 9.86$), $t(38) = 0.61$, $p = .55$, Cohen's $d = 0.2$, and overall to that of the critical sentences, $t(80) = 0.26$, $p = .79$, Cohen's $d = 0.06$, and to the fillers they were replacing: /s/-sentences: $t(38) = 0.23$, $p = .82$, Cohen's $d = 0.07$; /f/-sentences: $t(38) = 0.43$, $p = .67$, Cohen's $d = 0.14$. The contrast filler sentences were also matched in spoken lexical frequency ($M = 36{,}249$) to the critical words, $t(69.24) = 0.8$, $p = .43$, Cohen's $d = 0.18$, and to the fillers they were replacing: /s/-sentences: $t(38) = 1.3$, $p = .2$, Cohen's $d = 0.41$; /f/-sentences: $t(38) = 0.85$, $p = .4$, Cohen's $d = 0.26$. The contrast filler sentences had been normalized along with all other sentences in intensity.

Procedure. The procedure and design were the same as in Experiment 1, except that the contrast fillers replaced 20 of the previously used fillers. Participants who heard the ambiguous sound replace /s/ in the critical words received the /f/-endpoint contrast fillers; participants who heard the ambiguous sound replace the /f/ in the critical words received the /s/-endpoint contrast fillers.

Results and Discussion

Figure 2 shows the results of the phonetic categorization task by exposure condition. Participants showed a significant effect of exposure (Exposure: $\beta = -1.38$, $SE = 0.59$, $z = -2.34$, $p = .02$), indicating phonetic retuning. Participants who had heard the ambiguous sound in sentences disambiguated as /s/ gave more [s] responses at test than did participants who had heard the same sound before in sentences disambiguated as /f/. The size of this aftereffect was not modulated by step (Exposure × Step: $\beta = -0.30$, $SE = 0.36$, $z = -0.83$, $p = .4$), but participants gave overall more [s] responses for more /s/-like steps of the continuum (Step: $\beta = 2.42$, $SE = 0.20$, $z = 12.11$, $p < .00001$). Experiment 2 thus replicates the results of Experiment 1, in that sentence context elicited phonetic retuning. The aftereffects in Experiment 1 (and in Experiment 2) therefore are likely to reflect phonetic retuning rather than probabilistic learning.

Though probabilistic learning did not induce the observed aftereffects in Experiment 1, it could have contributed to these aftereffects of sentence-guided retuning. If that were the case, then aftereffects should be larger in Experiment 1 than in Experiment 2. On the other hand, the unambiguous contrastive phonemes presented during exposure in Experiment 2 could have served as perceptual anchors that provided additional help to the listeners in retuning their phonetic categories. Prior work has shown that aftereffects of lexically guided phonetic retuning are increased by the presence of contrastive phonemes during exposure (van Linden

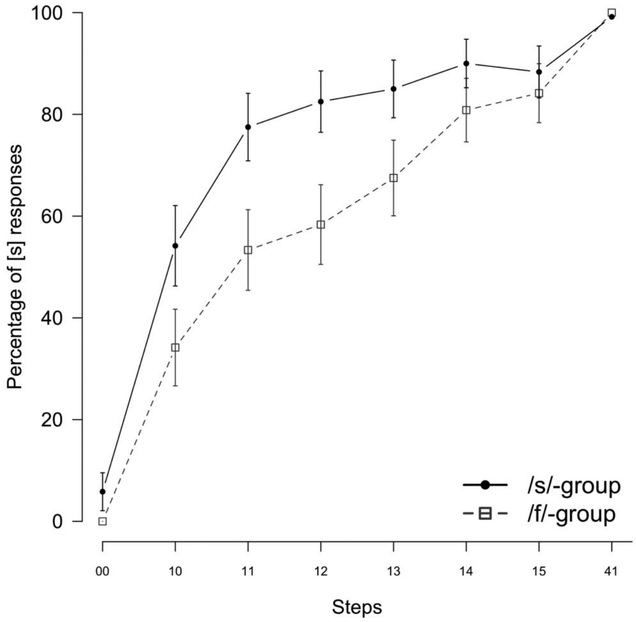

Figure 2. Percentage of [s] responses at test as a function of continuum step and exposure condition in Experiment 2. Error bars show the standard error of the means.

& Vroomen, 2007). If contrastive phonemes took on a similar role here as well, then aftereffects should be larger in Experiment 2 than in Experiment 1.

To test the role of contrastive phonemes in phonetic retuning guided by sentence context, we conducted a cross-experiment statistical comparison with experiment added to the model as a contrast-coded fixed factor (with endpoints [Experiment 2] = −0.5; without endpoints [Experiment 1] = 0.5). Results showed that the size of the retuning effect did not differ across experiments (Exposure × Experiment: $\beta = -0.03$, $SE = 0.26$, $z = -0.11$, $p = .91$). Whether or not contrastive phonemes were provided during exposure therefore did not change phonetic retuning. The triple interaction of exposure, experiment, and step was also not significant (Exposure × Experiment × Step: $\beta = 0.40$, $SE = 0.25$, $z = -1.55$, $p = .12$). Rather, only an overall effect of exposure was found (Exposure: $\beta = -1.46$, $SE = 0.17$, $z = -8.65$, $p < .00001$), indicating recalibration of auditory phonetic categories through sentence context. This effect changed across step (Exposure × Step: $\beta = -0.41$, $SE = 0.15$, $z = -2.81$, $p < .005$). Also, overall, participants gave more [s] responses for more /s/-like steps (Step: $\beta = 2.07$, $SE = 0.13$, $z = 16.28$, $p < .00001$). The main effect of experiment and its interaction with step were not significant (Experiment: $\beta = -0.03$, $SE = 0.11$, $z = 0.23$, $p = .81$; Experiment × Step: $\beta = -0.01$, $SE = 0.11$, $z = 0.08$, $p = .93$).

In summary, these results provide evidence that preceding sentence context guides phonetic retuning, independently of whether or not listeners heard contrastive phonemes during exposure. Hearing contrastive phonemes is neither necessary nor beneficial.

Experiment 3

Both Experiments 1 and 2 show that the disambiguation of speech sounds through sentence context can lead to phonetic retuning. Whereas listeners in Experiment 1 did not receive contrastive phonemes during exposure, listeners in Experiment 2 did. Evidence of phonetic retuning was found in the aftereffects of both experiments, and these aftereffects were of similar size. The aftereffects in Experiment 1 had therefore not been observed merely because of the absence of contrastive phonemes; that is, aftereffects do not reflect an upweighting of the probability of one of the respective fricatives in that study but rather indeed reflect phonetic retuning. Furthermore, hearing contrastive phonemes was neither necessary nor able to enhance phonetic retuning. Rather, whether listeners heard contrastive phonemes or not had no influence on aftereffects in this paradigm. Aftereffects in Experiment 1 and 2 were of the same size. The aftereffects observed in Experiments 1 and 2 therefore reflect sentence-guided phonetic retuning.

In the sentence-induced phonetic retuning paradigm in Experiments 1 and 2, only sentence context could possibly disambiguate the identity of the critical ambiguous phonemes. The lexical item that carried the ambiguity itself did not provide any information, because it always formed a minimal pair with the other acoustically supported interpretation. To further test that it is indeed sentence context that guides phonetic retuning, the predictive sentences of Experiments 1 and 2 were replaced with neutral ones (*Click on the word _*) in Experiments 3a and 3b. If sentence context had guided phonetic retuning in Experiments 1 and 2, then aftereffects should no longer be observed in Experiments 3a and 3b, where sentences were made nonpredictive of the final word, thus not disambiguating the identity of the critical phonemes.

Hearing a contrastive phoneme has been shown not to be sufficient to produce the aftereffects of lexically guided phonetic retuning. In the lexically guided retuning paradigm in Norris et al.'s (2003) study, no shifts were found after exposure to ambiguous sounds in nonwords, though listeners had still heard the contrastive phonemes in words. To test whether contrastive phonemes can induce aftereffects in the absence of any disambiguating information in the present paradigm, listeners heard contrastive phonemes in Experiment 3b but not in 3a.

Method

Participants. Eighty monolingual native speakers of American English with no reported hearing, language, or attention deficit participated (mean age = 19.9 years; 16 men). A quarter of them were randomly assigned to each of two exposure conditions in each of the two experiments (3a, 3b). Data from three other participants were excluded from all analyses for not showing a continuum.

Materials and procedure. The same materials as in Experiment 1 were used in Experiment 3a; that is, no contrast phonemes were provided. However, the biasing sentences were replaced with the neutral sentence *Click on the word _* while keeping the original ambiguous sound and word from each biasing sentence. Exposure conditions now referred to whether the ambiguous sound (and word) had originally come from a sentence disambiguating its identity as /s/ or as /f/. Experiment 3b also used neutral sentences, but as in Experiment 2, contrast phonemes were provided in 20 of the fillers. The procedure for both experiments was the same as for Experiment 1 and 2.

Results and Discussion

Figure 3 shows the results of the phonetic categorization task at test by exposure condition for Experiments 3a and 3b. Without biasing sentences disambiguating the critical words, participants did not show a significant effect of exposure (Exposure: Experiment 3a: $\beta = 0.02$, $SE = 0.53$, $z = 0.04$, $p = .97$; Experiment 3b: $\beta = 0.21$, $SE = 1.22$, $z = 0.17$, $p = .86$). That is, participants showed no aftereffects indicating phonetic retuning. The size of this (null) effect was not modulated by step (Exposure × Step: Experiment 3a: $\beta = -0.04$, $SE = 0.16$, $z = -0.23$, $p = .82$; Experiment 3b: $\beta = 0.18$, $SE = 0.84$, $z = 0.21$, $p = .83$), but participants gave overall more [s] responses for more /s/-like steps of the continuum (Step: Experiment 3a: $\beta = 1.26$, $SE = 0.09$, $z = 14.33$, $p < .00001$; Experiment 3b: $\beta = 3.48$, $SE = 0.49$, $z = 7.06$, $p < .00001$).

Together, these two control experiments show that listeners cannot retune to the word-initial idiosyncrasies in these stimuli in the absence of disambiguating sentence context. Furthermore, neither the presence nor the absence of contrastive phonemes can induce aftereffects. The results of Experiments 3a and 3b therefore provide further evidence that the aftereffects observed in Experiments 1 and 2 were due to retuning guided by constraining sentence context.

A cross-experiment comparison with experiment as a contrast-coded fixed factor (with endpoints [Experiment 3b] = −0.5; without endpoints [Experiment 3a] = 0.5) also provided no evidence of phonetic retuning (Exposure: $\beta = -0.06$, $SE = 0.13$, $z = -0.43$, $p = .67$). None of the interactions including exposure were significant (Exposure × Step: $\beta = -0.12$, $SE = 0.14$, $z = -0.88$, $p = .38$; Exposure × Experiment: $\beta = 0.77$, $SE = 0.41$, $z = 1.9$, $p = .06$; Exposure × Step × Experiment: $\beta = 0.41$, $SE = 0.35$, $z = 1.16$, $p = .25$). Categorizations differed, however, across experiments overall, in that when no endpoints had been presented during exposure, both exposure groups gave more [s] responses at test (Experiment: $\beta = 0.72$, $SE = 0.12$, $z = 6.19$, $p < .00001$). This difference across experiments was larger for more /s/-like steps (Experiment × Step: $\beta = 0.77$, $SE = 0.12$, $z = 6.21$, $p < .00001$). Overall, participants were sensitive to the continuum (Step: $\beta = 2.45$, $SE = 0.19$, $z = 12.82$, $p < .00001$).

General Discussion

The present study provides evidence that sentence context can guide the retuning of phonetic categories. In Experiment 1 and 2, the identity of an ambiguous phoneme in minimal pairs (e.g., *sight* vs. *fight*) was disambiguated by preceding sentence content as either /s/ or /f/ in a passive-listening task. Sentence constraints had been created in a variety of ways in the present materials (e.g., idiom completion, semantic priming). Lexical context itself was not informative. At a subsequent test, listeners in both experiments categorized more steps of a /sɑ/-/fɑ/-continuum in line with their prior exposure. Sentence context was thus sufficient to induce the retuning of phonetic categories. Only little exposure (i.e., 20 critical sentences) to the critical idiosyncrasy was needed. Two control experiments (3a and 3b) in which the same words were embedded in neutral sentence contexts did not show evidence of retuning, further supporting the notion that the sentence context had guided listeners in their adjustments to the speaker in Experiments 1 and 2. The absence or presence of contrastive phonemes during exposure did not modulate (Experiments 1 and 2) or induce aftereffects (Experiments 3a and 3b). Taken together, the results show that listeners can obtain sufficient information from sentence context on how to retune their phonetic categories to accommodate a speaker.

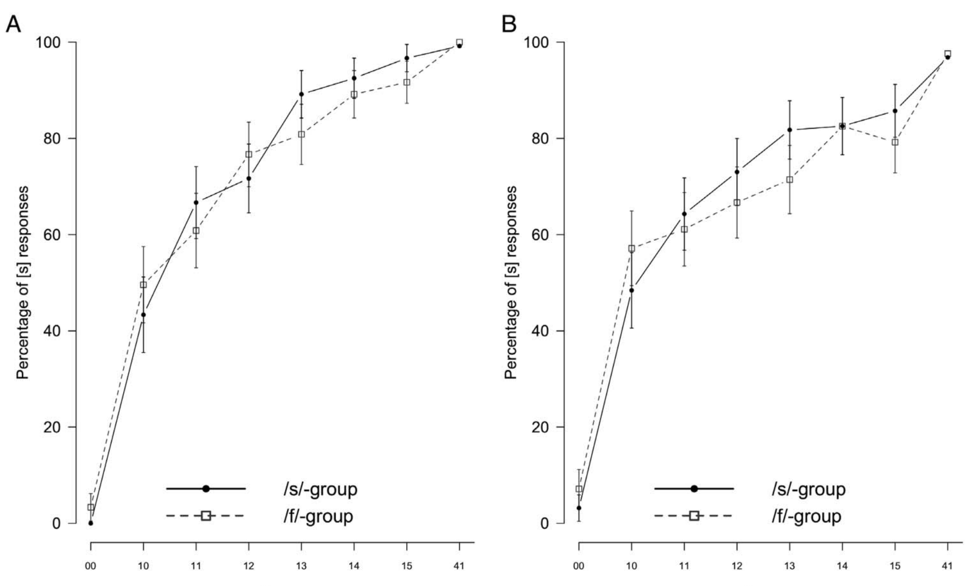

Figure 3. Percentage of [s] responses at test as a function of continuum step and exposure condition in Experiment 3a (without contrastive phonemes during exposure; Panel A) and 3b (with contrastive phonemes during exposure; Panel B). Error bars show the standard error of the means.

This study adds sentence context to the list of sources of information (lexical knowledge, visual speech information, imagery, phonotactics, and text) that listeners can use to guide their retuning of phonetic categories to accommodate talker idiosyncrasies (Baart & Vroomen, 2010; Bertelson, Vroomen, & de Gelder, 2003; Cutler et al., 2008; Keetels, Schakel, Bonte, & Vroomen, 2016; Norris et al., 2003; Scott, 2016; van der Zande et al., 2013). Sentence context provides a valuable addition to that list of listener tools because it can help overcome the limitations of lexically guided phonetic retuning. One such limitation is that for lexical information to guide phonetic retuning it must eliminate all possible interpretations of the idiosyncrasy except for one (e.g., only /s/ is lexically permissible following *platypu*). Lexical knowledge cannot guide retuning when several interpretations are lexically viable (e.g., *lice–life*). As shown in the present study, sentence context can help listeners in this situation. Furthermore, according to the timing hypothesis (Jesse & McQueen, 2011), for information to guide retuning, it must become sufficiently available as—or shortly after—the critical sound is presented. In line with this timing hypothesis, lexical information cannot guide phonetic retuning to idiosyncrasies in onsets of longer words that only later become lexically unique (Jesse & McQueen, 2011). Informative preceding sentence context could support lexically guided retuning when the availability of lexical information from the word itself is reduced or delayed. Sentence context therefore expands the utility of lexically guided retuning as a tool for the listener.

The mechanisms by which sentence content guides phonological processing are still largely unclear (e.g., Kuperberg & Jaeger, 2016). Likewise, it is unclear how sentence context guides phonetic retuning. However, it seems reasonable to assume that sentence context does not retune phonetic categories directly, but rather that it affects lexical processing and therefore in return aids lexically guided retuning. Both interactive and feedforward models of spoken-word recognition can provide an account of how sentence content may guide retuning. In interactive accounts (McClelland & Elman, 1986), sentence context influences lexical activation, which in turn top-down influences activation at the prelexical level and thereby alters current prelexical processing and, through retuning, subsequent prelexical processing (Mirman, 2008; Mirman, McClelland, & Holt, 2006). In feedforward accounts (Massaro & Cohen, 1991; Norris & McQueen, 2008; Norris, McQueen, & Cutler, 2000), lexical and prelexical information are combined at the decision level. In these accounts, lexical information retunes prelexical representations over time (Norris et al., 2003). Sentential information would likewise influence decisions and thus guide offline retuning (Norris & McQueen, 2008; Norris, McQueen, & Cutler, 2016). The current study was not designed to test between these two accounts, but its results show that any account of phonetic retuning needs to explain how sentence context contributes. Further testing of how sentence context guides phonetic retuning could also provide insights into how lexical knowledge can influence lower level perceptual processing.

The present study provides evidence that contrastive phonemes during exposure are neither necessary nor sufficient for phonetic retuning through sentence context to occur and do not modulate the size of its aftereffects. Since the seminal study by Norris and colleagues (2003), listeners in the lexically guided phonetic retuning paradigm have been exposed to the ambiguous sound as well as to a clear version of the other possible interpretation of the ambiguous phoneme. For example, listeners heard unambiguous /f/-tokens if the ambiguous sound replaced all occurrences of /s/ but unambiguous /s/-tokens if the ambiguous sound replaced all occurrences of /f/. Our results for sentence-guided phonetic retuning are in line with prior work on lexically guided retuning (van Linden & Vroomen, 2007), in that the presence of contrastive phonemes does not seem to be necessary for phonetic retuning to occur. Recalibration guided by sentence context occurred even when no contrastive phonemes were presented during exposure. Contrastive phonemes are therefore not needed to enable phonetic retuning.

The results of the present study also expand on previous work on the role of contrastive phonemes in lexically guided retuning. First, the results demonstrate that contrastive phonemes were not sufficient to induce aftereffects. No aftereffects were found in the absence of constraining sentences in Experiment 3, independent of whether contrastive phonemes had been included during exposure. Second, the aftereffects found in Experiment 1 when sentence context constrained the identity of the critical sound during exposure did not depend on the absence of contrastive phonemes. Not including contrastive phonemes during exposure in Experiment 1 could, in theory, induce a bias as listeners learn that one of the two fricatives is more likely to occur, and therefore this bias could lead to a shift in categorization. However, such a bias explanation can be ruled out because aftereffects also occurred in Experiment 2 when contrastive phonemes were presented during exposure. The observed aftereffects in Experiments 1 and 2 are therefore not contingent on, or even due to, the presence of clear contrastive phonemes. Last, whether or not contrastive phonemes were presented during exposure did not modulate the size of the aftereffects of sentence-guided phonetic retuning. That is, hearing the contrastive phonemes did not enhance sentence-guided phonetic retuning. Contrastive phonemes can enhance the aftereffects of lexically guided phonetic retuning, but only before dissipation from repeated testing (within the first 18 trials) has occurred (van Linden & Vroomen, 2007). To test whether this could also be the case here, we compared statistically the size of the aftereffects for the first two repetitions (14 trials) across Experiments 1 and 2. No effect of the presence or absence of contrastive phoneme on the size of the aftereffects was found, however. Taking all of these results together, hearing contrastive phonemes was therefore neither necessary nor beneficial for phonetic retuning in this paradigm.

In conclusion, the findings of the present study provide strong evidence that sentence context can guide phonetic retuning. This retuning may be lexically mediated, in that sentence context helps constrain possible word candidates. Lexical knowledge can then disambiguate an idiosyncrasy and guide retuning. This help can become particularly important when the lexical information provided by the word alone cannot disambiguate the identity of the phoneme, or at least not in time to guide retuning. Alternatively, sentence context could also potentially guide phonetic retuning directly. Future work is needed to distinguish these possible underlying mechanisms. Either way, sentence context provides listeners with another powerful tool to adjust their phonetic categories to a speaker and thus facilitate speech perception.

References

Abada, S. H., Baum, S. R., & Titone, D. (2008). The effects of contextual strength on phonetic identification in younger and older listeners. *Experimental Aging Research, 34,* 232–250. http://dx.doi.org/10.1080/03610730802070183

Allen, J. S., Miller, J. L., & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *Journal of the Acoustical Society of America, 113,* 544–552. http://dx.doi.org/10.1121/1.1528172

Baart, M., & Vroomen, J. (2010). Do you see what you are hearing? Cross-modal effects of speech sounds on lipreading. *Neuroscience Letters, 471,* 100–103. http://dx.doi.org/10.1016/j.neulet.2010.01.019

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67,* 1–48. http://dx.doi.org/10.18637/jss.v067.i01

Bertelson, P., Vroomen, J., & De Gelder, B. (2003). Visual recalibration of auditory speech identification: A McGurk aftereffect. *Psychological Science, 14,* 592–597. http://dx.doi.org/10.1046/j.0956-7976.2003.psci_1470.x

Boersma, P., & Weenink, D. (2016). Praat: Doing phonetics by computer (Version 6. 0.19) [Computer software]. Retrieved from http://www.praat.org/

Borsky, S., Tuller, B., & Shapiro, L. P. (1998). "How to milk a coat": The effects of semantic and acoustic information on phoneme categorization. *Journal of the Acoustical Society of America, 103,* 2670–2676. http://dx.doi.org/10.1121/1.422787

Brock, J., & Nation, K. (2014). The hardest butter to button: Immediate context effects in spoken word identification. *Quarterly Journal of Experimental Psychology, 67,* 114–123. http://dx.doi.org/10.1080/17470218.2013.791331

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America, 116,* 3647–3658. http://dx.doi.org/10.1121/1.1815131

Clarke-Davidson, C. M., Luce, P. A., & Sawusch, J. R. (2008). Does perceptual learning in speech reflect changes in phonetic category representation or decision bias? *Perception & Psychophysics, 70,* 604–618. http://dx.doi.org/10.3758/PP.70.4.604

Clopper, C. G., & Pisoni, D. B. (2004). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics, 32,* 111–140. http://dx.doi.org/10.1016/S0095-4470(03)00009-3

Connine, C. (1995). Effects of sentence context and lexical knowledge in speech processing. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 281–294). Cambridge, MA: MIT Press.

Connine, C. M., Blasko, D. G., & Hall, M. (1991). Effects of subsequent sentence context in auditory word recognition: Temporal and linguistic constraints. *Journal of Memory and Language, 30,* 234–250. http://dx.doi.org/10.1016/0749-596X(91)90005-5

Cutler, A., McQueen, J. M., Butterfield, S., & Norris, D. (2008). Prelexically-driven perceptual retuning of phoneme boundaries. In J. Fletcher, D. Loakes, R. Goecke, D. Burnham, & M. Wagner (Eds.), *Proceedings of the 9th Annual Conference of the International Speech Communication Association* (*INTERSPEECH 2008*; p. 2056). Red Hook, NY: Curran.

Dahan, D., & Tanenhaus, M. K. (2004). Continuous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 498–513. http://dx.doi.org/10.1037/0278-7393.30.2.498

Davies, M. (2008). *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present.* Retrieved from https://www.english-corpora.org/coca/

Drouin, J. R., Theodore, R. M., & Myers, E. B. (2016). Lexically guided perceptual tuning of internal phonetic category structure. *Journal of the Acoustical Society of America, 140,* EL307–EL313. http://dx.doi.org/10.1121/1.4964468

Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics, 67,* 224–238. http://dx.doi.org/10.3758/BF03206487

Eisner, F., Melinger, A., & Weber, A. (2013). Constraints on the transfer of perceptual learning in accented speech. *Frontiers in Psychology, 4,* 148. http://dx.doi.org/10.3389/fpsyg.2013.00148

Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics, 34,* 409–438. http://dx.doi.org/10.1016/j.wocn.2005.08.002

Fox, N. P., & Blumstein, S. E. (2016). Top-down effects of syntactic sentential context on phonetic processing. *Journal of Experimental Psychology: Human Perception and Performance, 42,* 730–741. http://dx.doi.org/10.1037/a0039965

Garnes, S., & Bond, Z. S. (1976). The relationship between semantic expectation and acoustic information. *Phonologica, 3,* 285–293.

Gow, D. W., Jr., & Olson, B. B. (2016). Sentential influences on acoustic-phonetic processing: A Granger causality analysis of multimodal imaging data. *Language, Cognition and Neuroscience, 31,* 841–855. http://dx.doi.org/10.1080/23273798.2015.1029498

Guediche, S., Salvata, C., & Blumstein, S. E. (2013). Temporal cortex reflects effects of sentence context on phonetic processing. *Journal of Cognitive Neuroscience, 25,* 706–718. http://dx.doi.org/10.1162/jocn_a_00351

Heald, S. L. M., & Nusbaum, H. C. (2014). Talker variability in audio-visual speech perception. *Frontiers in Psychology, 5,* 698.

Jesse, A., & McQueen, J. M. (2011). Positional effects in the lexical retuning of speech perception. *Psychonomic Bulletin & Review, 18,* 943–950. http://dx.doi.org/10.3758/s13423-011-0129-2

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America, 108,* 1252–1263. http://dx.doi.org/10.1121/1.1288413

Keetels, M., Schakel, L., Bonte, M., & Vroomen, J. (2016). Phonetic recalibration of speech by text. *Attention, Perception, & Psychophysics, 78,* 938–945. http://dx.doi.org/10.3758/s13414-015-1034-y

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology, 51,* 141–178. http://dx.doi.org/10.1016/j.cogpsych.2005.05.001

Kraljic, T., & Samuel, A. G. (2006). Generalization in perceptual learning for speech. *Psychonomic Bulletin & Review, 13,* 262–268. http://dx.doi.org/10.3758/BF03193841

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience, 31,* 32–59. http://dx.doi.org/10.1080/23273798.2015.1102299

Laver, J., & Trudgill, P. (1979). Phonetic and linguistic markers in speech. In K. R. Scherer & H. Giles (Eds.), *Social markers in speech* (pp. 1–32). New York, NY: Cambridge University Press.

Massaro, D. W., & Cohen, M. M. (1991). Integration versus interactive activation: The joint influence of stimulus and context in perception. *Cognitive Psychology, 23,* 558–614. http://dx.doi.org/10.1016/0010-0285(91)90006-A

McAuliffe, M., & Babel, M. (2016). Stimulus-directed attention attenuates lexically-guided perceptual learning. *Journal of the Acoustical Society of America, 140,* 1727–1738. http://dx.doi.org/10.1121/1.4962529

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1–86. http://dx.doi.org/10.1016/0010-0285(86)90015-0

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. http://dx.doi.org/10.1016/j.jml.2008.07.002

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science, 30,* 1113–1126. http://dx.doi.org/10.1207/s15516709cog0000_79

McQueen, J. M., Norris, D., & Cutler, A. (2006). The dynamic nature of speech perception. *Language and Speech, 49,* 101–112. http://dx.doi.org/10.1177/00238309060490010601

Miller, J. L., Green, K., & Schermer, T. M. (1984). A distinction between the effects of sentential speaking rate and semantic congruity on word identification. *Perception & Psychophysics, 36,* 329–337. http://dx.doi.org/10.3758/BF03202785

Mirman, D. (2008). Mechanisms of semantic ambiguity resolution: Insights from speech perception. *Research on Language and Computation, 6*(3–4), 293–309. http://dx.doi.org/10.1007/s11168-008-9055-5

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. http://dx.doi.org/10.3758/BF03213909

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. http://dx.doi.org/10.1121/1.1348009

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review, 115,* 357–395. http://dx.doi.org/10.1037/0033-295X.115.2.357

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences, 23,* 299–325. http://dx.doi.org/10.1017/S0140525X00003241

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47,* 204–238. http://dx.doi.org/10.1016/S0010-0285(03)00006-9

Norris, D., McQueen, J. M., & Cutler, A. (2016). Prediction, Bayesian inference and feedback in speech recognition. *Language, Cognition and Neuroscience, 31,* 4–18. http://dx.doi.org/10.1080/23273798.2015.1081703

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. http://dx.doi.org/10.1121/1.1906875

R Core Team. (2014). R: A language and environment for statistical computing (Version 3.6.1) [Computer software]. Retrieved from https://www.R-project.org/

Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General, 110,* 474–494. http://dx.doi.org/10.1037/0096-3445.110.4.474

Samuel, A. G. (2016). Lexical representations are malleable for about one second: Evidence for the non-automaticity of perceptual recalibration. *Cognitive Psychology, 88,* 88–114. http://dx.doi.org/10.1016/j.cogpsych.2016.06.007

Samuel, A. G., & Kraljic, T. (2009). Perceptual learning for speech. *Attention, Perception, & Psychophysics, 71,* 1207–1218. http://dx.doi.org/10.3758/APP.71.6.1207

Scott, M. (2016). Speech imagery recalibrates speech-perception boundaries. *Attention, Perception, & Psychophysics, 78,* 1496–1511. http://dx.doi.org/10.3758/s13414-016-1087-6

Sjerps, M. J., & McQueen, J. M. (2010). The bounds on flexibility in speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 195–211. http://dx.doi.org/10.1037/a0016803

Szostak, C. M., & Pitt, M. A. (2013). The prolonged influence of subsequent context on spoken word recognition. *Attention, Perception, & Psychophysics, 75,* 1533–1546. http://dx.doi.org/10.3758/s13414-013-0492-3

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. http://dx.doi.org/10.1121/1.3106131

van der Zande, P., Jesse, A., & Cutler, A. (2013). Lexically guided retuning of visual phonetic categories. *Journal of the Acoustical Society of America, 134,* 562–571. http://dx.doi.org/10.1121/1.4807814

van Linden, S., & Vroomen, J. (2007). Recalibration of phonetic categories by lipread speech versus lexical information. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 1483–1494. http://dx.doi.org/10.1037/0096-1523.33.6.1483

Vroomen, J., van Linden, S., de Gelder, B., & Bertelson, P. (2007). Visual recalibration and selective adaptation in auditory-visual speech perception: Contrasting build-up courses. *Neuropsychologia, 45,* 572–577. http://dx.doi.org/10.1016/j.neuropsychologia.2006.01.031

Weber, A., & Crocker, M. W. (2012). On the nature of semantic constraints on lexical access. *Journal of Psycholinguistic Research, 41,* 195–214. http://dx.doi.org/10.1007/s10936-011-9184-0

Yakel, D. A., Rosenblum, L. D., & Fortier, M. A. (2000). Effects of talker variability on speechreading. *Perception & Psychophysics, 62,* 1405–1412. http://dx.doi.org/10.3758/BF03212142

Appendix

Critical and Contrast Filler Sentences

/f/-Biasing Sentences

Critical items. Your math teacher called: Your grade got lower and lower. Try harder; try not to fail.

That she had met Lady Gaga at the airport would be her only claim to fame.

The light had been red, when you crashed into my car. The crash had been entirely your fault.

The chicken grew hungry. Time to get them their homemade organic chicken feed.

I would like to go to the party without you. how would that make you feel?

She did not know whether or not he had been lying. All she had, had been a gut feeling.

Amy rarely went anywhere without a shoe on; she did not like showing her bare feet.

Her ankle got broken. She had tripped on a cable and fell.

Laura had been arguing all day, and Tom did not want the argument to turn into a real fight.

The company had a job opening they urgently needed to fill.

Nemo had a lucky fin.

He did not get jail time, but the judge had demanded he would pay a three hundred dollar fine.

Emma had gained weight lately and now her jacket had gotten too tight. It would no longer fit.

They were married now, but she had met him during an internship and had thought their relationship would be nothing more than a quick fling.

I know you do not like the plan. But calm down and go with the flow.

Early in the morning down by the ocean, one could not make out a thing, due to the weather being too foggy.

Tina wrapped up the partially eaten burrito in aluminum foil.

All day long, she had not been able to locate her key. But then her roommate called to tell her not to worry—it had been found.

Tara enjoyed the trip to the beach. At the beginning, she didn't like it there, but gradually she had more and more fun.

The short hair on an animal, like on a cat, can be called fur.

Contrast filler items. On the airplane on her way home, she worried about the landing during the entire flight.

When he took my picture at the party with the camera, I got temporarily blinded by the bright flash.

When the tiger approached him, he began to shake in fear.

Her dream had been to become a pilot. Growing up, she had wanted nothing more than to fly.

She called it a romantic comedy. He called it a chick flick.

He wanted the dirty water gone. He pulled the handle and the toilet would flush.

At the pool, he jumped in doing a belly flop.

When I had P.E., I really liked track and field.

To buy alcohol when you are under age, you do not show a real ID. You show a fake.

That alcohol can damage your health should be a well-known fact.

Do not look in the pocket in the back. Turn it around and look in the pocket in the front.

I do not need to hear a reply right now, but I do need to hear one in the near future.

Luckily she didn't break her whole hand. She only broke her ring finger.

He brought the bait he had prepared to the lake to catch fish.

He returned the beer to the bar tender. On top it had too much white foam.

To keep the lunch meat cool, she put it back into the fridge.

I do not want to eat the noodle dish with my hand. I'd like a fork.

Upon completing a law degree at Yale, Tom began working in New York at a large law firm.

Her apartment had been on the third floor.

To complete their work, the accountant at H&R Block will need your W2 form.

/s/-Biasing Sentences

Critical items. The boat could be pushed by the wind; it had a big billowing white sail.

The shepherd could tell the sheep apart. But to Tom, they all looked the same.

The dish didn't need more pepper, but it did need more salt.

The gardener wanted to grow a tomato plant. He made hole in the ground and planted a seed.

The lid on a jam jar should maintain an airtight seal.

In order to shine light down on your head, the light should be mounted in the room on the ceiling.

Worrying that the train would be crowded at night, he had booked a window seat.

One inmate would go home in the morning, but the other one would remain one more night in her cell.

Tim could not locate the car key, though the whole time it had been in plain sight.

To put the plant in better light, she put it right up to the pane, on the wooden window sill.

According to the church and the Bible, murder would be a mortal sin.

Turn right at the yield sign.

She got a chair and told him to sit.

(*Appendix continues*)

When Debby hurt her shoulder, the doctor recommended that she kept the arm in a sling.

There are children playing here, and a big hump in the road. When you take your car through, you need to go really slow.

The cardboard had been out in the rain. It had become all wet and soggy.

Dirt in the garden, where your plant will grow, can be called soil.

The eye can detect light. The ear can detect sound.

Brenda had a daughter and a son.

The polite way to begin a letter directed at a woman would be “Dear Madam.” The polite way to begin a letter directed at a man would be “Dear Sir.”

Contrast filler items. Prior to her death, she had been a holy woman helping the poor. Now that the church could credit her with a miracle, they would declare her to be a saint.

To make the black bean dish, you need to put them in water and let them soak.

To unwind, many people like to hang out naked in a hot room called a sauna.

You cannot tell anyone. Pledge not to a tell a soul.

When walking by the monkey cage, the elderly woman held on tight to her daughter, who pointed to a banana peel on the ground and warned her she could slip.

She did not want the letter that had come. She wrote on it “Return to sender.”

She had no dishwasher. She washed her plate and tea cup in the kitchen sink.

You might giggle when you are happy. You cry when you are sad.

Amy told the waiter she wanted barbecued meat with the cole slaw.

To attack the opponent, the knight unsheathed the sword.

When I had a cold, my mom would cook me homemade chicken soup.

A man could wear a kilt but probably would not wear a mini skirt.

I had been telling a good joke, but Jim did not crack a smile.

Do not talk like that again or I’ll make you wash your mouth out with soap.

He didn’t like La Land. He did not like it when they broke out in song.

In the cold, cold winter, we would not get any rain but we could get snow.

In the winter, the blackbird, the robin, the hummingbird and many more all migrate south.

Her grandpa kept on bragging that in Pacman, Pong, and Bejeweled he had the all-time high score.

To mail a letter, in the upper right corner, you’ll need to put on a stamp.

When they made that elegant kimono, they did not take linen, wool, or cotton, but only silk.

Received April 29, 2019
Revision received November 16, 2019
Accepted November 18, 2019 ■