

Prelexical Adjustments to Speaker Idiosyncrasies: Are they Position-specific?

Alexandra Jesse¹, James M. McQueen¹

¹Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

alexandra.jesse@mpi.nl, james.mcqueen@mpi.nl

Abstract

Listeners use lexical knowledge to adjust their prelexical representations of speech sounds in response to the idiosyncratic pronunciations of particular speakers. We used an exposure-test paradigm to investigate whether this type of perceptual learning transfers across syllabic positions. No significant learning effect was found in Experiment 1, where exposure sounds were onsets and test sounds were codas. Experiments 2-4 showed that there was no learning even when both exposure and test sounds were onsets. But a trend was found when exposure sounds were codas and test sounds were onsets (Experiment 5). This trend was smaller than the robust effect previously found for the coda-to-coda case. These findings suggest that knowledge about idiosyncratic pronunciations may be position specific: Knowledge about how a speaker produces sounds in one position, if it can be acquired at all, influences perception of sounds in that position more strongly than of sounds in another position.

Index Terms: perceptual learning, speaker adaptation, prelexical representations, speech perception

1. Introduction

Listeners adjust their prelexical representations of speech sounds to the peculiarities of the speech of particular speakers [1]. These adjustments can be guided by lexical knowledge. In an exposure-test paradigm, listeners made lexical decisions to a list of words and nonwords, and then categorized a test continuum of speech sounds. A fricative sound ‘/?’ that was ambiguous between /s/ and /f/ tended to be categorized as /f/ after exposure where it had replaced the final /f/ in Dutch words (e.g., *olij?*, from *olijf*, ‘olive’; *olijf* is not a Dutch word). But the same ambiguous sound tended to be heard as /s/ after exposure when it replaced word-final /s/ (e.g., *radij?*, from *radijs*, ‘radish’; *radijf* is not a Dutch word). This learning has two important properties. It is, at least under some circumstances, speaker specific [2,3,4]. Secondly, it seems to have a prelexical locus. Learning of speaker idiosyncrasies transfers to novel words [5]. Learning can only readily be applied to the perception of new words spoken by the same speaker if adjustments are made to representations at the prelexical level, but not if they are made at lexical or postlexical levels. Thus, although it is lexical knowledge that guides these adjustments, it appears that they are made to prelexical rather than lexical representations.

We investigated here whether these perceptual adjustments are position specific or generalize across positions within a syllable. Previous studies presented the ambiguous sounds in the same (coda or word-medial) position during both exposure and test [1-5], although an effect on test stimuli in coda position has been found when the position of the ambiguous sound during exposure was free to vary [6]. In Experiment 1, we tested whether there was lexically-guided learning when critical phonemes were in onset position during exposure and in coda position at test. The experiment was

modelled closely on previous experiments using the same exposure-test paradigm (e.g., [1]). In the exposure phase, two groups of listeners performed an auditory lexical decision task on a list of words and nonwords, including a critical set of /f/- and /s/-initial words. For one group, the initial phoneme in all the /f/-words was replaced with an ambiguous sound midway between /f/ and /s/. For the other group, the same ambiguous sound replaced the /s/ in the /s/-words. In the test phase, all listeners categorized sounds on an /ɛs/-/ɛf/ continuum that had been used previously [2].

If listeners can learn to adjust their fricative categories based on exposure to sounds in onset position, and transfer that learning to fricatives in coda position, then listeners in the first group should categorize more tokens on the test continuum as /f/ than listeners in the second group. We could also compare the size of any such effect with that observed with the same test continuum when exposure was to ambiguous sounds in coda position [2].

2. Experiment 1

2.1. Methods

2.1.1. Participants

Thirty-two native Dutch speakers from the MPI participant pool were paid for their participation.

2.1.2. Stimuli

Twenty Dutch /f/-initial and 20 Dutch /s/-initial words were selected so that sets were equated on word frequency [7], number of syllables, lexical stress pattern, and on the vowel following the initial fricative. /f/ and /s/ did not occur anywhere other than in the initial position of these words (e.g., *finaal*, ‘final’, and *sigaar*, ‘cigar’). Words in each set did not have lexical neighbours starting with the other fricative (e.g., *sinaal* and *figaar* are not Dutch words). An additional 60 filler words and 100 filler nonwords were selected that were matched on number of syllables with the critical words. Nonwords were phonotactically legal in Dutch. None of these fillers included the critical fricatives.

All stimuli were recorded by the same female native Dutch speaker as in [2]. The critical words were also recorded with an /x/ instead of their initial fricatives. The nonsense syllables /sø/, /fø/, and /xø/ were also recorded. The fricative sounds of the syllables /sø/ and /fø/ were equated in duration and then mixed to create a 21-step continuum [see 1 for details]. These continuum steps were spliced onto the /ø/ from a /xø/ syllable. A pilot study with this /sø/-/fø/ continuum showed that step 9 (with 49% f-responses) was the most ambiguous sound. The ambiguous exposure words were made by replacing the /x/ in the /x/-initial versions of the critical words with this ambiguous fricative.

2.1.3. Procedure

During the exposure phase, participants completed an auditory lexical decision task. For half of the participants (the *f-bias group*), the 20 natural /s/-words were presented with the 20 ambiguous /f/-words and all 160 fillers. For the other participants (the *s-bias group*), the /f/-words were natural, the /s/-words were ambiguous, and the fillers were the same. In the immediately following test phase, both groups categorized six times five steps of the /εs/-/εf/ continuum used in [2] (steps 12, 17, 20, 23, and 28 of a 41-step continuum).

Table 1. Lexical decision latencies (mean reaction times, in ms, from word offset).

Expt.	Natural		Ambiguous	
	f-words	s-words	f-words	s-words
1	210	228	239	303
2	219	221	227	313
3	342	226	226	442
4	185	286	339	321
5	81	129	266	179

Table 2. Lexical decision accuracy (percentage correct responses).

Expt.	Natural		Ambiguous	
	f-words	s-words	f-words	s-words
1	91%	94%	81%	74%
2	94%	96%	85%	72%
3	95%	94%	89%	75%
4	91%	82%	88%	82%
5	94%	97%	91%	88%

2.2. Results and discussion

One subject was excluded from further analysis since she rejected more than 50% of the ambiguous words in the lexical decision task. Tables 1 and 2 list, for the remaining subjects, the average response latencies for correct responses, measured from word offset, and the percentage of correct responses, for natural and ambiguous versions of the /f/- and /s/-words. Most of the ambiguous items were accepted as words, even though reaction times and error rates were slightly higher for them than for the natural tokens. Participants were also more likely to accept as words items where the ambiguous sound replaced /f/ than where it replaced /s/, as has been found before [1].

The phonetic categorization data are shown in Figure 1. Listeners were sensitive to the continuum manipulation ($F(4, 26)=10.29$; $p<.0001$). Responses changed similarly in both groups across the steps of the continuum ($F(4,26)=1.77$, $p=.17$). More importantly, there was no lexically-guided perceptual learning ($F(1,29)=.20$, $p=.66$). A comparison of the current results to those in [2] for coda-to-coda learning showed in addition that the size of the bias varied as a function of the position of the ambiguous phoneme during exposure ($F(1,56)=5.96$, $p<.02$).

These results might suggest that there is no transfer of learning of speaker idiosyncrasies across phoneme position. There is some indication, however, that this could be due to the acoustic dissimilarity between exposure and test sounds. Groups showed a tendency to label step 17 consistently with exposure ($t(29)=1.42$, $p=.08$). This step had the same proportional mix of /s/ and /f/ as the ambiguous sound used during exposure.

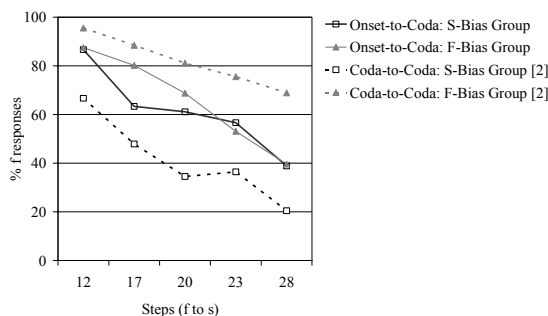


Figure 1: Onset exposure and coda test (Expt. 1), compared to coda exposure and coda test [2].

3. Experiment 2

The lack of a significant perceptual learning effect in Experiment 1 could be because there is no transfer of learning across syllabic positions. This would suggest that learning may be about position-specific allophones, rather than phonemes. Alternatively, however, the failure could be caused by the decrease in similarity between the ambiguous sound heard during exposure and the test continuum. In [1,2], for example, the ambiguous exposure sound was also the most ambiguous sound in the test continuum. The results do indeed show that acoustic similarity might influence the effect (see [3] for a similar argument). More fundamentally, it is possible that lexically-guided learning cannot be induced from sounds in onset position. If so, then one could not expect learning to transfer to another syllabic position. Experiment 2 addressed both of these concerns. We used the same exposure materials as in Experiment 1, but used a test continuum with fricatives in onset position. Furthermore, this continuum was acoustically matched to the exposure ambiguous sound. If lexically-guided learning can be induced from onset-position sounds, and is applied at least to acoustically similar test stimuli, listeners in the *f-bias group* should categorize more sounds as /f/ than listeners in the *s-bias group*.

3.1. Methods

Twenty-four native Dutch speakers from the same population as in Experiment 1 were tested. Materials and procedure were the same as in Experiment 1, except for the test continuum. Five steps of the onset /sø/-/fø/ continuum (steps 5, 7, 9 (the ambiguous sound), 11, 13) were used in the phonetic categorization task. A pretest consisting of the same task and materials as the test phase was run before the exposure phase.

3.2. Results and discussion

Three participants were excluded from the analysis since they rejected more than 50% of the ambiguous words. The lexical decision results for the remaining subjects replicate those from Experiment 1 (see Tables 1 and 2). The categorisation data are shown in Figure 2. Participants were sensitive to the continuum in the pretest ($F(4,16)=3.76$, $p<.05$) and the posttest ($F(4,16)=10.74$, $p<.0001$). Performance changed across the continuum similarly for the two groups in the posttest ($F(4,16)=.73$, $p=.59$). This difference was marginally significant in the pretest ($F(4,16)=2.83$, $p=.06$). Groups were matched in their perception before the experiment, as the

pretest results indicate ($F(1,19)=.59, p=.45$). However, there was no indication of learning ($F(1,19)=.10, p=.75$). There were more f-responses in the posttest than in the pretest for both the *s-bias group* and the *f-bias group* ($F(1,19)=15.92, p<.01$), but groups did not differ as a function of test ($F(1,19)=.16, p=.69$).

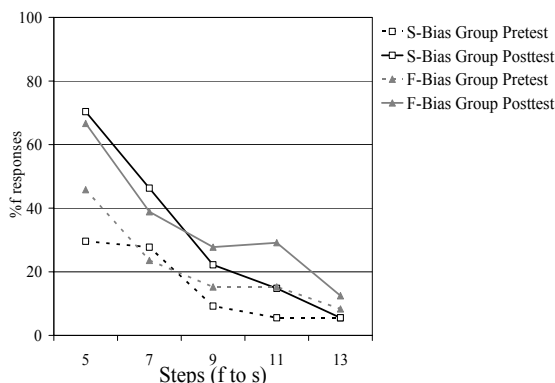


Figure 2: Onset exposure and onset test (Expt. 2).

4. Experiment 3

Experiment 2 failed to provide convincing evidence for onset induced learning. There may be a floor effect, however: particularly at pretest, listeners labelled most sounds on the test continuum as /s/. This may have obscured any lexically-guided learning effect. Experiment 3 therefore replicated Experiment 2, but with a more balanced continuum.

4.1. Methods

Ten native Dutch speakers participated in this study. This experiment was an exact replication of Experiment 2, except that the steps of the continuum were changed to steps 3, 5, 9 (ambiguous sound), 13, and 15.

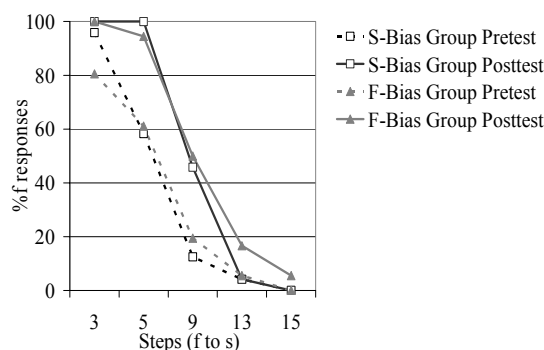


Figure 3: Onset exposure and onset test (Expt. 3).

4.2. Results and discussion

One participant was excluded from the analysis due to a high error rate (>50%) for words containing the ambiguous sound. The lexical decision results, at least those for accuracy, replicated the previous patterns (see Tables 1 and 2). The categorisation results (Figure 3) were similar to those of Experiment 2, in that both groups show a shift to more f-responses between pretest and posttest ($F(1,16)=12.85, p<.05$), but this does not vary as a function of group ($F(1,16)=.17, p=.68$). The continuum was, however, more

balanced (pretest: $F(4,5)=34.54, p<.0001$; posttest: $F(4,5)=148.50, p<.0001$) and performance across the continuum did not change as a function of group in pretest ($F(4,5)=.78; p=.58$) or in posttest ($F(4,5)=.47; p=.76$). There was no difference between groups in either pretest ($F(1,8)=.01; p=.93$) or posttest ($F(1,8)=.42; p=.54$). It seems therefore that no lexically-guided learning is induced from onset-positioned sounds.

5. Experiment 4

Experiments 1-3 suggest that sounds in onset position do not induce lexically-guided learning. In Experiments 2 and 3, however, there was a substantial shift between pretest and posttest towards more f-responses in both groups. This could have been a strategic reaction to the large /s/ bias during the pretest. In addition, the pretest introduced the listeners to a large spectrum of fricatives falling between /f/ and /s/ that could apparently be produced by the speaker. This exposure might have discouraged the listeners from learning anything about the speaker's fricatives. Experiment 4 was therefore an exact replication of Experiment 3, but without a pretest.

5.1. Methods

Twenty native Dutch speakers participated. The experiment was in all aspects identical to Experiment 3, except that it did not contain a pretest.

5.2. Results and discussion

One participant had to be excluded from the further analysis due to a high error rate (>50%) for words containing the ambiguous sound. The lexical decision results replicated the patterns found before (see Tables 1 and 2). One difference, however, is that /s/-words with an ambiguous sound were not processed more slowly or rejected more than /f/-words with the ambiguous sound. The categorization data are shown in Figure 4. The effect of continuum was again significant ($F(4,15)=37.14, p<.0001$) and did not vary as a function of group ($F(4,15)=2.08, p=.13$). These results replicate Experiments 2 and 3 in providing further evidence that onset-positioned ambiguous sounds do not lead to lexically-guided phoneme adjustments ($F(1,18)=.65, p=.43$), even for onset-positioned test fricatives.

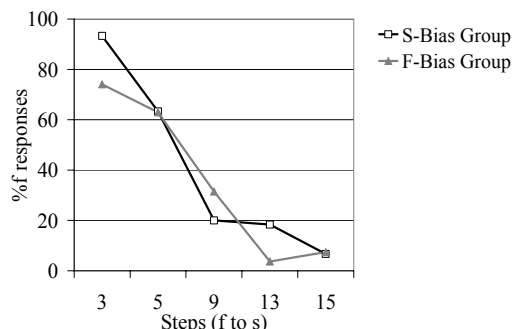


Figure 4: Onset exposure and onset test (Expt. 4).

6. Experiment 5

The previous four experiments failed to show an onset-induced perceptual adjustment. This result was not due to a lack of similarity between exposure and test sounds, nor the presence of a pretest, and it persisted whether we tested for transfer of learning across positions or tested for within-position learning. So it is still possible that learning transfers

across position, but that this cannot be shown when the ambiguous sound in exposure is in onset position – simply because there is no learning that could transfer under these conditions. In Experiment 5, therefore, the ambiguous exposure sound was spliced into coda position (in /s/- and /f/-final words). Previous studies [1,2] have shown robust lexically-guided learning from exposure sounds in coda position. If learning can generalize across positions then learning with the same exposure conditions should be found for test sounds in onset position.

6.1. Methods

Nineteen native Dutch speakers participated. The same onset continuum was used during test as in Experiments 3 and 4. In the lexical decision task, however, two new sets of twenty words with either /f/ or /s/ in coda position were selected (e.g., *olijf* and *radijs*). As before, these words had no word neighbours with the other fricative in coda position (e.g., *olijf* and *radijs* are nonwords). The sets were matched in word frequency, number of syllables, lexical stress pattern, and the immediate vowel context of the fricative to each other, and to the onset word sets used previously. They were recorded by the same speaker and prepared as in Experiment 1. Word and nonword fillers were the same as before. Note that the same ambiguous sound as in Experiments 1-4 was spliced into coda position (thus creating a better acoustic match between exposure and test stimuli). The procedure was identical to the earlier experiments.

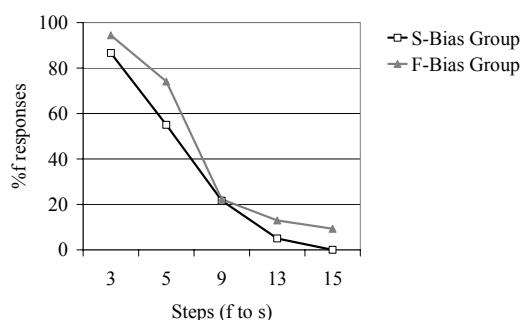


Figure 5: Coda exposure and onset test (Expt. 5).

6.2. Results and discussion

No subjects were excluded. The lexical decision results (see Tables 1 and 2) indicate that words with the ambiguous sound in coda position were still highly acceptable as words. Unlike in the previous experiments, participants took longer to accept the ambiguous sound as an /f/ than as an /s/, although error rates for both types of item were similar. This difference could suggest that the ambiguous sound fitted differently in coda position than in onset position, where it was slightly easier to accept as /f/ than as /s/. The categorization data are shown in Figure 5. Responses varied along the continuum ($F(4,14)=190.56$, $p<.0001$), but did not change as a function of group ($F(4,14)=1.11$, $p=.39$). Most interestingly, there is a trend for a difference between groups ($F(1,17)=2.22$, $p=.16$). This suggests that learning may weakly transfer across syllabic positions.

7. General Discussion

This series of experiments investigated whether lexically-guided learning of speaker idiosyncrasies can be transferred

across syllabic (onset, coda) positions. The first experiment failed to show transfer of learning from onset to coda position. Experiments 2-4 suggested, however, that this might be due to a failure to induce lexically-guided learning when listeners are exposed to ambiguous sounds in onset position. These onset-to-onset experiments showed no evidence of learning even though both exposure and test stimuli were in onset position, and the exposure and test stimuli were acoustically similar. The failure to find learning was replicated with different continuum steps, and both with and without a pretest. It may be that the influence of lexical knowledge on word-initial fricative perception is too weak to induce a perceptual learning effect.

Transfer across position from coda to onset was therefore examined in Experiment 5. The trend observed in this experiment contrasts with previous clear demonstrations of coda-to-coda learning (e.g., [1,2]). This suggests that if there is transfer of learning, then it is reduced when adjustments are applied to new syllabic positions. One potential drawback of the current study was that the ambiguous sound was recorded in onset position but then spliced into coda position. It might have been the case that the sound was not sufficiently ambiguous. An indication for this is that the ambiguous /s/-words were accepted more readily as words than the ambiguous /f/-words, while the opposite was the case when the ambiguous sound was in onset position. Further research is required to establish how appropriate the ambiguous sound was as a coda consonant, and, more generally, whether the trend observed in Experiment 5 is robust.

We draw two conclusions from the present investigation. First, adjustments to talker idiosyncrasies based on lexical knowledge do not appear to be made for sounds in onset position. Second, if transfer of this type of perceptual learning across positions is possible at all, then it is certainly reduced compared to the within-position effect. This suggests that learning in this paradigm, at least for the fricatives studied here, may concern position-specific allophones rather than more abstract phonemes.

8. References

- [1] Norris, D., McQueen, J.M., and Cutler, A. "Perceptual learning in speech", *Cog. Psych.*, Vol. 47, 2003, pp 204-238.
- [2] Eisner, F. and McQueen, J.M., "The specificity of perceptual learning in speech processing", *Perc. & Psychophys.*, Vol. 67, 2005, pp 224-238.
- [3] Kraljic, T. and Samuel, A.G. "Perceptual learning in speech: Is there a return to normal?", *Cog. Psych.*, Vol. 51, 2005, pp 141-178.
- [4] Kraljic, T. and Samuel, A.G. "Generalization in perceptual learning for speech", *Psych. Bull. & Rev.*, Vol. 13, 2006, pp 262-268.
- [5] McQueen, J.M., Norris, D., and Cutler, A., "Phonological abstraction in the mental lexicon", *Cog. Sci.*, Vol. 30, 2006, pp 1113-1126.
- [6] Eisner, F., and McQueen, J.M. "Perceptual learning in speech: Stability over time", *J. Acoust. Soc. Amer.*, Vol. 119, 2006, pp 1950-1953.
- [7] Baayen, R.H., Piepenbrock, R., and Gulikers, L "The CELEX Lexical Database (Release 2)" [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [Distributor], 1995.