Orienting attention during phonetic training facilitates learning

Eric Pederson and Susan Guion-Anderson
Department of Linguistics, University of Oregon, Eugene, OR 97403-1290
epederso@cas.uoregon.edu, guion@uoregon.edu

Abstract: The role of consciously-directed attention toward speech input in learning has not yet been determined. Previous phonetic learning studies have manipulated acoustic signals and response feedback, but not conscious control over attentional orienting. This study tests whether directed attention facilitates learning of novel phonetic information. Two monolingual English-speaking groups were trained with feedback on the same auditory stimuli: Hindi words. One group was instructed to attend to the consonants and the other to the vowels. The consonant-oriented group, but not the vowel-oriented group, demonstrated post-training improvement in consonant perception, confirming a role for consciously-directed attentional mechanisms during phonetic learning.

PACS numbers: 43.71.Hw, 43.71.Ft, 43.71.Es

1. Introduction

The purpose of the current study is to explore the role of attention in the learning of novel phonetic categories by adults. It is well known that listeners have difficulties with some non-native phonetic contrasts due to prior native-language experience. Presumably listeners have learned to allocate their attentional resources to best fit the particulars of their native language (Strange and Shafer, 2008). This acquired processing can lead to interference in learning novel phonetic contrasts. For example, Iverson and Kuhl (1995) propose that linguistic experience warps perceptual space in that, with acoustically equidistant tokens, those categorized as good exemplars of a phonetic category cluster closely together in perceptual space, while poorer exemplars are more distant perceptually. Such warping of the perceptual space is thought to affect perceptual processing such that the relative weighting of specific acoustic cues is affected by linguistic experience. That is, perceptual sensitivity along an acoustic dimension is affected by the characteristics of one’s native language (Iverson et al., 2003).

It is widely believed that attention plays an important role in phonetic learning. That is, accurate perception of at least some non-native phonemes requires reorienting of attention. For example, Pisoni and colleagues have demonstrated that laboratory training can effectively modify the ability to discriminate non-native phonetic categories and have argued that sampling over the greater stimulus variability of multiple talkers helps focus the learner’s attention on those acoustic cues most relevant to the novel categories (Bradlow et al., 1997; Logan et al., 1991). Other studies have manipulated characteristics of the stimuli on the assumption that such manipulation aids the listener in directing attention to some acoustic cues in preference to others (Jamieson and Morosan, 1986; McCandliss et al. 2002). These authors have found that enhancement of phonetic features improves learning and hypothesize that this enhancement helps participants orient their attention to these features, thus improving learning. Similarly, some studies have manipulated training conditions to allow comparison of specific acoustic cues, with the assumption that listeners would attend to these cues. Christensen and Humes (1997) conclude that listeners shifted their attention according to the speech-like phonetic cues that were emphasized by different training conditions. Francis and colleagues have provided listeners with feedback about the category membership of the stimuli and results demonstrated that trained acoustic cues were, in fact, weighted more heavily in post-training tests (Francis et al., 2000; Francis and Nusbaum, 2002).

Note, however, that while each of the studies above posit an attentional component to learning and use stimuli that promote attention to certain cues, none directly manipulate listeners' attention during training. Very few studies of phonetic learning have attempted to directly manipulate attention. Polka (1992) and Werker and Tees (1984) manipulated attention during *testing* and found no effect. In a first attempt, Guion and Pederson (2007) manipulated attention during *training* to either phonetic form or meaning and found that attention enhanced learning but they did not explore the effects of differential attention within the phonetic domain.

The current study manipulates endogenous orienting of attention through varying instructions that encourage participants to orient more strongly to one or another phonetic segment in identical stimuli. The control of orienting presumably increases 1) the likelihood of detection of the relevant phonetic information, 2) the processing in short term memory and 3) the transfer to long-term memory (Posner, 1980; Posner and Peterson, 1990; Logan et al., 1996).

Here, English monolinguals were randomly assigned to either of two groups: *consonant-attending* or *vowel-attending*. Phonetic categories were selected from the phonemic inventory of Hindi. It is hypothesized that greater attention to targeted segments of the stimuli during training will improve learning for those segments and this will be reflected in greater discrimination ability in post-training tests.

Importantly, both groups were exposed to identical auditory stimuli during training and testing. This ensures that any differences revealed by post-training testing are the results of differential attention during training rather than differential exposure during training. In order for the participants to maintain orienting either to consonants or to vowels during training, an identification task with feedback was used in which participants had to identify either the consonant or the vowel (depending on group assignment) for each presentation of a Hindi word.

2.Method

2.1 Participants

Forty-two (35 female) monolingual English speakers (mean age 22.4) participated for course credit. None had lived in a non-English speaking region for more than six months, nor formally studied another language for more than three years, nor had any appreciable exposure to any South Asian language. All reported normal hearing. Participants were randomly assigned to one of two groups: consonant-attending (N=21) and vowel-attending (N=21).

2.2 Materials

2.2.1 Material for discrimination tests. Hindi has a number of phonetic contrasts not found in English among both vowels and consonants. Vowel discrimination was expected to be fairly accurate, as the contrasts employ duration and quality differences similar to those employed in English. On the other hand, the consonant discrimination was expected to be less accurate, as some of the contrasts employ novel phonetic distinctions known to be difficult for English speakers (Polka, 1991).

For the discrimination test, a female Hindi speaker produced minimal pair stimuli of the following Hindi initial stop contrasts: /b/-/bʰ/, /d̪/-/ʈ/, /k/-/g/, /t̪/-/t̪ʰ/, /t̪/-/ʈ/ and the following medial vowel contrasts: /a/-/aː/, /e/-/o/, /i/-/iː/, /o/-/a/, /u/-/uː/. For example, /buːr/ 'sawdust' vs. /bʰ uːr/ 'fountain' was one of the minimal pairs recorded for the /b/-/bʰ/ contrast and /ʈap/ 'hood' vs. /ʈaːp/ 'hoof' was one of the minimal pairs recorded for the /a/-/aː/ contrast. Each word was said three times and each production was edited into a separate wav file.

2.2.2 Materials for identification training. The training paradigm used four other native Hindi speakers (2 men, 2 women) reading 27 monosyllabic words beginning with one of eight consonants /k g ʈ t̪ t̪ʰ d̪ b bʰ/ and with one of eight medial vowels /a aː i iː u uː e o/. Thus, the same set of words was used for both consonant and vowel training. For example, for the word /bʰaːr/ ‘burden’, the consonant-attending group was to select the symbol corresponding to /bʰ/, while the vowel-attending group was to select the symbol corresponding to /aː/. None of the training words were the same as those used for the discrimination test. Each word was repeated three times by each speaker, digitally recorded, and edited into separate files.

2.3 Procedure

A discrimination pretest-posttest with intervening training design was used. Both groups of participants took the same pretest-posttest, which consisted of both vowel and consonant discrimination trials. Participants came to three sessions over the course of two weeks. In the first session, the pretest was administered, followed by identification training with feedback. The second session continued the training and in the final session, there was a final training, a short break, and then the posttest was administered. Identification training with discrimination testing was used for two reasons: (1) Identification testing was not possible because participants could not be expected to label unknown categories in the pretest nor in the posttest for the segment type (consonant or vowel) they were not trained on. (2) Identification training has been shown to improve phonetic discrimination (Wayland and Li, 2008).

2.3.1 Training. Both groups of participants were trained in an identification task using the *same* Hindi stimuli presented individually over headphones in a forced-choice identification task with feedback. The consonant-attending group was given explicit instructions to attend to word-initial consonants and the vowel-attending group to attend to word-medial vowels. The choices consisted of eight buttons displaying arbitrarily-assigned non-orthographic characters. When the correct button was chosen, positive auditory feedback was given. When the incorrect button was chosen, the correct button flashed silently. There were three training sessions, each consisting of four blocks of 81 trials (the three repetitions of 27 words). Blocks alternated between the male and female speakers.

2.3.2 Discrimination test. An identical discrimination test was used for both the pre- and posttest with counterbalanced blocks. An AXB categorical discrimination procedure was used to test the five Hindi consonant and five Hindi vowel contrasts. Participants had to respond correctly to four practice trials, representing contrasts that were not subsequently tested, before taking the test.

The stimuli were presented in triads over headphones. Each triad consisted of three words, the first and third were two contrastive words constituting a minimal pair. The middle word was a different production by the same speaker of either the first or the third word. Half of these trials used minimal pairs differing by the initial consonant (e.g., buːr$_3$ buːr$_1$ bʰ uːr$_3$) and the other half used minimal pairs differing by the medial vowel (e.g., ʈap$_2$ ʈaːp$_2$ ʈaːp$_1$). There was no signaling of the type of contrast tested for each trial. No feedback was given.

The test consisted of 160 trials (10 contrasts X 4 minimal pairs X 4 trials) in two pseudo-randomized, counterbalanced blocks. Each of the four possible orders (AAB, ABB, BBA, BAA) was presented the same number of times for each contrast type. The inter-stimulus interval was 2 seconds and the inter-trial interval was 2 seconds after response. The participants were to decide whether the middle word was more like the first or the last word presented and to click with a computer mouse either of two buttons on the computer monitor labeled “first” and “last”.

3. Results

3.1 Identification training

Both groups demonstrated improvement in the identification of Hindi segments across the three training sessions and, for both groups, some segment types were more difficult to identify than others. ANOVAs (α=.05) returned significant main effects of Session (3) and Segment Type (8) for both groups. Neither group displayed an interaction between Session and Segment Type [Consonant-attending: Session $F(2,480)=81.33$, $p < .001$; Segment Type $F(7,480)=58.58$, $p <.001$, Session*Segment Type $F(14, 480)=0.59$; Vowel-attending: Session $F(2,480)=80.55$, $p < .001$; Segment Type $F(7,480)=23.88$, $p <.001$, Session*Segment Type $F(14, 480)=0.23$], indicating that identification improved equally across all segments for both groups.

For the Consonant-attending Group, the mean correct response rate for session one was 57.3%, session two 74.1%, session three 80.9%, with the segments ranking in increasing difficulty: /g/ /bʰ/ /b/ /d̪/ /t̪ʰ/ /k/ /ʈ/ /t̪/. For the Vowel-attending Group, the mean correct response rate for session one was 51.3%, session two 69.5%, session three 75%, with the segments ranking in increasing difficulty: /iː/ /aː/ /i/ /o/ /e/ /a/ /uː/ /u/.

3.2 Pre-training discrimination test

Examining the pretest, the two groups showed no difference in their discrimination scores prior to training. Separate ANOVAs for vowel and consonant contrasts with the factors of Group (consonant-attending vs. vowel-attending) and Contrast (5 vowel and 5 consonant contrast types) showed no effect of Group [vowels $F1(1,40)=1.31$, $F2(1,150)=0.71$; consonants $F1(1,40)=2.22$, $F2(1,150)=0.98$] nor was there an interaction between Group and Contrast [vowels $F1(4,160)=1.68$, $F2(4,150)=0.57$; consonants $F1(4,160)=0.25$, $F2(4,150)=0.12$]. However, the main effect of Contrast was significant [vowels $F1(4,160)=4.56$, $p=.002$ $F2(4,150)=3.11$, $p=.017$; consonants $F1(4,160)=49.78$, $p<.001$, $F2(4,150)=18.03$, $p<.001$]. This indicates that some contrast types were more accurately discriminated than others for both vowel and consonant contrast types. As predicted, the combined groups responded to the vowel contrasts more accurately (97% correct) than the consonant contrasts (73% correct).[1]

3.3 Post-training evaluation

The effect of the two orienting conditions during training on discrimination of both vowels and consonants was investigated. The difference between the post- and pretest scores were submitted to a Group (consonant-attending vs. vowel-attending) by Trial Type (consonant vs. vowel discrimination) ANOVA. A significant interaction between the factors was found [$F1(1,40)=5.31$, $p=.026$, $F2(1,16)=4.42$, $p=.052$], indicating that training conditions affected the discrimination of vowels and consonants differentially. This interaction was investigated with separate ANOVAs for vowel and consonant discrimination contrasts. In these ANOVAS, contrast type was included as a factor to explore whether the different accuracy levels for the different contrasts observed in the pretest interacted with training condition.

For vowels, the mean posttest and pretest scores indicated no effect of training for either group (see Figure 1). The difference in vowel scores between the posttest and the pretest was submitted to an ANOVA with factors of Group and Contrast ([a]-[aː], [e]-[o], [i]-[iː], [o]-[a], [u]-[uː]). There was no effect of Group [$F1(1,40)=0.16$, $F2(1,150)=3.09$] nor Contrast [$F1(4,160)=1.38$, $F2(4,150)=1.34$] and Group and Contrast did not interact [$F1(4,160)=1.01$, $F2(4,150)=0.59$]. This lack of a group effect for the vowels was expected given their already high pretest scores (97% correct).

However, for consonants, the difference scores did differ across the groups. The consonant-attending group demonstrated an improvement after training, whereas the vowel-attending group did not (see Figure 1). The difference in consonant scores was submitted to an ANOVA with the factors of Group and Contrast ([b]-[bʱ], [d̪]-[ʈ], [k]-[g], [t̪]-[t̪ʰ], [t̪]-[ʈ]). There was a significant effect of Group [*F1*(1,40)=5.03, *p*=.03, *F2*(1,150)=4.15, p=.04] but no effect of Contrast [*F1*(4,160)=0.71, *F2*(4,150)=0.78], nor was there an interaction between Group and Contrast [*F1*(4,160)=0.47, *F2*(4,150)=1.02]. These results indicate that orienting attention toward consonants, but not vowels, during training improved the subsequent consonant discrimination. Furthermore, the effect was not limited to individual contrast types, but was found across all contrasts.

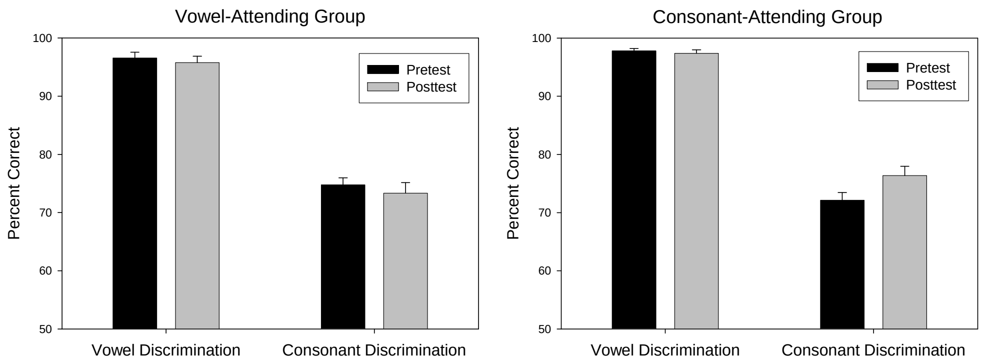

Fig. 1. Percentage correct on pretest and posttest for the consonant- and vowel-attending groups.

To further test the effect of training condition on learning, each group was examined to determine whether there was a significant difference between pre- and posttest scores for consonants. The pretest and posttest consonant discrimination scores were submitted to paired, one-tailed t-tests. For the vowel-attending group, no significant difference between the post- and pretest scores was found [*t1*(20)= -0.76, *t2*(79)= -0.26]. For the consonant-attending group, a significant difference between the post- and pretest consonant discrimination scores was found in the predicted (positive) direction [*t1*(20)=2.49, *p*=.01, *t2*(79)=2.54, *p*=.007]. These results indicate that orienting attention toward consonants during training facilitated learning.

4. Discussion

Orienting attention is well known to facilitate grammatical aspects of language learning (see Schmidt, 2001, for a review). Attention has been hypothesized to facilitate phonetic learning as well (Christensen and Humes, 1997; Francis and Nusbaum, 2002), although this hypothesis has never been tested explicitly. In this study, the group instructed to attend to consonants during training demonstrated learning of consonant contrasts, whereas, the group instructed to attend to vowels instead did not demonstrate learning of consonant contrasts. (Of course, since vowel discrimination was high at pretest, no learning for either group could be expected for vowels.) These results confirm that orienting attention during phonetic training facilitates learning of the specific class of stimuli to which the participants are instructed to attend.

The mechanisms of phonetic learning merit further exploration. Previous training studies have manipulated the types of stimuli (e.g., McCandliss et al. 2002), but not directly manipulated the participants' attention to those stimuli. This study represents an early attempt to manipulate conscious attention during the training of phonetic categories by varying instructions to attend to certain class-specific stimuli. These instructions varied the endogenously-controlled orienting of attention across the two groups. Orienting is known to generally have differential effects on the detection and manipulation of stimuli in working memory (Posner and Peterson, 1990), and/or the likelihood of transfer to long-term memory (Logan et al., 1996). Thus, presumably, the attentional orienting manipulated in this study modulated the selective uptake and storage of phonetic information. Many processes are undoubtedly involved from auditory detection to final encoding and we do not yet know which specific processes are the more dependent on attentional orienting. Further development of this line of research could include the use of different phonetic contrasts, more extensive training periods, and different methods of manipulating attention. Additionally, attention could be directed toward specific acoustic features, rather than whole segments. For now, however, it is noteworthy that the simple method of participant-directed orienting of attention can have a measurable effect on phonetic learning.

Acknowledgments

We would like to thank the five anonymous Hindi speakers for recording our speech materials, Megha Sundara for helpful comments, and Brad Biehl and Michael Peterson for their invaluable assistance running the experiment. Funding for this work was kindly provided by the University of Oregon.

Note

[1] The higher accuracy in the vowel discrimination test than in the vowel identification training indicates that juxtaposition of different vowel tokens facilitates discrimination, whereas individually presented vowel tokens are relatively difficult to categorize.

References and links

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., and Tohkura, Y. (**1997**). “Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production,” J. Acoustic. Soc. Am. **101**, 2299-2310.

Christensen, L. A., and Humes, L. E. (**1997**). “Identification of multidimensional stimuli containing speech cues and the effects of training,” J. Acoustic. Soc. Am. **102**, 2297-2310.

Francis, A. L., Baldwin, K., and Nusbaum, H. C. (**2000**). “Effects of training on attention to acoustic cues,” Percept. Psychophys. **62**, 1668-1680.

Francis, A. L., and Nusbaum, H. C. (**2002**). “Selective attention and the acquisition of new phonetic categories,” J. Exper. Psych.: Human Percept. Perform. **28**, 349-366.

Guion, S. G., and Pederson, E. (**2007**). “Investigating the role of attention in phonetic learning,” *Language Experience in Second Language Speech Learning*, edited by O.-S. Bohn and M. Munro (John Benjamins, Amsterdam), pp. 57-77.

Iverson, P., and Kuhl, P. K. (**1995**). “Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling,” J. Acoustic. Soc. Am. **97**, 553-562.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E. Tohkura, Y., Kettermann, A., and Siebert, C. (**2003**). “A perceptual interference account of acquisition difficulties for non-native phonemes,” Cognition **87**, B47-B57.

Jamieson, D.G., and Morosan, D.E. (**1986**). “Training non-native speech contrasts in adults: Acquisition of the English D-T contrast by francophones,” Percept. Psychophys. **40**, 205-215.

Logan, G. D., Taylor, S. E., and Etherton, J. L. (**1996**). “Attention in the acquisition and expression of automaticity,” J. Exper. Psych.: Learn. Mem. Cog. **22**, 620-638.

Logan, J. S., Lively, S. E., and Pisoni, D. B. (**1991**). “Training Japanese listeners to identify English /r/ and /l/: A first report,” J. Acoustic. Soc. Am. **89**, 874-886.

McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., and McClelland, J. L. (**2002**) “Success and failure in teaching the [r]-[l] contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization in spoken language perception,” Cog., Affect., Behav. Neurosci. **2**, 89-108.

Polka, L. (**1991**). “Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions,” J. Acoustic. Soc. Am. **89**, 2961-2977.

Polka, L. (**1992**). “Characterizing the influence of native language experience on adult speech perception,” Percept. Psychophys. **52**, 37-52.

Posner, M. I. (**1980**). “Orienting of attention.” Quarter. J. of Exper. Psych. **32**, 2-25.

Posner, M. I., and Peterson, S. E. (**1990**). “The attention system of the human brain,” Annual Rev. Neurosci. **13**, 25-42.

Schmidt, R. (**2001**). “Attention,” in *Cognition and Second Language Instruction*, edited by P. Robinson (Cambridge Univ. Press), pp. 3-32.

Strange, W. and Shafer, V. L. (**2008**). “Speech perception in second language learners: The re-education of selective perception,” in *Phonology and Second Language Acquisition*, edited by J. G. Hansen Edwards, and M. L. Zampini (John Benjamins, Amsterdam), pp. 153-191.

Wayland, R. P., and Li, B. (**2008**). “Effects of two training procedures in cross-language perception of tones,” J. Phon. **36**, 250-267.

Werker, J. F., and Tees, R. C. (**1984**). “Phonemic and phonetic factors in adult cross-language speech perception,” J. Acoustic. Soc. Am. **75**, 1866-1878.