EFFECTS OF ATTENTION AND TRAINING METHOD ON THE IDENTIFICATION OF AMERICAN ENGLISH VOWELS AND CODA NASALS BY NATIVE JAPANESE LISTENERS

Takeshi Nozawa

Ritsumeikan University
t-nozawa@ec.ritsumei.ac.jp

ABSTRACT

The accuracy with which native Japanese listeners identified American English vowels and coda nasals was assessed before and after training. The listeners were divided into four groups, each of which received a different type of training. Two of the four groups were vowel-oriented; one of these groups received vowel identification training (VI), while the other received vowel discrimination training (VD). The other two groups were nasal-oriented. One of the nasal-oriented groups received nasal identification training (NI), and the other received nasal discrimination training (ND).

The results revealed that the VI group made more gains in its ability to identify vowels than the other groups after training. However, training appeared to have no effect on nasal identification, and no significant difference among the groups was observed.

Keywords: identification, training, attention, vowels, nasals.

1. INTRODUCTION

In perception experiments, participants are typically presented with a syllable rather than the target sound in isolation. For instance, words like “right” and “light” are presented when a listener’s task is to identify the English phonemes /r/ and /l/. In such experiments, the listeners hear the following vowel and the consonant in addition to the initial target consonant.

This study investigates whether listeners’ sensitivity to non-native vowels and coda nasals is affected by whether their attention is focused on the vowel or the coda nasal when they hear each stimulus. Pederson, Guion-Anderson [4] trained two monolingual English listener groups with the same stimuli. One group was instructed to attend to the vowels while the other was instructed to attend to the consonants. Pederson, Guion-Anderson [4] demonstrated that only the consonant-oriented group showed improvement in consonant perception after training.

Also of interest is whether discrimination training can improve a listener’s identification accuracy. Wayland and Li [5] demonstrated that both discrimination and identification training effectively improve the accuracy with which non-native listeners discriminate Thai mid and low tone contrasts. Ingram and Park [2], however, suggest that identification tests require more phonological processing of signals than discrimination tests.

In the present study, the participants (listeners) identified 1) American English vowels before stop consonants, 2) American English vowels before nasal consonants, and 3) American English postvocalic nasals, which they heard in the first and last sessions (the pretest and posttest). After the initial sessions, the listeners were divided into four groups, each of which received a different type of training: 1) vowel identification training (VI group), 2) vowel discrimination training (VD group), 3) nasal identification training (NI group), and 4) nasal discrimination training (ND group). In the training sessions, the listeners heard the same set of stimuli in the pretest and posttest.

If the focus of the training is an important factor, the VI and VD groups should outperform the NI and ND groups on the vowel identification task in the posttest. Likewise, the NI and ND groups should outperform the VI and VD groups on the nasal identification task. All of the listeners heard prenasal vowels and postvocalic nasals. Therefore, if repeated exposure to stimuli can enhance listeners’ identification accuracy, each of the four groups should show some improvement in their ability to identify prenasal vowels and postvocalic nasals. However, the NI and ND groups only heard stimuli that contain nasal consonants in the training sessions. Therefore, these groups are not expected to show improvement in their identification of vowels before stop consonants.

If discrimination training is as effective as identification training in improving identification accuracy, then the VD and ND groups should perform as well as the VI and NI groups, respectively.

2. EXPERIMENT

2.1. Stimuli

Four native speakers of American English served as talkers. These talkers produced /i, ɪ, eɪ, ɛ, æ, ɑ, ʌ/ in /bVb/, /bVd/, /bVg/, /bVm/, /bVn/ frames, and /ɪ, æ, ɑ, ʌ/ in /bVŋ/ frame. Their utterances were digitally recorded and edited to be used as stimuli.

2.2. Listeners

Thirty-two native speakers of Japanese participated as listeners. All of the listeners were undergraduate students at a private university in Japan.

2.3. Procedure

2.3.1. Pretest

All of the listeners performed the following three identification tasks. The order of the tasks was counterbalanced across listeners. In each identification task, a listener heard one stimulus per trial. The inter-trial interval was fixed at 1,000 ms.

The listeners identified American English vowels before stop consonants /b, d, g/. Seven choices were presented using English spelling, each of which represents one vowel: bee-, bi-, bay-, be-, ba-, bo-, and bu-. The listeners were told that "bee-" represents the vowel sound in "beat", and so on. They were told that the final consonant would be /b/, /d/ or /g/ but that no matter what the final consonant was, when they heard the vowel sound used in "beat", they should choose "bee-", and so on. The listeners were also told that nonwords were included and that they should not be distracted by the lexical meaning of a stimulus. After receiving the instructions, the listeners participated in practice trials. This identification task consisted of 84 trials (4 speakers × 7 vowels × 3 following consonants).

The listeners also identified American English vowels in prenasal positions. The same seven choices were given. The listeners were told that the choice "bee-" represents the vowel sound in "beam", "bean", and so on. They were not told that /biŋ/, /beɪŋ/ and /bɛŋ/ were not included in the stimuli because knowing that fewer choices are available before /ŋ/ could make it easier to identify the vowel. This identification task consisted of 72 trials (4 speakers × 7 vowels × 2 following nasal consonants + 4 speakers × 4 vowels × 1 following nasal consonant).

The listeners also identified postvocalic nasals. The same set of stimuli used for the prenasal vowel identification was used. Three choices were given: "m", "n", and "ng". This identification task consisted of 72 trials.

The listeners heard the stimuli over a headset and responded by moving the cursor to and clicking on the box that represented the vowel they heard.

Figures 1-3 show the percentages of correct responses on the pretest's three identification tasks.

Figure 1: The percentages of correct vowel identifications in /bVb/, /bVd/, /bVg/ contexts

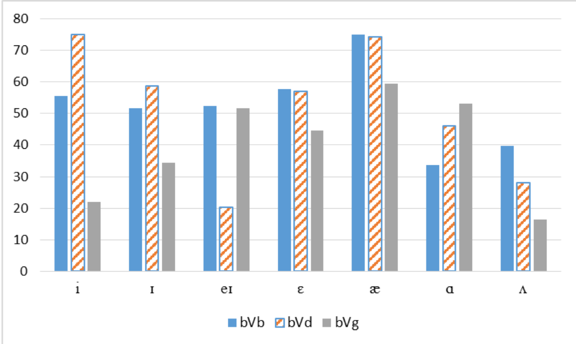


Figure 2 The percentages of correct vowel identifications in /bVm, /bVn, /bVŋ/contexts

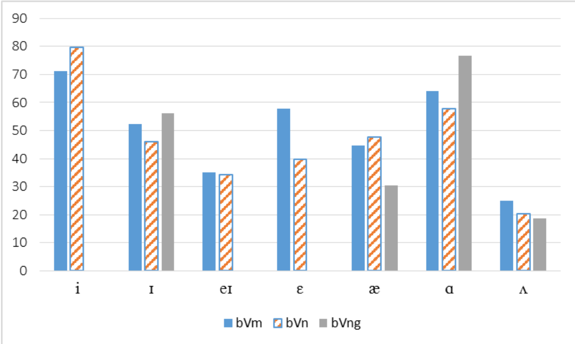
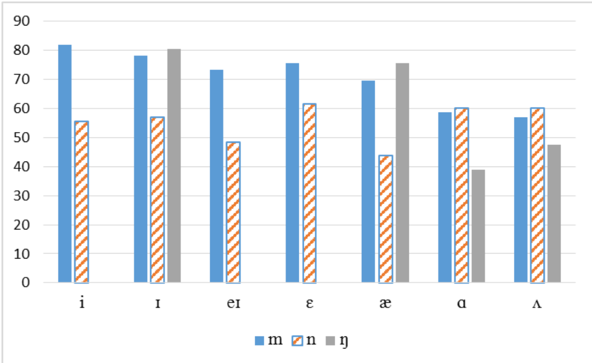


Figure 3 The percentages of correct responses of nasal identification



A three-way ANOVA (4 Listener Groups × 7 vowels × 3 stop consonants) of vowel identification before stop consonants revealed significant main effects of vowels [$F(6, 167)=11.30$, $p < .001$] and following stop consonants [$F(2, 56)=13.04$, $p < .001$] but an insignificant main effect of Listener Group [$F(3, 28)=0.68$, $p=.567$]. Post-hoc pair-wise comparisons revealed that /æ/ was identified the most accurately ($p < .05$) and that /ʌ/ was identified less accurately than /i/, /ɪ/, /æ/ and /ɑ/ ($p < .05$). It was also revealed that vowels were identified better before /d/ than before /b/ and /g/ ($p < .05$).

In the statistical analysis of the identification of prenasal vowels, /i/, /eɪ/ and /ɛ/ were excluded because these three vowels do not occur before /ŋ/. A three-way ANOVA (4 Listener Groups × 4 vowels × 3 nasal consonants) yielded a significant main effect of vowels [$F(3, 84)=5.42$, $p < .001$]. However, no significant effects of the following nasal consonants [$F(2, 56)=0.997$, $p=.376$] or of Listener Group were shown [$F(3, 28)=1.90$, $p=.152$]. Post-hoc pair-wise comparisons revealed that /ʌ/ was identified less accurately than /ɪ/, /æ/ and /ɑ/ and that /æ/ was identified less accurately than /ɑ/.

In the statistical analysis of the identification of postvocalic nasals, nasals after /i/, /eɪ/ and /ɛ/ were excluded because /ŋ/ was never included after these vowels. A three-way ANOVA (4 Listener Groups × 4 preceding vowels × 3 nasal consonants) yielded significant main effects of the preceding vowel [$F(3, 84)=10.01$, $p < .001$] and of nasals [$F(2, 56)=12.76$, $p < .001$], but no significant main effect of Listener Group was shown [$F(3, 28)=0.78$, $p=0.515$]. Post-hoc pair-wise tests revealed that nasals were less accurately identified after /æ/ than after /ɪ/ and /ɑ/ and after /ʌ/ than after /ɪ/ and /ɑ/. It was also revealed that /m/ was identified more accurately than /n/ and /ŋ/.

2.3.2. Training sessions

The listeners were divided into four groups, and each group received different training. Each listener participated in two training sessions, each of which lasted 90 minutes. The stimuli used in the training sessions had been used in the pretest.

The listeners numbered 1, 5, 9, 13, 17, 21, 25, and 29 received vowel identification training. As in the pretest, listeners heard one stimulus per trial and chose the vowel they heard from seven choices. However, in the training sessions, feedback was given each time a listener gave a response. When a listener made an error, the correct choice blinked on the computer screen. This group of listeners heard vowels before stop and nasal consonants.

The listeners numbered 2, 6, 10, 14, 18, 22, 26, and 30 received vowel discrimination training. In this training, an AXB format was adopted. The listeners heard three stimuli per trial and indicated whether they believed the second stimulus belonged to the same category as the first or the third stimulus by clicking on either the "first" or the "last" box on the screen. The three stimuli in each trial were taken from the utterances of three different speakers. The following eight vowel pairs were used: /i/-/ɪ/, /i/-/eɪ/, /ɛ/-/ɪ/, /ɛ/-/eɪ/, /æ/-/ɛ/, /æ/-/ɑ/, /æ/-/ʌ/, and /ɑ/-/ʌ/. Vowels before the velar consonants /g/ and /ŋ/ were not included because it was impossible to use the same eight vowel pairs before /ŋ/. Trials were blocked according to the following consonants. Each vowel pair appeared in eight trials. Thus, 128 (8 vowel pairs × 8 trials × 2 following consonants) trials were created for the /bVb/ and /bVd/ frames. Likewise, 128 trials were created for the /bVm/ and /bVn/ contexts. Feedback was given each time a listener gave a response. The inter-stimulus interval and inter-trial interval were both 1,000 ms.

The listeners numbered 3, 7, 11, 15, 19, 23, 27, and 31 received nasal identification training. As in the pretest, listeners heard one stimulus per trial and chose the nasal they heard from among three choices. As in the other trainings, feedback was given.

The listeners numbered 4, 8, 12, 16, 20, 24, 28, and 32 received nasal discrimination training. As in the vowel discrimination training, an AXB format was adopted. The listeners heard three stimuli per trial and indicated whether they believed the second stimulus contained categorically the same nasal as the first stimulus or the third stimulus. The three stimuli in each trial were taken from the utterances of three different speakers. The following nasal pairs were created: /m/-/n/ beam-bean, bim-bin, bame-bane, bam-ban, and bum-bun; /m/-/ŋ/ beam-bing, bim-bing, bame-bang, bam-bang, and bum-bung; and /n/-/ŋ/ bean-bing, bin-bing, bane-bang, ban-bang, and bun-bung.

In American English, /ɪ/ and /æ/ are said to be raised to the region of /i/ and /eɪ/. Pairs like beam-bing and bane-bang were included because the phonetic difference between the preceding vowels could offer an acoustic cue to help listeners differentiate coda nasals. See, for example, Johnson and Dicanio [3] and Baker, Mielke, Archangeli [1].

Each nasal pair was tested in eight trials. Thus, 120 trials were created (8 trials × 15 pairs). The inter-stimulus interval and inter-trial interval were both 1,000 ms.

2.3.3. Posttest

All of the listeners completed the same task as that presented in the pretest, i.e., vowel identification before stop consonants, vowel identification before nasal consonants, and postvocalic nasal identification. The order of the tasks was counterbalanced across listeners.

2.3. Results

Table 1 shows the rate of increase of vowel identification accuracy before stop consonants. A three-way ANOVA (4 Listener Groups × 7 vowels × 3 following stops) yielded significant main effects of Listener Groups [$F(3, 28)=4.06$, $p=.016$] and vowels [$F(6, 168)=3.4$, $p=.003$]. However, the main effect of the following stops was not significant [$F(2, 56)=0.56$, $p=.571$]. Post-hoc pair-wise comparisons revealed that the vowel identification group had a significantly better rate of increase than the vowel discrimination and nasal identification groups ($p < .05$). It also revealed that the increase ratio of the identification of /ʌ/ was significantly lower than that of /eɪ/ and /ɛ/.

Table 1: Average increase ratio, in percent, of vowel identification before stop consonants.

	/bVb/	/bVd/	/bVg/
VI	16.07	13.39	19.64
VD	4.46	0.45	2.23
NI	-3.125	8.04	2.68
ND	-3.125	2.23	8.48

Table 2 shows the rate of increase of accurate prenasal vowel identification. A three-way ANOVA (4 Listener Groups × 4 vowels (/ɪ, æ, ɑ, ʌ/) × 3 following nasals) yielded a significant main effect of Listener Groups [$F(3, 28)=6.08$, $p=.003$], but the main effects of vowels [$F(3, 84)=1.43$, $p=.239$] and of following nasals [$F(2, 56)=0.17$, $p=.846$] were not significant.

Table 2: Average increase ratio, in percent, of vowel identification before nasal consonants.

	/bVm/	/bVn/	/bVŋ/
VI	18.75	23.2	12.5
VD	8.036	3.13	7.03
NI	7.14	1.79	2.34
ND	9.82	1.34	-4.69

Table 3: Average increase ratio, in percent, of three postvocalic nasals.

	/m/	/n/	/ŋ/
VI	1.79	10.71	6.25
VD	-1.34	8.04	3.91
NI	6.25	16.52	14.06
ND	6.70	24.55	9.38

Table 3 shows the rate of increase of accurate postvocalic nasal identification. A three-way ANOVA (4 Listener Groups × 3 following nasals × 4 vowels (/ɪ, æ, ɑ, ʌ/)) yielded a significant main effect of nasals [$F(2, 56)=5.35$, $p=.007$], but the main effects of the preceding vowels [$F(3, 84) = 0.50$, $p=.686$] and Listener Groups [$F(2, 56) = 0.167$, $p=.846$] were not significant. Post-hoc pair-wise comparisons revealed that the identification of /n/ increased significantly more than that of /m/ ($p < .05$).

3. SUMMARY

Overall, the results show that attention can significantly improve the accuracy with which listeners identify target sounds. However, the results of the vowel and nasal identification tasks show a slightly different tendency. In the two vowel identification tasks, the VI group showed significantly more improvement than the other three groups. The VD group, though vowel-oriented, did not improve significantly more than the two nasal-oriented groups. In this study, discrimination training was not as effective as identification training in improving listeners' identification accuracy.

The nasal identification task showed slightly different results. The identification accuracy of the two nasal-oriented groups improved more than that of the two vowel-oriented groups, but no statistically significant difference was observed. Unlike on the vowel identification tasks, the ND group achieved a similar or even greater increase than did the NI group. Therefore, discrimination training was as effective as identification training for nasal identification. Additionally, to a lesser extent, the two vowel-oriented groups increased the accuracy with which they identified nasals, indicating that exposure to stimuli can enhance listeners' sensitivity even when their attention is not on the target sound.

The small size of the sample makes it imprudent to draw conclusions. Further research is necessary to determine how attention and training affect perception.

4. REFERENCES

[1] Baker, A., Mielke, J., Archangeli, D. 2008. More velar than[g]: Consonant coarticulation as a cause of diphthonization,. *Proceedings of the 16th West Coast Conference on Formal Linguistics*, 60-68

[2] Ingram, J.C. L., Park, S. G. 1998. Language, context, and speaker effects in the identification and discrimination of English /r/ and /l/ by Japanese and Korean listeners. *J. Acoust. Soc. Am*. 103, 1161–1174.

[3] Johnson, K., DiCanio, C. 2007. The acoustic and visual phonetic basis of place of articulation in excrescent nasals. *UC Berkeley Phonology Lab Annual Report*, 529-561

[4] Pederson, E., Guion-Anderson, S. 2010. Orienting attention during phonetic training facilitates learning. *J. Acoust. Soc. Am*. 127, EL54–59.

[5] Wayland, R.P., Li, B. 2008. Effects of two training procedures in cross-language perception of tones, *J. Phonetics* 36, 250–267