INDIVIDUAL DIFFERENCES IN THE ACQUISITION OF ENGLISH FRICATIVE DISTINCTIONS BY NATIVE SPEAKERS OF JAPANESE

William Martens & Stephen Lambacher
The University of Aizu, Aizu-Wakamatsu, JAPAN

ABSTRACT

Both at the beginning and at the end of their first year of education at the University of Aizu, a group of 80 native speakers of Japanese participating in an English language program were administered a five-alternative, forced-choice (5AFC) test to measure changes in their ability to distinguish between English voiceless fricatives The 80 pairs of the resulting 5-by-5 confusion matrices were submitted to cluster analysis over individuals. The two cluster solution was interpreted through an examination of changes in listener sensitivity and changes in response biases. Though an ANOVA gave no evidence that these two groups differed in their overall mean identification performance, there was a significant interaction showing that the members of these two groups used the response categories in a different manner. The signal-detection-theoretic analysis showed that the groups differed primarily in terms of bias toward particular fricative response categories, but also showed contrasting effects of vowel context on sensitivity versus response bias.

1. INTRODUCTION

A previous study by Lambacher *et al* [7] examined the ability of native Japanese listeners to distinguish between the English voiceless fricatives /f/, /s/, /sh/, /θ/ and /h/. The experiment stimuli consisted of 75 nonsense syllables in which the five fricatives (/f/, /s/, /sh/, /θ/ /h/) were presented in five vowel environments (/i E a o u/) and three different consonant positions relative to the vowel: consonant-vowel (CV), vowel-consonant-vowel (VCV), and vowel-consonant (VC). A five-alternative, forced-choice (5AFC) answer sheet was passed out to each of the subjects, and the subjects marked one of five possible responses for each stimulus. Identification scores ranged from a maximum of 88 % for the all the /sh/ stimuli to 55 % for all the /θ/ stimuli. The results showed that the confusion rates were highest for the /θ/ - /s/, /sh/ - /s/ and /f/ - - /h/ contrasts. The results also revealed that the subjects judgments were highly context-dependent, with both vowel environment and consonant position having a significant effect on the subjects' perception of the target fricatives.

2. METHOD

2.1. Subjects The subjects were 80 native Japanese adults ranging in age from 18 to 20 years old. All the subjects were first year students majoring in computer science at the University of Aizu in Fukushima Prefecture, Japan. As is common in Japan, all the subjects had six years of prior English training at the junior and senior high school levels. None of the subjects had any reported history of speech or hearing impediments.

2.2. Training. All of the Japanese subjects were enrolled in English as a Foreign Language (EFL) courses at the university's Center for Language Research (CLR). During the first two years of study, students must take two EFL classes each semester (4 semesters) for a total of eight classes. Although all of the Japanese subjects had received six years of prior English training at junior and senior high school levels, the focus was on reading and grammar; very little attention was given to listening, speaking, and pronunciation. At the time of the first experiment, the subjects were taking a pronunciation course, but English fricative sounds had not been covered. In addition, they were taking two other courses in writing and conversation/listening, respectively. After completing the first experiment, the subjects had four more English courses – one in conversation/listening, one in writing, and two in reading. Each English course was taught by a native speaker and met once a week for 90 minutes for a total of 13 weeks. The total hours in English class was between the first and second experiments was 78 (4 courses x 13 weeks x 1.5 hours). The subjects had other opportunities for English listening practice. For example, many of the computer science courses at the University of Aizu are conducted in English by foreign teachers, some of whom were native speakers.

2.3. Stimuli. The stimulus set consisted of a total of 75 syllables presented in consonant-vowel (CV), vowel-consonant-vowel (VCV), and vowel-consonant (VC) syllables spoken within a varied-vowel environment /i E a o u/ by three native speakers of English. These syllables were spoken at a normal rate by three phonetically-trained male native speakers of English. The stimuli were recorded in three randomized blocks.

2.4. Recording Method. The speech samples were generated using a binaural recording system located in a large anechoic chamber at the University of Aizu. The talker was located approximately two meters from the Bruel & Kjaer

(B&K) Head and Torso Simulator (Type 4128), using the Right Ear Simulator (Type 4158) and the Left Ear Simulator (Type 4159). The talker was positioned at ear level, but slightly off the median plane of the manikin, at an azimuth angle of approximately 15 degrees. This avoided the rather unnatural sensation associated with monaural headphone listening. The samples were recorded at a 48 kHz sampling rate using a Denon Model DTR-80P digital audio recorder. The samples were reproduced via Sony HL-90 headphones in the University of Aizu Language Media Laboratory and were presented at an average presentation level of 69 dB.

2.5. Procedure. A brief introduction to each of the five fricatives was given to the subjects before each of the two identification tests, in the form of practice trials. The subjects were told to identify each syllable they heard as containing /f/, /s/, /sh/, /θ/ or /h/. A five-alternative, forced-choice (5AFC) answer sheet was passed out to each of the subjects, and the subjects marked one of five possible responses for each stimulus. Four separate groups of subjects participated in the test, and they were asked to give a response to each item even if they were unsure of the identity of the fricative they heard. The 75-tokens were randomly presented three times to each of the subjects with a pause of about five minutes between each presentation. The stimuli were presented to the subjects with a pause of five seconds between each speech sample.

3. RESULTS

The resulting 5-by-5 confusion matrices were submitted to cluster analysis, which revealed that within the sample of 80 subjects, two subsets of these subjects exhibited distinctly different patterns of results. Though ANOVA gave no evidence that these two groups differed in their overall mean identification performance, a significant interaction was observed that indicated that the members of these two groups used the response categories in a different manner. The response rates of these two groups of subjects were analyzed in terms of receiver operating characteristics, so that changes in fricative identifiability (measured by the index of sensitivity, d prime) could be distinguished from changes in the likelihood of given fricative responses (measured by the index of bias, β).

First, note that the main effect of group membership on ratings was not statistically significant. The ANOVA gave F(1,85)=0.076, p = 0.783. We are forced to conclude that the two groups have the same mean response (if we were to reject the null hypothesis here, there is a 78 % chance that we make a type I error). Also, the interaction between group and stimulus was very small, giving $F(4, 340) = 1.336, p = 0.256$. Again, we must conclude that the effect on stimulus on ratings did not depend upon group membership. But the effect of response category on ratings did depend upon group membership. ANOVA for this interaction gave $F(4, 340) = 7.734, p < 0.001$. An even bigger ratio was obtained for the three-way interaction of group, stimulus, and response: $F(16, 1360) = 38.046, p < 0.001$. The message is that paying attention only to student performance, as measured, for example, by percent correct on examination, will likely not reveal the differences that exist between the students.

4. DISCUSSION

4.1. Individual differences in L2 speech learning. When addressing individual differences in the L2 speech acquisition, the question may arise why some subjects' sensitivity to the fricative distinctions was greater than others. Is it a just a case of some subjects having superior hearing ability than others, or are there more subjective factors involved? A brief look at the philosophy of individual learning differences may serve to offer some reasons for this.

Research has shown that learners vary in their learning traits and processes of thinking, and in their aptitudes for learning, willingness, and styles of learning which all impact the learning process. Each learner approaches the learning process with his/her own filter, and each learner filters instruction through their own set of filters or lenses. For example, in the case of L2 speech acquisition, a built-in phonological filtering mechanism or process may prevent L2 learners from accurately perceiving nonnative phonemes ([3] [12]. Phonological filtering makes distinguishing the acoustic differences between native and nonnative sounds for L2 learners more difficult. According to [12], a learner's phonological filter acts like "sieve" and passes only information useful to categorizing sounds in the L1. As a result, L2 learners have a tendency to disregard allophonic variations which distinguish L1 sounds as well as the subcategorical phonetic differences which distinguish similar sounds in L1 and L2 [11].

In this study, two groups of Japanese learners were involved – each group having different sensitivity to some fricative distinctions and also a different pattern of biases. Significant individual differences exist between these two groups regarding learning attributes they bring into the classroom. How these differences might affect their ability to successfully or unsuccessfully distinguish between English fricatives is worth addressing here. One major difference is in student motivation. For example, some learners are more active participants than others. This may be due to several reasons. The more active learner may want to integrate into the L2 culture and may therefore be more of a risk-taker, not worrying about sounding more like a native speaker.

For the teacher of English as a Second/Foreign Language (ESL/EFL), an awareness of individual learning differences can be helpful in prioritizing training objectives, as well as help teachers to better understand the difficulties in relation to specific tasks. Individual learning differences can be attributed to the mental processing required by the specific learning task being performed. Thus, it is essential that the nature of the learning task be seriously considered. Thinking and learning will differ based on the type of task being used. Learners possess varying aptitudes for different learning outcomes. Although learners must be willing or motivated to learn as well as have the ability to learn, but they must also be exposed a classroom that effectively fosters learning in order to be successful learners.

4.2. L2 speech training for individual differences. How can a EFL/ESL teacher address individual differences in the classroom and provide an environment that helps learners be more sensitive to English voiceless fricatives distinctions, and at the same time consider both types of learners to that they both can benefit from the instruction? For example, how does one take into consideration the strong bias factor of one group of subjects in the the development of a training program to help Japanese to learn to distinguish L2 speech contrasts? According to [6], individuals react to different forms of instruction in diverse ways, and learning outcomes are affected by the form of instruction. Thus, different instructional activities will differentially affect learning outcomes. In the case of L2 speech training, training the subjects who are already more sensitive to begin with is much less difficult. The main difficulty lies with training the biased subjects as they will need more focused training.

Researchers agree that a combination of auditory and visual feedback can be helpful in teaching sound segmentals, suprasegmentals, and other aspects of pronunciation programs ([10] and [1]). The use of visual acoustic feedback in addition to auditory only feedback can be effective in L2 speech learning. Many programs include a dual display with top and bottom screens which help learners to objectively evaluate their speech errors and progress through analyzing and visually comparing their own pronunciation with a native speaker's.

Exposing Japanese subjects to gairaigo or "loan words" containing voiceless fricatives and corresponding English words can help them to perceive the phonetic differences between them. For example, the English-language pronunciation of "milk shake" differs from the common pronunciation of the term by native speakers of Japanese. The /shE/ is replaced by the /sE/ sound of the katakana character. When these loanwords were adopted, the unusual English fricative sounds were replaced by the closest sounds in the Japanese syllabary.

Another effective method is the use tokens produced by multiple tokens to provide more variability to subjects. For example, research of [9] and [4] suggests that variability in speakers and tokens during training results in improved category formulation. In their experiments, Japanese learners were better able to form robust categories for the /r/ - /l/ distinction if they are exposed to tokens produced by multiple speakers rather than a single speaker.

Another method in L2 speech training should be a focus on L2 speech production of sounds and not just perception. Research shows that there is a direct link between speech perception and production, and that improvement in perception of L2 speech contrasts transfers over to the production domain and vice versa [2]. According to the motor theory (e.g., [8]) listeners perceive speech in terms of their own articulatory gestures. Also, the direct-realist theory of speech perception (e.g., [5]) posits that listeners directly perceive the articulatory gestures of the speaker in terms of the structure they impart to the acoustic medium.

5. CONCLUSION

Cluster analysis found that within the sample of 80 subjects, two subsets of these subjects exhibited distinctly different patterns of results. Though these two groups did not differ in their overall mean identification performance, the members of these two groups used the response categories in a different manner. An appreciation of these results can be beneficial to the teacher of English as a second language, as pedagogical methods that take into account the observed individual differences in sensitivity and bias can meet student needs more effectively.

REFERENCES

[1] Anderson-Hsieh, J. (1992). Using electronic visual feedback to teach suprasegmentals. *System*. 20:1, 51-62.

[2] Bradlow, A, Pisoni, D. Yamada, R. & Tohkura, (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*. 101 (4), 2299-2310.

[3] Flege, J. E. (1988). The production and perception of foreign language speech sounds. In: Human communication and its disorders, edited by H. Winitz, pp. 224-401, New Jersey: Ablex.

[4] Flege, J. Takagi, N. & Mann, V. (1995). Japanese adults can learn to produce English /r/ and /l/ accurately. *Language and Speech* 38 (1) 25-55.

[5] Fowler, C. (1986). An event approach to the study of speech perception from a direct-realist perspective, *J. Phonetics* 14, 3-28.

[6] Jonassen, D. & Grabowski, B. (1993). *Handbook of individual differences, learning, and instruction.* New Jersey: Lawrence Erlbaum Associates.

[7] Lambacher, S., Martens, W., Nelson, B. & Berman, J. (1997). Perception of English voiceless fricatives by native speakers of Japanese. In J. Leather & A. James (Eds.), *NEW SOUNDS 97: Proceedings of the Third International Symposium on the Acquisition of Second Language Speech.* University of Klagenfurt, Austria.

[8] Liberman, A. & Mattingly, I. (1989). A specialization for speech perception, *Science* 245, 489-494.

[9] Lively, S., Logan, J., and Pisoni, (1993). Training Japanese listeners to identify English /r/ and /l/. ll: The role of phonetic environments and talker variability in learning new perceptual categories.*Journal of the Acoustical Society of America*. 94, 1242-1255.

[10] Molholt, G. (1990). Spectrographic analysis and patterns in pronunciation, *Computers and the Humanities*, 24, 81-92.

[11] Redmond, L. D. (1977). Learning to recognize foreign speech sounds: Strategies and auditory processing. Unpublished Ph.D. thesis, University of California, California.

[12] Trubetzkoy, N. (1939/1969). Principles of phonology. Translated by C. A. Baltaxe, Berkeley, CA University of California Press.