THE EFFECTS OF PHONETIC DISTANCE, LEARNING CONTEXT AND LEARNER PROFICIENCY ON L2 PERCEPTION OF ENGLISH LIQUIDS

Sally Chen and Janice Fon

National Taiwan University
{d93142002,jfon}@ntu.edu.tw

ABSTRACT

This study aims to investigate the effects of phonetic distance, learning context and learner proficiency on L2 perception of English liquids. Reaction time difference between the pre- and post-tests was analyzed. Results showed that the natural context induced the most progress for participants of a lower L2 proficiency level, while no preference was shown for those of a higher proficiency level. In general, L2 learners showed more progress for liquids occurring in novel phonotactic structures. Phone effect was significant only when L2 learners of lower proficiency perceived liquids in the singleton position.

Keywords: L2 perception, phonetic distance, learning context, SLM, PAM.

1. INTRODUCTION

Perception of unfamiliar L2 sounds has always been an obstacle that learners of the language have to strive to conquer. Both the speech learning model (SLM) [9, 10] and the perceptual assimilation model (PAM) [2, 4, 5] predict the degree of success listeners have in perceiving nonnative sounds is based on the perceived phonetic distance between L1 and L2 sounds. The two models differ in that the SLM focuses on the aspect of L2 learning, while the PAM is established on the fact that some foreign sounds are harder to perceive than others [12].

Regarding language contrasts, the SLM proposes that the greater the perceived dissimilarity between an L2 sound and its closest L1 counterpart, the more likely a new category is to be formed for the L2 sound [9, 10, 12]. Modeling from another perspective, the PAM proposes that sounds in a foreign language are perceived in accordance with their similarities to the closest native language sounds articulatorily [2, 4, 5, 12]. When two sounds are assimilated to two different categories, the perceptual performance is expected to be good. However, when two sounds are assimilated to one single category, the performance will be unsatisfactory [16]. In addition, based on a logic similar to the Native Language Magnet theory [14, 15], when two sounds are assimilated to the same category, it is believed that in cases where only one of the two sounds is regarded as a good exemplar of that target category, the performance will be much better than in those where both sounds are regarded as comparable candidates for the same target category [12].

2. AIMS OF THE STUDY

There are three specific aims of this study. The first is to examine the effect of phonetic distance. The two English liquids, /l/ and /r/, impose differential levels of difficulty on Mandarin speakers [7]. The former has a closer counterpart in Mandarin than the latter ([l̪] *vs.* [z̪]). According to the SLM, this would predict that L2 listeners are more likely to form perceptual categories for /r/ than for /l/. On the other hand, according to the PAM, both categories should be perceived equally well. This study investigates which model better interprets the perception of English liquids by Mandarin speakers of English.

The second aim is to investigate the effect of learning contexts. In Taiwan, the audiolingual method is used very often. Students learn new words and sentences by repeating after the instructor, often without contexts. However, in recent years, more and more EFL instructors, especially those outside the regular school system, are using some combination of language immersion strategies [6], believing that this can promote greater success in L2 learning. Therefore, it would be interesting to see whether the presentation method of the stimuli affects L2 sound perception.

Finally, this study looks into the effect of speaker proficiency. As listening is not a skill that

is much emphasized in Taiwan's mandatory education system, it is often hard to assess accurately students' listening comprehension abilities. However, it is worth investigating whether differences in general language proficiency affect L2 sound perception after a short-term training period.

3. METHOD

3.1. Participants

135 participants were recruited for this study, including 27 native speakers of English (with mean age of 22.89 years old and standard deviation of 4.31) and 108 EFL college students (with mean age of 21.00 years old and standard deviation of 1.46) with Mandarin as their native language. Half of the EFL learners were of high proficiency level and the other half were of low proficiency level as determined by the General English Proficiency Test (GEPT), a national English proficiency test held annually by the Language Training and Testing Center. The high proficiency level group passed the high-intermediate level of the GEPT while the low proficiency level group did not.

3.2. Materials

56 monosyllabic words were chosen from *the 1000 Basic English Vocabularies for Elementary and Junior High School Students* [8] announced by the Ministry of Education in Taiwan. These words either begin with an /l/ or /r/ (e.g. *last* and *rain*) or a Cl- or Cr- cluster [e.g. *climbed* and *broke*]. Another 56 phonotactically-matched monosyllabic pseudo words were also included as fillers (e.g., *laped* and *reim*). In total, there were 56 (targets) + 56 (fillers) = 112 stimuli. All materials were recorded by a female speaker of American English. Recordings were done in a sound-treated room with a sampling rate of 48 kHz and, were later downsampled to 16 bit 22050 kHz using Adobe Audition 2.0.

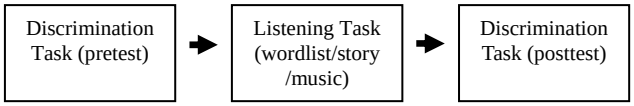
3.3. Equipment

A SONY DAT PCM-M1 digital recorder along with a Maxell DM120 tape and a SHURE SM10A-CN head-mounted microphone were used for recording stimuli. E-Prime 1.0 and its accompanying PST Serial Response Box (SR Box) Model 200a were used for the perception tasks.

3.4. Procedure

Participants were seated in a quiet room and were given a three-block perception task, as shown in Figure 1.

Figure 1: Procedure of the experiment.



The first block was a 2-alternative-force-choice (2AFC) speeded perception task on target words. Participants were instructed to respond by pressing the corresponding button on the SR Box as accurately and quickly as possible. The block lasted about 5 minutes.

The second block was a listening task, in which participants were randomly assigned to one of the three audio conditions, i.e. wordlist, story, and music (Table 1). The wordlist condition contained all 56 target words (inter-stimulus interval = 3s), and the story condition was a tailor-made children's story containing all 56 target stimuli. All other words in the story were carefully chosen so as to avoid any /l/ and /r/ sounds. The music condition was a piece of classical music and served as a control. All three conditions lasted for 3 minutes and 30 seconds. Half of the L2 speakers listened once (Short version) while the other half listened twice (Long version). Native speakers only listened to the short version to avoid boredom.

Table 1: Excerpts of the 3 audio conditions in the listening task.

Audio condition	Excerpt of the content
Wordlist	**Black, round, live, group, place, love, rice, like, read, play, dream**….
Story	My name is Andy. I am a mouse but I am big and **black** and **round**. I **live** with my mom and a **group** of mice in a nice **place**. We **love** to eat **rice**. We **like** to **read** books and **play** games. Each day I **dream** of catching a cat….
Music	(Handel's *Le Rejouissance*)

The third block was again a 2AFC speeded perception task, as in Block 1. However, in this posttest, both target words and fillers were included. The block lasted about 10 minutes.

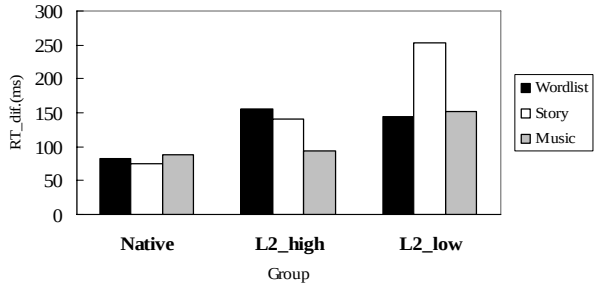
4. RESULTS

4.1. General contrast

A Proficiency (3) × Task (3) × Position (2) × Phone (2) four-way ANOVA was performed for the Short version on RT difference between the pre- and post- tests to examine the effect of listening conditions. Results showed that except for Phone, all the other main effects were significant [Proficiency: F (2, 4285) = 41.23, p < .0001; Task: F (2, 4285) = 8.39, p < .001; Position: F (1, 4285) = 38.69, p < .001]. The two-way interaction between Proficiency and Task was also significant [F (4, 4285) = 8.53, p < .0001]. The three-way interaction concerning Group, Position, and Phone was marginally significant [F (2, 4285) = 2.85, p = .06]. The four-way interaction was not significant.

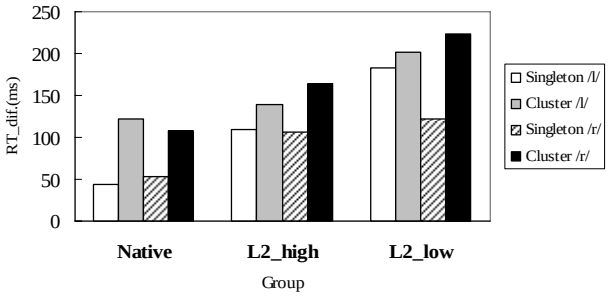
Figure 2 shows the interaction between L2 proficiency levels and different audio contexts in the listening task. Higher bars indicate more improvement. Given that native speakers were already fast in the pretest, they made relatively less progress between the pretest and the posttest. Post-hoc analyses of the two-way interaction showed that there was no significant difference among the three tasks for the native speakers. For L2 speakers of higher English proficiency, listening to music yielded the least progress (p < .01), while for those of lower English proficiency, the story condition yielded the best results (p < .0001).

Figure 2: Interaction between English proficiency and the audio tasks assigned.



Regarding the three-way interaction, as shown in Figure 3, Position is significant for all three proficiency groups. More progress was found for liquids in the cluster position [p < .01]. Phone contrast was significant only when L2 learners of lower proficiency perceived liquids in the singleton position: the sound /l/ was perceived with more progress than the sound /r/ between the pre- and post-tests [p < .05].

Figure 3: RT difference between /l/ and /r/ in different positions across 3 proficiency groups.

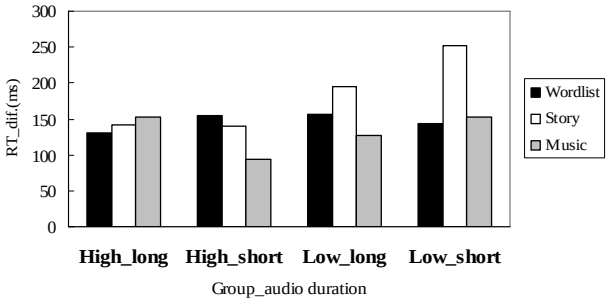


4.2. Durational contrast

A Proficiency (2) × Task (3) × Duration (2) × Position (2) × Phone (2) five-way ANOVA was performed on nonnative speakers to investigate whether the duration of audio input affected the RT difference between two perception tests. Results showed that Duration had a 2-way interaction with Proficiency [F (1, 5579) = 4.28, p < .05] and a three-way interaction with Proficiency and Task [F (2, 5579) = 4.66, p < .01].

Post-hoc analyses showed that for L2 speakers with a higher proficiency, Duration and Task had interactions: As shown in Figure 4, those who listened to the Short version in the listening task made more progress after given the "wordlist" or the "story" in the listening block (p < .01). On the other hand, for the lower proficiency group, Duration and Task effects were found significant. Longer listening times reduced their progress. In addition, those who listened to the "story" in the listening block showed the most progress, as previously mentioned in 4.1.

Figure 4: Interaction among English proficiency, Duration, and Task.



5. DISCUSSION

Unlike Japanese EFL speakers [1, 11], the two different liquid sounds were not perceived differently for Mandarin EFL learners; however, their interaction with phonotactic structures (singleton *vs.* cluster) did conform to the SLM in that for novel combinations, such as liquids in consonant clusters, L2 speakers were able to make more progress (same trend as native speakers) after being exposed to the listening task for a few minutes. In addition, the fact that those of lower L2 proficiency level made more progress on /l/ – the phone considered as a good exemplar to their L1 – supported the PAM.

Results in this study also showed that different learning contexts did facilitate L2 speakers' perception of L2 sounds to different extents. On top of that, L2 speakers' proficiency level also played a role – those of a higher proficiency level made comparable progress regardless of whether the listening material was natural or mundane; while those of a lower L2 proficiency level required that the material be interesting enough to keep them attentive. One explanation is that the basic linguistic unit for speech processing could be lexical items, instead of segments. As the lexical restructuring model proposes, as one becomes more experienced and gains better understanding of the relationship among lexical items, one gradually acquires better segmental representations [3, 13]. Considering that L2 speakers with a lower proficiency might not be as experienced with L2 words as those with a higher proficiency, providing a temporary artificial setting in which they could more easily define the distinctions among the stimulus words (such as the "story") thus boosted their performance on the differentiation of the two liquid sounds.

In sum, results of this study clearly indicated that all fundamental differences between L1 and L2 phonological systems need to be taken into consideration when designing EFL training materials. Further studies are required to claim whether the most fine-grained L2 perception unit should be of the segmental or the lexical level.

6. REFERENCES

[1] Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics 32*, 233-250.

[2] Best, C. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*, 171-204. Timonium, MD: York Press.

[3] Bradlow, A. R. & Pisoni, D. B. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors. *Journal of the Acoustical Society of America* 106 (4), 2074-2085.

[4] Best, C. T., McRoberts, G. W., Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listeners' native phonological system. *Journal of the Acoustical Society of America* 109 (2), 775-792.

[5] Best, C. T., McRoberts, G. W, Sithole, N. (1988). The phonological basis of perceptual loss for non-native contrasts: maintenance of discrimination among Zulu clicks by English speaking adults and infants. *Journal of Experimental Psychology, Human Perception Performance* 14, 345-360.

[6] Chan, Y. -j., Lin, S. -H. (2001). The empirical study of whole language teaching in English class of junior high. *Journal of Taiwan Normal University* 46 (2), 233-253.

[7] Chen, S., Fon, J. (2005). To play or to pray?–The influences of initial exposure age and environment on perception of Mandarin speakers of English on devoiced liquids [Abstract]. *Journal of the Acoustical Society of America*, 118 (3), Pt. 2, 1898-1899.

[8] Dept. of Elementary Education, Ministry of Education, Taiwan. http://www.boe.tyc.edu.tw/1000words/%B3%CC%B0%F2%A5%BB%A4@%A4d%A6r%B5%FC.htm

[9] Flege, J. E. (1995). Second Language Speech Learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*. 233-277. Timonium, MD: York Press.

[10] Flege, J. E., Birdsong, D., Bialystok, E., Mack, M., Sung, H., Tsukada, K. (2006). Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics* 34, 153-175.

[11] Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "l" and "r". *Neuropsychologia* 9, 317-323.

[12] Guion, S. G., Flege, J. E., Akahane-Yamada, R., Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America* 107 (5), 2711-2724.

[13] Imai, S., Walley, A. C., Flege, J. E. (2005). Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners. Journal of Acoustical Society of America 117 (2), 896-907.

[14] Kuhl, P. K. (1993). Innate Predictions and the effect of experience in speech perception: the native language magnet theory. In B. de Boysson-Bardies et al. (eds.), *Developmental neurocognition: Speech and face processing in the first year of life*, 259-274. Dordrecht: Kluwer Academic Publishers.

[15] Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Science*, 97, 11850-11857.

[16] Sebastian-Galles, N., Soto-Faraco. (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition* 72, 111-123.