The Efficacy of EPG Assisted L2 Pronunciation Instruction: An Audio-Perceptual Analysis of the Speech of Native Japanese Learners of English

Emily Louise Peterson

Brigham Young University

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

Part of the Education Commons

BYU ScholarsArchive Citation

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.
The Efficacy of EPG Assisted L2 Pronunciation Instruction: An Audio-Perceptual Analysis
of the Speech of Native Japanese Learners of English

Emily Louise Peterson

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of
Master of Science

Shawn Nissen, Chair
Douglas Petersen
K. James Hartshorn

Department of Communication Disorders
Brigham Young University

Copyright © 2020 Emily Louise Peterson
All Rights Reserved
ABSTRACT

The Efficacy of EPG Assisted L2 Pronunciation Instruction: An Audio-Perceptual Analysis of the Speech of Native Japanese Learners of English

Emily Louise Peterson
Department of Communication Disorders, BYU
Master of Science

As there is a clear correlation between one’s degree of proficiency in the English language and one’s subsequent financial compensation in the workplace setting and in interpersonal relationships, improving one’s speaking abilities can be highly valuable from both a financial and emotional perspective. This study examines the efficacy of an electropalatography (EPG) assisted pronunciation training program in helping native Japanese learners of English acquire and improve the /r/ and /l/ sound contrast in American English, as rated by a group of listeners. Additionally, it evaluates whether or not the degree of improvement varied across word position, task type, or assessment period. Four native Japanese speaking learners of English participated in a four-week program which included seven 45-minute training sessions enhanced with visual biofeedback from the EPG Samples of their productions of the target phonemes were obtained at baseline, at posttreatment, and at follow-up assessment periods. Using a visual analogue scale, 36 adult listeners listened to these recordings and provided comparative auditory perceptual ratings. Overall, subjects showed greater improvement in their production of the phoneme /l/ than in the phoneme /r/. Phoneme-specific patterns emerged in terms of word position, task type, and assessment period. For the phoneme /l/, more improvement was seen in final position than initial position, more improvement was seen in nonsense syllables than in words, and improvements were maintained across posttreatment to follow-up assessment periods. For the phoneme /r/, roughly equal levels of improvement were seen across word position, while greater improvement was seen in the context of words in sentences than in nonsense syllables, and posttreatment showed greater levels of improvement than did follow-up assessment periods. These results are promising as it indicates that EPG assisted pronunciation training may be an effective vehicle to help L2 English language learners acquire and improve their productions of the /r/-/l/ phonemic contrast. This is significant, as the Japanese L2 population has typically been found to be highly resistant to more traditional forms of intervention.

Keywords: electropalatography, Japanese, second-language acquisition, accent reduction, liquid contrast, auditory perceptual ratings
ACKNOWLEDGMENTS

I am very grateful to everyone who has helped me throughout the course of this project and contributed to its success. Dr. Nissen, my committee chair, has been a valuable mentor and I greatly appreciate his optimistic nature and dedicated help in guiding me through the process of both graduate school and this thesis. I owe gratitude as well to Dr. Hartshorn and Dr. Petersen, whose feedback and insights have helped shape the direction of this work. I am grateful to have worked with Summer Price and Nicole McDougal in this project, and am grateful both for their friendship and for all of their hard work in designing and implementing the EPG Assisted Pronunciation Training program that was used to train the Japanese speakers. To the many students and friends who gave perceptual ratings in this study, I thank you as well. The research assistants, Emily Wright and Mac Christensen, were invaluable in helping me collect data from participants and also in organizing the data. I’m grateful for my cohort and for all of the professors who have mentored me throughout this program. I am also grateful for my family, who have supported me in my dreams. Finally, I am grateful to God and for His help in making my graduate school experience and this project a reality.
TABLE OF CONTENTS

TITLE PAGE ................................................................................................................................... i
ABSTRACT .................................................................................................................................... ii
ACKNOWLEDGMENTS ............................................................................................................. iii
TABLE OF CONTENTS ............................................................................................................... iv
LIST OF TABLES ......................................................................................................................... vi
LIST OF FIGURES ...................................................................................................................... vii
DESCRIPTION OF THESIS STRUCTURE AND CONTENT ..................................................... viii
Introduction ......................................................................................................................................1
  Approaches to L2 Instruction .....................................................................................................2
    Auditory methods ................................................................................................................2
    Visual methods ....................................................................................................................3
  Japanese Learners of English .....................................................................................................7
  Purpose of this Study .................................................................................................................7
Method .............................................................................................................................................9
  Participants .................................................................................................................................9
  Instruction ................................................................................................................................10
  Recordings ................................................................................................................................13
  Stimuli ......................................................................................................................................13
  Procedures ................................................................................................................................14
  Statistics ...................................................................................................................................16
Results ............................................................................................................................................17
  MK ...........................................................................................................................................17
LIST OF TABLES

Table 1  *Listener Demographic Information* .................................................................11

Table 2  *Perceptual Rating Means and Standard Deviations as a Function of Task Type and*  
*Word Position for Speaker MK’s /l/ Productions* ........................................................18

Table 3  *Perceptual Rating Means and Standard Deviations as a Function of Task Type and*  
*Word Position for Speaker MK’s /r/ Productions* ........................................................20

Table 4  *Perceptual Rating Means and Standard Deviations as a Function of Task Type and*  
*Word Position for Speaker NH’s /l/ Productions* ........................................................22

Table 5  *Perceptual Rating Means and Standard Deviations as a Function of Task Type and*  
*Word Position for Speaker NH’s /r/ Productions* ........................................................23

Table 6  *Perceptual Rating Means and Standard Deviations as a Function of Task Type and*  
*Word Position for Speaker TA’s /l/ Productions* ........................................................25

Table 7  *Perceptual Rating Means and Standard Deviations as a Function of Task Type and*  
*Word Position for Speaker TA’s /r/ Productions* ........................................................27

Table 8  *Perceptual Rating Means and Standard Deviations as a Function of Task Type and*  
*Word Position for Speaker YM’s /l/ Productions* ........................................................28

Table 9  *Perceptual Rating Means and Standard Deviations as a Function of Task Type and*  
*Word Position for Speaker YM’s /r/ Productions* ........................................................31

Table 10  *Auditory Perceptual Rating Means as a Function of Task Type and Word Position for*  
*All Speakers’ /l/ Productions* ..........................................................................................32

Table 11  *Perceptual Rating Means as a Function of Task Type and Word Position for All*  
*Speakers’ /r/ Productions* ...............................................................................................33
LIST OF FIGURES

Figure 1. Visual Analogue Scale used by listeners to submit auditory perceptual ratings ...........15

Figure 2. Mean listener ratings comparing MK’s productions of /l/ across linguistic context,
word position, and assessment period.................................................................18

Figure 3. Mean listener ratings comparing MK’s productions of /r/ across linguistic context,
word position, and assessment period.................................................................20

Figure 4. Mean listener ratings comparing NH’s productions of /l/ across linguistic context,
word position, and assessment period.................................................................22

Figure 5. Mean listener ratings comparing NH’s productions of /r/ across linguistic context,
word position, and assessment period.................................................................23

Figure 6. Mean listener ratings comparing TA’s productions of /l/ across linguistic context,
word position, and assessment period.................................................................25

Figure 7. Mean listener ratings comparing TA’s productions of /r/ across linguistic context,
word position, and assessment period.................................................................27

Figure 8. Mean listener ratings comparing YM’s productions of /l/ across linguistic context,
word position, and assessment period.................................................................28

Figure 9. Mean listener ratings comparing YM’s productions of /r/ across linguistic context,
word position, and assessment period.................................................................30

Figure 10. Mean listener ratings comparing aggregate productions of /l/ across linguistic
context, word position, and assessment period.....................................................31

Figure 11. Mean listener ratings comparing aggregate productions of /r/ across linguistic
context, word position, and assessment period.....................................................32
DESCRIPTION OF THESIS STRUCTURE AND CONTENT

This thesis, *The Efficacy of EPG Assisted L2 Pronunciation Instruction: An Auditory-Perceptual Analysis of the Speech of Native Japanese learners of English*, is part of a larger study exploring the effectiveness of EPG assisted English as a second language (L2) pronunciation instruction. Portions of this thesis may be submitted for publication, with the thesis author being included in the list of contributing coauthors. An annotated bibliography is provided in Appendix A, and the consent form used in this study is provided in Appendix B.
Introduction

As English is widely known as the language of science, business, and medicine, many advantages are gained by acquiring proficiency in its use. In fact, it has even been suggested that learning English is becoming not only an advantage, but a basic and necessary skill, and currently, there are at least one billion individuals who speak or are learning to speak English as a second language (L2; Education First, 2018).

Several social and emotional benefits occur as a result of developing a high level of proficiency in English as an L2, especially when one’s accent is perceived as native-like. Research suggests that when controlling for other factors, the degree of an L2 learner’s speech proficiency impacts the way native English speakers perceive the English learner’s credibility (Lev-Ari & Keysar, 2010). Not only are foreign accented speakers of English perceived as being less credible and truthful than their non-accented counterparts, research also suggests that foreign accented speech negatively impacts affective and behavioral interactions with others in the workplace. Russo, Islam, and Koyuncu (2017) suggested that managers at work may view accented L2 speakers with less positive regard, showing fewer displays of acceptance, warmth, and value, and that this may result in accented L2 speakers feeling excluded, devalued, and unappreciated.

The level of English proficiency of L2 language learners has been found to be directly correlated with their earning potential. McManus, Gould, and Welch (1983) hypothesized that the wage gap typically associated with an individual’s ethnicity, country of birth, and time spent living in the United States is actually caused by differences in English language proficiency. Additionally, Day and Shin (2005) performed a study based on results of the U.S. Census and found a correlation that individuals with the lowest levels of English also experienced the lowest
rates of employment and median earnings. Furthermore, a wage gap of $7,000 existed between individuals speaking English “very well” and “well.” Individuals with the highest levels of English abilities had the highest earnings, approaching the earnings of native English speakers (Day & Shin, 2005). A non-native accent and lower English language skills can not only impact an individual’s wages and earnings, but also one’s career satisfaction and advancement. A model created by Russo et al. (2017) suggested that managers of accented L2 speakers, due to reduced perceptions of cognitive fluency, may hold negative perceptions of accented L2 speakers including lower expectations of their job performance abilities. Furthermore, they posit that the managers may subsequently assume a controlling management style and assign unimportant job tasks to the accented L2 speakers, which could lead to less career satisfaction and advancement for the accented L2 speaker.

There is a clear need to develop intervention procedures that can improve English in L2 speakers. Several intervention approaches have been researched, with mixed results for native Japanese L2 English learners.

**Approaches to L2 Instruction**

**Auditory methods.** Specific methods for pronunciation training include a wide range of techniques. Auditory discrimination training (Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999) has been used to help L2 learners improve their auditory discrimination as a prerequisite to discriminately producing sounds. Hazan, Sennema, Iba, and Faulkner (2005) utilized a hybrid approach of audiovisual perceptual training to help L2 learners improve their auditory discrimination. Other techniques, as described by Celce-Murcia, Brinton, and Goodwin (1996), include that of listening and imitating in order to develop stronger perceptual and productive skills, minimal pair drills to aid in categorization of distinct phonemes, and contextualized
minimal pairs to assist L2 learners in the generalization of phonemic categorization. Additionally, visual aids, tongue twisters, and developmental approximation drills have also been used. As correct pronunciation includes not only the correct production of consonants, but also vowels, prosody, and intonation, other traditional methods include practicing vowel production, altering stress patterns, and reading aloud. To help the L2 learner develop greater self-awareness, some clinicians also record speech productions for later auditory feedback and review.

Traditional methods for L2 pronunciation training include programs such as Compton Pronouncing English as a Second Language Program (Compton P-ESL) and the Speech Perception Assessment and Training System (SPATS-ESL). Compton P-ESL is an approach used to assess and evaluate the pronunciation of non-native speakers, and also includes designing evidence-based training programs, web-based tools, and post-training analysis. SPATS-ESL is a software system that helps individuals test and train bottom-up and top-down speech perception skills. It was initially designed to be used by individuals with hearing loss, but has been modified to be used by English L2 learners. While these auditory and traditional methods have helped many L2 learners improve their English pronunciation, additional recent advances have shown further ways to aid individuals who may not benefit strongly from traditional auditory approaches alone.

**Visual methods.** Recently, clinicians and researchers have begun to implement various forms of biofeedback to help L2 language learners visualize their articulatory movement patterns during sounds which typically are otherwise difficult to conceptualize.

Ultrasound has been used as a form of biofeedback to help both L1 and L2 speakers visualize their tongue movements. Specifically, it has also been studied in its efficacy to help
Japanese speakers visualize their tongue movement from sagittal and horizontal angles as they attempt to correctly produce the /r/-/l/ phoneme contrast (Tsui, 2012). However, Gick, Bernhardt, Bacsfalvi, and Wilson (2008) note several core issues and limitations in ultrasound research, including a lack of standardization in measurement protocol across researchers. Additionally, Gick et al. mentioned that with a temporal resolution of only 30 frames per second, ultrasound is too slow to adequately capture some types of movement.

Electropalatography (EPG) has been used as a tool to aid in both native language remediation and in L2 instruction. It has been used to treat children with persistent sound system disorders (Dent, Gibbon, & Hardcastle, 1995). Within the field of L2 pronunciation instruction, EPG has been researched as a tool to aid in pronunciation training in native speakers of Thai (Schmidt & Beamer, 1998), Spanish (Bright, 1999), Korean, Taiwanese, Chinese, and Japanese (Li, 2017), and Japanese (Gibbon, Hardcastle, & Suzuki, 1991).

Schmidt and Beamer (1998) were two of the early researchers to study the role of EPG in pronunciation training. In this case study, three native Thai English L2 learners received EPG treatment targeting the phoneme contrasts /s/-/ʃ/, /t/ - /θ/, and /l/-/ɹ/. Participants attended 45-minute sessions at a frequency of twice per week between 39 to 48 times; the sessions included conversational practice, testing on paired contrast productions, and practice of phonemic contrasts with visual support from the EPG and guidance from the clinician. As a result, each of the participants demonstrated articulatory changes which applied to both previously learned sounds as well as to new sounds. The authors suggested that even if L2 language learners had habituated errored consonantal productions over the course of many years, that EPG could provide feedback necessary to help change and modify the errored patterns, and that this progress can occur more quickly with support of EPG than with traditional types of treatment.
Bright (1999) researched the impact of EPG instruction on native Spanish speaking L2 learners of English. In this case study, which included three participants, each of the native Spanish speakers attended twelve 90-minute sessions (at a frequency of three sessions per week). During the training sessions, both the clinician and treatment participants used EPG palates. The clinician provided training on correct lingual placement for each of the two target phonemes and presented models to help the participants further develop their auditory discrimination skills. Then, the clinician provided feedback on each of the phonemes produced by the participants and provided models when needed. After the participants established a contrastive pronunciation pattern, the L2 learners practiced self-correction with each other. Phonemes were practiced first as minimal pairs in isolation until reaching 90% accuracy, then targets were practiced at the word level, and finally, at the sentence level. The *Comp*, a test which consists of background history in learning English, reading a list of words aloud, a spontaneous speech sample, and an oral reading passage, was administered to participants both before and after training. Each participant demonstrated progress in reducing accent in target phonemes at both the word and sentence level. As measured by composite scores on “goodness of pronunciation” (of words) from 10 listeners, a least mean square comparison revealed an overall improvement in standardized scores. A statistical test of aggregate baseline and posttreatment data across participants was significant, the two-tailed T tests for individual participants were not statistically significant. From this study, researchers concluded that EPG could be used as an effective tool to help L2 learners improve their discrimination and production of target phonemes.

Li (2017) sought to further investigate the L2 learner’s experience in and perceptions of EPG assisted pronunciation training. Researchers collected both quantitative and qualitative data
from nine L2 English learners from Korea, Taiwan, China, and Japan. Over seven weeks, each of the L2 English learners participated in twelve 40-minute sessions of pronunciation instruction. Overall, the participants reported that they perceived the EPG assisted pronunciation instruction as an effective tool to help them become more aware of correct lingual tongue placement. Additional positive feedback included perceptions that EPG would be socially acceptable, that the software was user-friendly, and that they appreciated the attention and organization of their instructors. Some participants reported initial articulatory distortions as they adjusted to the feeling of the sensor in their mouths. One participant reported that it hurt the upper side of their mouth, while another participant reported that it hurt their tongue, “but not too bad.” One participant reported that the initial discomfort did not last, while another participant reported that they did not feel discomfort while wearing the EPG sensor.

Gibbon et al. (1991) researched the efficacy of EPG as a treatment tool to aid Japanese L2 English learners in the acquisition of the /r/-/l/ phonemic contrast. Each of the two participants read a list of words containing /l/ and /r/, including a list of minimal pairs. Then, the participants read a list of words in Japanese which included examples of the Japanese /r/ in varying phonetic contexts. Finally, the speakers participated in four treatment sessions which included visual biofeedback from the EPG device. Treatment focused on awareness, articulatory patterns, establishing the sound in different vowel contexts, and in producing the sound in complex phonetic contexts. After treatment, a follow-up recording was taken of each participant as they read the same word lists that they had read prior to the initiation of the study. Before intervention began, the Japanese learners of English demonstrated many similarities between their productions of /l/ and /r/; additionally, they produced a variety of patterns for the Japanese /r/ phoneme. After treatment, the participants demonstrated the ability to make consistent
contrasts between the liquids /r/ and /l/ in controlled contexts, and they also seemed to be better able to identify and change their articulatory movements for the liquids /r/ and /l/. Gibbon et al. concluded that further research is warranted into the impact of EPG as a tool for pronunciation training, especially as to its impact on helping improve phonemic contrasts in fluent speech.

**Japanese Learners of English**

One group of speakers that commonly study English as an L2 are native Japanese speakers. Even though Japan has an established history of teaching English in schools, Japan’s average score on the TOEFL, a test of English proficiency, has been the second lowest in Asia. (Education Testing Service, 2016). It is posited that this may be due partially to the traditional utilization of the grammar-translation method of English instruction in public schools which focuses on “teaching for the test,” rather than communication-based approaches that focus on teaching for developing communicative competence (Morita, 2017). In fact, a comparative study of native Japanese L2 English learners and teachers of English in Japan showed that while over 90% of students indicated that it was important to learn correct pronunciation (Matsuura, Chiba, & Hilderbrandt, 2001) and preferred to learn English by watching or listening to native speakers of English (Hayashi & Cherry, 2004) less than 70% of teachers at the university level reported teaching pronunciation in their classrooms (Matsuura et al., 2001).

One English phonemic sound contrast with which native Japanese speakers often struggle is the /r/-/l/ contrast. To further elucidate the ways in which members of the Japanese population categorically perceive the liquids /r/ and /l/, a study was conducted by Miyawaki et al. (1975). The researchers noted that when Japanese adults were presented with a series of synthetic sounds along a continuum ranging from /ra/ to /la/, that the adults’ categorical rating scores remained near chance, indicating that they did not detect any change or difference along the /ra/ to /la/
continuum. Rather, the participants all perceived the sounds as the Japanese “r.” Further research has shown that while native Japanese speakers often assimilate the English phonemes /r/ and /l/ into the same phonemic category, the category assimilation for this contrast is slightly asymmetrical; the liquid phoneme /l/ is more strongly assimilated into the category of the Japanese apicoalveolar tap than is the liquid phoneme /r/ (Hattori & Iverson, 2009). This may be due partially to the fact that although infants can discriminate phonetic contrasts in all languages until the age of around three months, there is a decline in an infant’s ability to accurately perceive consonants in foreign languages around the age of 11 months. By adulthood, adults cannot always discriminate between the phonetic units used in all languages (Kuhl, 2004). As the Japanese language’s phonemic categories for speech sounds do not include a classification of /r/ and /l/ as separate phonemes, Japanese infants are less likely to be exposed to the /r/-/l/ phonemic contrast, which limits the probability of maintaining the ability to accurately discriminate between these two target phonemes through adulthood. When Japanese children typically have the opportunity to begin learning English in grade school, many of their English teachers have learned English as an L2 and thus may not produce a native-like model of the target phonemic contrast. Thus, their exposure to develop the categorical perception of the /l/-/r/ phonemic contrast continues to be limited.

**Purpose of this Study**

As the demand for efficient L2 language instruction and pronunciation training increases, this study is designed to evaluate the efficacy of EPG, a relatively new form of visual biofeedback, to assist native Japanese speakers in the acquisition of the /l/-/r/ phonemic contrast. This study will specifically evaluate the following research questions:
1. Does EPG assisted instruction help native Japanese L2 English learners to acquire the /r/-/l/ phonemic contrast, as measured by an auditory perceptual analysis?

2. Does the efficacy of EPG assisted instruction vary as a function of linguistic task type?

3. Does the efficacy of EPG assisted pronunciation training vary as a function of word position?

**Method**

**Participants**

Speech recordings evaluated in this study were produced by four Japanese learners of English who ranged in age from 19 to 23 years, with a mean age of 21.2. The group included two males and two females, all of whom had normal dentition and with oral cavity anatomy deemed sufficient for functional speech production. No speech or hearing problems were reported by any of the participants. All of the speakers enrolled in this study were simultaneously enrolled at the English Language Center at Brigham Young University. While participating in this study, the Japanese L2 English learners also participated in four 65-minute classes per day at the English Language Center, which focused on speaking, listening, reading, and writing.

The speech sound recordings were evaluated by 36 native English-speaking listeners, ranging in age from 19 to 31 years, with a mean age of 23.9 years. At the time of participating in the auditory perceptual rating task, each of the participants signed a consent form and passed a pure-tone air conduction hearing screening at octave frequencies between 500 and 8,000 Hz. Although three participants reported a history of speech sound delay, none of the listeners presented with any current speech or language disorders. Six of the listeners were monolingual English speakers, while 27 of the listeners spoke one or more
additional languages, including Arabic, American Sign Language (ASL), Bulgarian, Chinese, Japanese, Portuguese, Russian, and Spanish. As delineated in Table 1, nine of the listeners were currently enrolled in a phonetics course, while 15 had completed one prior to the study and 12 of the listeners had never been enrolled in a phonetics course. As data obtained from visual analogue scales can be impacted by the handedness of the raters, demographic data on handedness was obtained. Thirty-five of the listeners were right-handed, while one of the listeners was left-handed. The Institutional Review Board at BYU has reviewed and approved all procedures used in this study.

**Instruction**

Each of the four Japanese speakers participated in individual treatment sessions. Utilizing a hybrid approach, the sessions included auditory modeling, audio-perceptual training, and visual biofeedback from an EPG device. For a total of four weeks, participants received two 45-minute sessions per week. For the first five minutes of the session, the participants were given time to adapt to the EPG device. Then, for the next five minutes, the speakers practiced the phonemes /r/ and /l/ in isolated nonsense words and were also invited to auditorily distinguish between the /r/-/l/ phonemic contrast. During this time, they had the support of EPG software, which allowed the speakers to see real-time biofeedback of their own lingual palatal contact patterns in contrast to an idealized contact pattern model. For the next 10 minutes, the speakers practiced correctly producing target phonemes within the context of words in sentences with the support of the EPG device. Every two sessions, the instructional stimuli used to elicit the target phonemes (such as minimal word pairs, sentences, and spontaneous speech) were changed (McDougal, 2019). After practicing phoneme production in tandem with visual biofeedback support, the EPG device was
Table 1

**Listener Demographic Information**

<table>
<thead>
<tr>
<th>ID</th>
<th>Gender</th>
<th>Age</th>
<th>Dominant Hand</th>
<th>Languages Spoken in Addition to English</th>
<th>Phonetics Experience</th>
<th>aHearing Screening</th>
<th>bSpeech Disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>24</td>
<td>R</td>
<td>Portuguese</td>
<td>Current</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>22</td>
<td>R</td>
<td>Spanish</td>
<td>Current</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>31</td>
<td>R</td>
<td>Spanish</td>
<td>Yes</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>26</td>
<td>L</td>
<td>Spanish</td>
<td>Yes</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>25</td>
<td>R</td>
<td>Spanish</td>
<td>Yes</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>24</td>
<td>R</td>
<td>Spanish</td>
<td>Yes</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>23</td>
<td>R</td>
<td>Russian</td>
<td>Yes</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>22</td>
<td>R</td>
<td>Portuguese</td>
<td>Current</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>19</td>
<td>R</td>
<td>-</td>
<td>Current</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>25</td>
<td>R</td>
<td>Spanish</td>
<td>No</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>11</td>
<td>F</td>
<td>23</td>
<td>R</td>
<td>Spanish</td>
<td>Yes</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>26</td>
<td>R</td>
<td>Spanish</td>
<td>Yes</td>
<td>25 dB HL</td>
<td>/r/</td>
</tr>
<tr>
<td>13</td>
<td>F</td>
<td>23</td>
<td>R</td>
<td>ASL</td>
<td>Yes</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>14</td>
<td>F</td>
<td>26</td>
<td>R</td>
<td>Spanish, ASL</td>
<td>Yes</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>15</td>
<td>F</td>
<td>20</td>
<td>R</td>
<td>Spanish</td>
<td>No</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>16</td>
<td>M</td>
<td>20</td>
<td>R</td>
<td>Spanish</td>
<td>No</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>17</td>
<td>F</td>
<td>20</td>
<td>R</td>
<td>Spanish</td>
<td>No</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>18</td>
<td>F</td>
<td>21</td>
<td>R</td>
<td>Spanish</td>
<td>No</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>19</td>
<td>F</td>
<td>27</td>
<td>R</td>
<td>Chinese, Bulgarian</td>
<td>Yes</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>20</td>
<td>M</td>
<td>22</td>
<td>R</td>
<td>Spanish</td>
<td>No</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>21</td>
<td>F</td>
<td>21</td>
<td>R</td>
<td>Spanish</td>
<td>No</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>22</td>
<td>F</td>
<td>25</td>
<td>R</td>
<td>-</td>
<td>Yes</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-------</td>
<td>----</td>
<td>------</td>
<td>----</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>F</td>
<td>23</td>
<td>R</td>
<td>Japanese</td>
<td>Yes</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>24</td>
<td>F</td>
<td>24</td>
<td>R</td>
<td></td>
<td>Yes</td>
<td>30 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>25</td>
<td>F</td>
<td>19</td>
<td>R</td>
<td></td>
<td>Current</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>26</td>
<td>F</td>
<td>21</td>
<td>R</td>
<td>ASL</td>
<td>Current</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>27</td>
<td>M</td>
<td>29</td>
<td>R</td>
<td></td>
<td>No</td>
<td>30 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>28</td>
<td>F</td>
<td>25</td>
<td>R</td>
<td>Spanish</td>
<td>No</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>29</td>
<td>F</td>
<td>23</td>
<td>R</td>
<td>Spanish</td>
<td>Current</td>
<td>30 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>30</td>
<td>F</td>
<td>21</td>
<td>R</td>
<td>Spanish</td>
<td>Current</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>31</td>
<td>F</td>
<td>20</td>
<td>R</td>
<td></td>
<td>Current</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>32</td>
<td>F</td>
<td>24</td>
<td>R</td>
<td>Portuguese, Spanish</td>
<td>Yes</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>33</td>
<td>F</td>
<td>27</td>
<td>R</td>
<td></td>
<td>Phonics</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
<tr>
<td>34</td>
<td>M</td>
<td>26</td>
<td>R</td>
<td></td>
<td>No</td>
<td>25 dB HL</td>
<td>/s/, /z/</td>
</tr>
<tr>
<td>35</td>
<td>F</td>
<td>26</td>
<td>R</td>
<td></td>
<td>Yes</td>
<td>25 dB HL</td>
<td>/r/, /s/</td>
</tr>
<tr>
<td>36</td>
<td>F</td>
<td>21</td>
<td>R</td>
<td>Arabic, Spanish</td>
<td>No</td>
<td>25 dB HL</td>
<td>NR</td>
</tr>
</tbody>
</table>

Note. *Hearing Screening Threshold (dB HL) bHistory of remediated speech Disorder: NR = not reported. removed, and the speakers spent the remainder of the session practicing the same stimuli without visual biofeedback support, for the purpose of helping speakers generalize their practiced speech sounds. Baseline recordings (obtained one week prior to treatment), posttreatment recordings (obtained one week after termination of treatment), and) follow-up recordings (obtained six weeks after the posttreatment assessment) were gathered for each participant’s production of the /r/-/l/ contrast in the context of nonsense syllables, words in sentences, and spontaneous speech. It is important to note that the testing stimuli was not used as instructional stimuli, and instructional stimuli was not used as testing stimuli. Lesson plans for treatment sessions can be found in the theses of McDougal (2019) and Price (2019).
Recordings

In preparation for auditory perceptual analysis, each of the targets from selected task types were cut from raw recordings using Adobe Audition sound editing software. Then, to reduce the likelihood of listener fatigue, a random number generator was used to select one of the three target iterations for further perceptual analysis. Each of the randomly selected targets were then evaluated for recording quality; targets which included “pops” or “clicks” within the target, or over-aggressive cutting of the beginning or end of target were replaced with a higher-quality recording, if available. Any pops or clicks that occurred outside the target stimuli were individually removed. The duration of the target stimuli in both nonsense syllables and complete words was controlled and edited to include 500 milliseconds of silence preceding the onset of each sound, and additional silence was added after the sound to create a total duration of 1.25 seconds (+/- 1 millisecond) per target. The stimulus targets were normalized for intensity, and any remaining electronic noise or noise artifacts were filtered out.

Stimuli

The examined phonemes included the liquids /l/ and /r/. This study evaluated target phonemes within the task types of nonsense syllables (lee, eel, ree, eer) and words in sentences (leap, feel, reap, fear). For both of these task types, the Japanese learners of English produced each of the four nonsense syllables or four words three times in random order. Each of the stimulus targets also varied as a function of word position; target stimuli were placed in both initial and final position of consonant-vowel (CV) or vowel-consonant (VC) syllables and consonant-vowel-consonant (CVC) words, and were preceded or followed by the high front vowel /i/. Recordings of the stimuli were produced using a Yeti USB multi-pattern microphone with a sampling rate of 48 kHz and a quantization of 24 bits.
To compare the native Japanese speakers’ performance across time, “comparison” sound files were created. The comparison sound files were built to begin with 0.3 milliseconds of silence, followed by a production of the target phoneme from the baseline condition, followed by one second of silence, and then a comparative phoneme from either the posttreatment or follow-up assessment period. These paired comparative sound files were created to directly compare performance across specific phonemes, linguistic task types, word positions, and assessment periods. It should be noted that although only ratings from the baseline vs. posttreatment and baseline vs. follow-up were included for statistical analysis in this study, the listeners were also tasked with rating comparative files which were sequenced in the opposite order (i.e., posttreatment vs. baseline and follow-up vs. baseline).

**Procedures**

The listeners were asked to evaluate speech sound sample recordings of the Japanese learners of English, and to rate their perception of the quality of the target sound phonemes to standard productions of American English. The target phonemes /l/ and /r/ were embedded within the task types of nonsense syllables and words which were extracted from sentences. Listeners were instructed to disregard the perceived quality of the stimulus items as a whole (i.e., such as pronunciation quality of vowels or non-target consonants), and rather focus their ratings on the evaluation of the perceived quality and goodness of the target phonemes /l/ and /r/ only. The listeners were instructed that they would hear two words: the first word would be known as the anchor, and the perceived goodness of the target phoneme embedded in the second presentation should be compared to the perceived goodness of the target phoneme of the anchor. As shown in Figure 1, the ratings were measured using a visual analog scale presented via a custom computer program. Categorically, the participants were instructed to rate the second
sound in comparison to the anchor along a continuum ranging from “much worse” to “much better.” After using the sliding bar in the middle of the figure to select a judgement rating for each paired comparison, the listeners were instructed to submit their rating and proceed to the next listening pair of stimuli using a click of a computer mouse. Although the listeners submitted categorical and not numerical ratings, the ratings submitted were recorded quantitatively by the computer software program, corresponding numerically to the values -100 to 100.

Figure 1. Visual Analogue Scale used by listeners to submit auditory perceptual ratings.

The stimuli were presented to the listeners via Sennheiser 650 HD open back headphones, which received signal routed from a USB thumb drive which was inserted into a computer. The participants listened to the stimuli while seated in a double-walled sound booth. Although the starting intensity level for the stimuli was originally set to approximately 60 dB HL, each participant was allowed to self-select their preferred volume, within the limits of safe hearing levels. The listeners completed the auditory perceptual ratings within the time frame of
an approximately 40-minute session. In addition to the hearing screening and a practice session, the session included four testing periods: an approximately six-minute session evaluating the target phoneme /l/ within nonsense syllables, another six-minute session evaluating the target phoneme /r/ within nonsense syllables, another six-minute session evaluating the target phoneme /l/ within words extracted from sentences, and another six-minute session evaluating the target phoneme /r/ within words extracted from sentences. Each rating session included eight comparative recordings per speaker. To calculate levels of intra-rater reliability, each session also included seven additional repeated comparative recordings. Using a random number generator, the order of session presentation was randomized for each participant. The participants were offered a two-minute break during the testing process. These listener ratings, which categorically quantified the improvements in goodness of /l/ and /r/ phoneme production from baseline to posttreatment and from baseline to follow-up, were used to determine the efficacy of EPG assisted pronunciation training.

Statistics

Descriptive statistics of mean and standard deviations were reported for the dependent variables. Differences in the speakers’ ability to produce the /l/ and /r/ sounds in a native-like manner one week after the end of instruction (posttreatment) and seven weeks after instruction (follow-up) compared to before the EPG-assisted treatment began (baseline) were reported in terms of a percentage of improvement based on the mean ratings of the 36 listeners. The mean percentage of improvement ratings were categorically classified into the following divisions: much better (MB) indicating a >50% to 100% change, better (B) indicating a >1% to 50% change, same (S) indicating a >-1% to 1% change, worse (W) indicating a -1% to -50% change, and much worse (MW) indicating a <-50% to -100% change.
Differences across word position (initial or final), linguistic context (nonsense syllable or words in sentences), and assessment period (posttreatment or at a follow-up) were determined using a three-way repeated measures analysis of variance (ANOVA). Twenty percent of the tokens were presented to the listeners twice in order to measure intra-rater reliability. Comparisons of the two sets of data resulted in a Pearson Correlation of .76 ($p < .0001$), with an absolute mean difference between the two sets of measures of 1.08 on the -100 to +100 scale.

Results

This study was one part of a three-part project designed to evaluate the efficacy of EPG Assisted Pronunciation Training in helping Japanese L2 learners of English to improve their production of the liquid phonemes /l/ and /r/. The results of the auditory perceptual analysis will be presented first as four individual case studies examining the listeners’ perceptions of the speakers MK, NH, TA, and YM, with initial discussion focused on the L2 learners’ possible improvement producing the /l/ phoneme followed by the /r/ phoneme. Finally, aggregate data across all participants for the mean scores of phonemes /l/ and /r/ will be examined.

MK

Phoneme /l/. As illustrated in Figure 2, posttreatment and follow-up productions of /l/ were rated higher (i.e., better pronunciation) than baseline productions. In terms of the qualitative categorization, slight improvement was noted across all word positions, linguistic contexts, and assessment periods. Greater improvement was seen in the productions of /l/ for nonsense syllables (28.71%) when compared to sentences (19.17%). See Table 2 for listener ratings.
Figure 2. Mean listener ratings comparing MK’s productions of /l/ across linguistic context, word position, and assessment period.

Table 2

Perceptual Rating Means and Standard Deviations as a Function of Task Type and Word Position for Speaker MK’s /l/ Productions

<table>
<thead>
<tr>
<th>Word Position</th>
<th>Linguistic Context</th>
<th>Assessment Period</th>
<th>Mean</th>
<th>SD</th>
<th>Categorical Ratinga</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>35.34</td>
<td>27.91</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>8.02</td>
<td>33.91</td>
<td></td>
<td>B</td>
</tr>
<tr>
<td>Final</td>
<td>Sentences</td>
<td>Posttreatment</td>
<td>14.96</td>
<td>20.71</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>2.04</td>
<td>24.45</td>
<td></td>
<td>B</td>
</tr>
<tr>
<td>Final</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>34.77</td>
<td>29.81</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>36.74</td>
<td>24.14</td>
<td></td>
<td>B</td>
</tr>
<tr>
<td>Final</td>
<td>Sentences</td>
<td>Posttreatment</td>
<td>29.17</td>
<td>32.97</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>30.53</td>
<td>25.79</td>
<td></td>
<td>B</td>
</tr>
</tbody>
</table>

Note. *Categorical Rating: MB indicates > 50% to 100% change, B indicates > 1% to 50% change, S indicates > -1% to 1% change, W indicates -1% to -50% change, MW indicates < -50% to -100% change.
The ANOVA indicated a significant main effect of linguistic context on overall level of phoneme production improvement, $F(1,35) = 9.816, p < .001$, partial eta squared ($\eta^2_p$) = 0.219. Productions of /l/ embedded in nonsense syllables were rated as having 8.5% greater improvement than follow-up productions of /l/ (19.33%). However, when the phoneme was located in the final word position, improvement was slightly higher at the follow-up assessment period (33.64%) than at the posttreatment period (31.97%).

**Phoneme /r/**. Both posttreatment and follow-up productions of /l/ were rated higher than baseline productions, as can be seen in Figure 3. Categorically, all positions, contexts, and assessment periods were judged to be better than baseline. The posttreatment assessment period of /r/ in the final position of words in sentences was rated as “much better” than baseline by listeners. Greater improvement was seen in the productions of /r/ for words in sentences (35.4%) in comparison to the productions of /r/ within the nonsense syllables (21.98%).

The ANOVA indicated a significant three-way interaction between linguistic context, word position, and assessment period on overall level of phoneme production improvement, $F(1,35) = 25.09, p < .001$, $\eta^2_p = 0.418$. Listener ratings are reported in Table 3. In the context of nonsense syllables, phonemes produced in the initial position were rated higher at the post-treatment assessment than at the follow-up assessment, while phonemes produced in the final position were rated higher at the follow-up assessment than at the posttreatment assessment. However, phonemes produced in words showed greater improvement posttreatment ($M = 54.0\%$) than at follow-up ($M = 19.9\%$) across all positions.

**NH**

**Phoneme /l/**. As delineated in Figure 4, posttreatment and follow-up productions of /l/ were rated higher than baseline productions.
Categorically, all experimental conditions were rated as being “better” than baseline, with the posttreatment assessment period of /l/ in the final position of nonsense syllables and in both assessment periods in the final position of words in sentences rated “much better” than baseline.

Figure 3. Mean listener ratings comparing MK’s productions of /r/ across linguistic context, word position, and assessment period.

Table 3

<table>
<thead>
<tr>
<th>Word Position</th>
<th>Linguistic Context</th>
<th>Assessment Period</th>
<th>Mean</th>
<th>SD</th>
<th>Categorical Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>18.52</td>
<td>19.88</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>3.93</td>
<td>21.32</td>
<td>B</td>
</tr>
<tr>
<td>Sentences</td>
<td>Posttreatment</td>
<td>39.54</td>
<td>27.23</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>19.67</td>
<td>36.37</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>Final</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>29.98</td>
<td>31.45</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>35.47</td>
<td>22.72</td>
<td>B</td>
</tr>
<tr>
<td>Sentences</td>
<td>Posttreatment</td>
<td>62.64</td>
<td>21.95</td>
<td>MB</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>20.04</td>
<td>22.09</td>
<td>B</td>
<td></td>
</tr>
</tbody>
</table>

Note. *Categorical Rating: MB indicates > 50% to 100% change, B indicates > 1% to 50% change, S indicates > -1% to 1% change, W indicates -1% to -50% change, MW indicates < -50% to -100% change.
Greater improvement was seen in the productions of /l/ in the final position (56.73%) in comparison to the productions of /l/ in the initial position (18.32%). Listener ratings are reported in Table 4.

A significant interaction between word position and treatment condition on overall level of improvement of phoneme production was indicated by the three-way repeated measures ANOVA, $F(1,35) = 13.57, p = .001, \eta^2_p = 0.279$. Furthermore, productions of /l/ in the final position showed continuing improvement from the immediate posttreatment to follow-up treatment conditions, while productions of /l/ in the initial position showed more improvement in the immediate posttreatment than in the follow-up treatment condition.

**Phoneme /r/**. As visualized in Figure 5, posttreatment or follow-up productions of /r/ were rated higher than baseline productions in five of the eight conditions. The categorical ratings, seen in Table 5, note five conditions rated “better” than baseline, while three conditions were noted as “worse” than baseline. Greater improvement was noted in the productions of /r/ at the initial position (14.12%) in comparison to the productions of /l/ in the final position (-4.24%).

Statistical analysis indicated a significant three-way interaction between context, word position, and treatment condition, $F(1) = 14.53, p = .001, \eta^2_p = 0.293$. While /r/ phonemes produced in the context of nonsense syllables across both word positions in the posttreatment condition showed improvement in comparison to baseline, the follow-up conditions showed less improvement than posttreatment. /r/ phonemes produced in the context final position of nonsense syllables were actually seen to worsen during the posttreatment condition. The /r/ phonemes produced in the context of words for both word positions in the posttreatment condition exhibited a mean decrease ($M = -7\%$) in performance, yet during the follow-up treatment condition showed an improvement ($M = 25.2\%$) in performance in comparison to baseline.
Figure 4. Mean listener ratings comparing NH’s productions of /l/ across linguistic context, word position, and assessment period.

Table 4

Perceptual Rating Means and Standard Deviations as a Function of Task Type and Word Position for Speaker NH’s /l/ Productions

<table>
<thead>
<tr>
<th>Word Position</th>
<th>Linguistic Context</th>
<th>Assessment Period</th>
<th>Mean</th>
<th>SD</th>
<th>Categorical Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>26.64</td>
<td>31.00</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>16.51</td>
<td>28.24</td>
<td>B</td>
</tr>
<tr>
<td>Sentences</td>
<td>Posttreatment</td>
<td>25.27</td>
<td>31.73</td>
<td></td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>4.86</td>
<td>28.80</td>
<td></td>
<td>B</td>
</tr>
<tr>
<td>Final</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>49.41</td>
<td>32.40</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>64.63</td>
<td>24.52</td>
<td>MB</td>
</tr>
<tr>
<td>Sentences</td>
<td>Posttreatment</td>
<td>56.29</td>
<td>30.98</td>
<td></td>
<td>MB</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>56.60</td>
<td>25.74</td>
<td></td>
<td>MB</td>
</tr>
</tbody>
</table>

Note. *Categorical Rating: MB indicates > 50% to 100% change, B indicates > 1% to 50% change, S indicates > -1% to 1% change, W indicates -1% to -50% change, MW indicates < -50% to -100% change.
Figure 5. Mean listener ratings comparing NH’s productions of /r/ across linguistic context, word position, and assessment period.

Table 5

Perceptual Rating Means and Standard Deviations as a Function of Task Type and Word Position for Speaker NH’s /r/ Productions

<table>
<thead>
<tr>
<th>Word Position</th>
<th>Linguistic Context</th>
<th>Assessment Period</th>
<th>Mean</th>
<th>SD</th>
<th>Categorical Rating*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>26.64</td>
<td>29.69</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td></td>
<td>16.39</td>
<td>26.23</td>
<td>B</td>
</tr>
<tr>
<td>Sentences</td>
<td>Posttreatment</td>
<td></td>
<td>-8.86</td>
<td>21.50</td>
<td>W</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td></td>
<td>22.29</td>
<td>20.86</td>
<td>B</td>
</tr>
<tr>
<td>Final</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>5.19</td>
<td>27.63</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td></td>
<td>-44.48</td>
<td>24.43</td>
<td>W</td>
</tr>
<tr>
<td>Sentences</td>
<td>Posttreatment</td>
<td></td>
<td>-5.82</td>
<td>24.45</td>
<td>W</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td></td>
<td>28.15</td>
<td>27.08</td>
<td>B</td>
</tr>
</tbody>
</table>

Note. *Categorical Rating: MB indicates > 50% to 100% change, B indicates > 1% to 50% change, S indicates > -1% to 1% change, W indicates -1% to -50% change, MW indicates < -50% to -100% change.
Phoneme /l/. The posttreatment and follow-up productions of /l/ across all conditions were rated higher than baseline productions, as seen in Figure 6. As reported in Table 6, the initial position of nonsense syllables across both assessment periods were rated as being “much better” than baseline productions. Greater improvement was seen in the productions of /l/ in initial position (59.91%) in comparison to the productions of /l/ in final position (18.18%).

The ANOVA indicated a significant three-way interaction between context, position, and treatment condition on overall level of improvement of phoneme production, $F(1,35) = 21.59, p <.001, \eta^2_p = 0.382$. In the context of nonsense words, level of improvement increased from posttreatment to follow-up across all contexts, with the greatest levels of improvement seen in the initial position of words. In the context of words in sentences, the initial position showed an increase in performance from posttreatment to follow-up, while the final position showed higher performance at posttreatment than at follow-up.

Phoneme /r/. Posttreatment and follow-up productions of /r/ were rated to have better pronunciation than baseline productions. See Figure 7 for a visualization of this data. In terms of the categorical ratings, seven conditions were rated as “better” than baseline, while the follow-up assessment period of nonsense syllables in the final position was rated as the “same” as baseline. Listeners’ perceptual ratings and categorical ratings are displayed in Table 7.

The ANOVA indicated a significant main effect of context on goodness of phoneme production, $F(1,35) = 20.56, p < .001, \eta^2_p = 0.370$. TA demonstrated more improvement when the /r/ speech sound was produced in the linguistic context of words within sentences ($M = 26\%$) than in the context of nonsense syllables ($M = 7.9\%$). Additionally, the ANOVA also indicated a significant two-way interaction in listener ratings of pronunciation improvement between the
Figure 6. Mean listener ratings comparing TA’s productions of /l/ across linguistic context, word position, and assessment period.

Table 6

Perceptual Rating Means and Standard Deviations as a Function of Task Type and Word Position for Speaker TA’s /l/ Productions

<table>
<thead>
<tr>
<th>Word Position</th>
<th>Linguistic Context</th>
<th>Assessment Period</th>
<th>Mean</th>
<th>SD</th>
<th>Categorical Rating*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>85.02</td>
<td>17.67</td>
<td>MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>89.18</td>
<td>13.82</td>
<td>MB</td>
</tr>
<tr>
<td></td>
<td>Sentences</td>
<td>Posttreatment</td>
<td>25.41</td>
<td>28.66</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>41.19</td>
<td>29.35</td>
<td>B</td>
</tr>
<tr>
<td>Final</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>6.61</td>
<td>22.58</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>18.21</td>
<td>46.39</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Sentences</td>
<td>Posttreatment</td>
<td>44.32</td>
<td>27.11</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>3.56</td>
<td>42.32</td>
<td>B</td>
</tr>
</tbody>
</table>

Note. *Categorical Rating: MB indicates > 50% to 100% change, B indicates > 1% to 50% change, S indicates > -1% to 1% change, W indicates -1% to -50% change, MW indicates < -50% to -100% change.
word position and assessment period variables, $F(1,35) = 8.76, p < 0.01, \eta^2_p = 0.200$. Greater improvements were seen across all contexts in the final position than in the initial position. Greater improvements in the pronunciation of /r/ were found at the posttreatment assessment ($M = 32\%$) than at follow-up ($M = 9.3\%$).

**YM**

**Phoneme /l/.** YM’s posttreatment and follow-up productions of /l/ were rated higher than baseline productions for the majority of experimental conditions, as illustrated in Figure 8. The data in Table 8 indicate that pronunciation of /l/ across six conditions was rated categorically as “much better” than baseline, with one condition being rated as “better,” while the phoneme /l/ in the initial position of words in sentences was rated as “worse.” A higher degree of improvement was seen in the productions of /l/ at the level of nonsense syllables ($M = 67.78\%$) in comparison to the slight improvement seen in the productions of /l/ at the level of words in sentences ($M = 30.21\%$).

As indicated by the ANOVA, there was a significant three-way interaction found between context, word position, and assessment period on the pronunciation improvement of /l/ production, $F(1,35) = 54.99, p < .0001, \eta^2_p = 0.610$. As noted previously, greater improvement was noted at the level of nonsense syllables in comparison to words in sentences. In the context of nonsense syllables, phonemes in the initial position showed greater improvement in the posttreatment than in the follow-up condition, while the context of words showed greater improvement in the follow-up condition than in the posttreatment condition.

**Phoneme /r/.** Posttreatment and follow-up pronunciation of /r/ were rated improved compared to baseline productions in seven of the eight conditions, as can be seen in Figure 9.
Figure 7. Mean listener ratings comparing TA’s productions of /r/ across linguistic context, word position, and assessment period.

Table 7

Perceptual Rating Means and Standard Deviations as a Function of Task Type and Word Position for Speaker TA’s /r/ Productions

<table>
<thead>
<tr>
<th>Word Position</th>
<th>Linguistic Context</th>
<th>Assessment Period</th>
<th>Mean</th>
<th>SD</th>
<th>Categorical Ratinga</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>7.28</td>
<td>29.98</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td></td>
<td>1.61</td>
<td>37.67</td>
<td>B</td>
</tr>
<tr>
<td>Sentences</td>
<td>Posttreatment</td>
<td></td>
<td>25.34</td>
<td>31.59</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td></td>
<td>22.35</td>
<td>38.86</td>
<td>B</td>
</tr>
<tr>
<td>Final</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>22.40</td>
<td>28.00</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td></td>
<td>0.25</td>
<td>37.45</td>
<td>S</td>
</tr>
<tr>
<td>Sentences</td>
<td>Posttreatment</td>
<td></td>
<td>45.66</td>
<td>19.44</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td></td>
<td>13.10</td>
<td>22.92</td>
<td>B</td>
</tr>
</tbody>
</table>

Note. *Categorical Rating: MB indicates > 50% to 100% change, B indicates > 1% to 50% change, S indicates > -1% to 1% change, W indicates -1% to -50% change, MW indicates < -50% to -100% change.
Figure 8. Mean listener ratings comparing YM’s productions of /l/ across linguistic context, word position, and assessment period.

Table 8

Perceptual Rating Means and Standard Deviations as a Function of Task Type and Word Position for Speaker YM’s /l/ Productions

<table>
<thead>
<tr>
<th>Word Position</th>
<th>Linguistic Context</th>
<th>Assessment Period</th>
<th>Mean</th>
<th>SD</th>
<th>Categorical Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>81.73</td>
<td>19.22</td>
<td>MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>61.43</td>
<td>34.62</td>
<td>MB</td>
</tr>
<tr>
<td>Sentences</td>
<td>Posttreatment</td>
<td>-24.10</td>
<td>49.25</td>
<td></td>
<td>W</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>39.33</td>
<td>30.00</td>
<td></td>
<td>B</td>
</tr>
<tr>
<td>Final</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>56.76</td>
<td>35.34</td>
<td>MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>71.19</td>
<td>29.85</td>
<td>MB</td>
</tr>
<tr>
<td>Sentences</td>
<td>Posttreatment</td>
<td>51.19</td>
<td>26.46</td>
<td></td>
<td>MB</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>54.44</td>
<td>29.36</td>
<td></td>
<td>MB</td>
</tr>
</tbody>
</table>

Note. *Categorical Rating: MB indicates > 50% to 100% change, B indicates > 1% to 50% change, S indicates > -1% to 1% change, W indicates -1% to -50% change, MW indicates < -50% to -100% change.
In terms of the categorical ratings, major improvement was noted across three conditions, slight improvement was noted across four conditions, and a slight worsening was noted across one condition, as indicated in Table 9.

The ANOVA indicated three different significant two-way interactions. A two-way interaction was found for context and position, $F(1,35) = 73.86, p < .001, \eta^2_p = 0.678$. In the context of nonsense syllables, productions in initial position showed more improvement than productions in the final position. In the context of words in sentences, more overall improvement was seen in the final position than in the initial position.

Another two-way interaction was indicated between context and treatment condition, $F(1,35) = 79.146, p < .001, \eta^2_p = 0.693$. In the context of nonsense words, the posttreatment condition in the initial position showed greater improvement than in the final position, while the posttreatment condition in the final position showed a decrease in level of performance when compared to the follow-up condition. In the context of words in sentences, the posttreatment data showed greater improvement than follow-up treatment data across both positions. The final two-way interaction was indicated between position and treatment condition, $F(1,35) = 55.58, p < .001, \eta^2_p = 0.614$. In the initial position, greater improvement was noted in the posttreatment than in the follow-up condition, while in the final position, greater overall performance was seen in the follow-up condition than in the posttreatment condition.

**Overall**

**Phoneme /l/**. The aggregate listener ratings across all Japanese speakers are displayed in Table 10. As illustrated in Figure 10, posttreatment and follow-up pronunciations of /l/ were rated higher than baseline productions, with an average of 37.42% improvement across all speakers and all assessment periods. The context of nonsense syllables (46.39%) showed greater
improvement than in the context of words in sentences (28.44%). Greater improvement was noted in the final word position (41.53%) than in the initial word position (33.31%).

**Phoneme /r/.** The mean listener ratings across all Japanese speakers are seen in Table 11. As illustrated in Figure 11, posttreatment and follow-up productions of /r/ were rated higher than baseline productions, with an aggregate 21.47% improvement noted. The participants showed greater improvement at the posttreatment assessment period (25.89%) than at the follow-up assessment period (17.05%). Categorically, all productions in all conditions across posttreatment and follow-up assessment periods were rated as “better” than baseline. Greater improvement was observed in the context of words in sentences (27.74%) than in the context of nonsense syllables (15.21%). Similar levels of improvement were noted across word position, with 21.74% improvement seen in initial position and 21.20% improvement in the final position.

![Figure 9](image-url)

*Figure 9.* Mean listener ratings comparing YM’s productions of /r/ across linguistic context, word position, and assessment period.
Table 9

Perceptual Rating Means and Standard Deviations as a Function of Task Type and Word Position for Speaker YM’s /r/ Productions

<table>
<thead>
<tr>
<th>Word Position</th>
<th>Linguistic Context</th>
<th>Assessment Period</th>
<th>Mean</th>
<th>SD</th>
<th>Categorical Ratinga</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>50.85</td>
<td>28.90</td>
<td>MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>46.21</td>
<td>27.93</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Sentences</td>
<td>Posttreatment</td>
<td>50.95</td>
<td>28.38</td>
<td>MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>5.16</td>
<td>19.75</td>
<td>B</td>
</tr>
<tr>
<td>Final</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>-11.99</td>
<td>40.35</td>
<td>W</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>35.03</td>
<td>32.24</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Sentences</td>
<td>Posttreatment</td>
<td>56.02</td>
<td>27.37</td>
<td>MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>47.59</td>
<td>31.21</td>
<td>B</td>
</tr>
</tbody>
</table>

Note. *Categorical Rating: MB indicates > 50% to 100% change, B indicates > 1% to 50% change, S indicates > -1% to 1% change, W indicates -1% to -50% change, MW indicates < -50% to -100% change.

Figure 10. Mean listener ratings comparing aggregate productions of /l/ across linguistic context, word position, and assessment period.
Table 10

*Auditory Perceptual Rating Means as a Function of Task Type and Word Position for All Speakers’ /l/ Productions*

<table>
<thead>
<tr>
<th>Word Position</th>
<th>Linguistic Context</th>
<th>Assessment Period</th>
<th>Mean</th>
<th>SD</th>
<th>Categorical Rating&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>57.18</td>
<td>23.95</td>
<td>MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>43.79</td>
<td>27.65</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Sentences</td>
<td>Posttreatment</td>
<td>10.39</td>
<td>32.59</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>21.86</td>
<td>28.15</td>
<td>B</td>
</tr>
<tr>
<td>Final</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>36.89</td>
<td>30.03</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>47.69</td>
<td>31.23</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Sentences</td>
<td>Posttreatment</td>
<td>45.24</td>
<td>29.38</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>36.28</td>
<td>30.80</td>
<td>B</td>
</tr>
</tbody>
</table>

<sup>a</sup>Categorical Rating: MB indicates > 50% to 100% change, B indicates > 1% to 50% change, S indicates > -1% to 1% change, W indicates -1% to -50% change, MW indicates < -50% to -100% change.

Figure 11. Mean listener ratings comparing aggregate productions of /r/ across linguistic context, word position, and assessment period.
Table 11

Perceptual Rating Means as a Function of Task Type and Word Position for All Speakers’ /r/ Productions

<table>
<thead>
<tr>
<th>Word Position</th>
<th>Linguistic Context</th>
<th>Assessment Period</th>
<th>Mean</th>
<th>SD</th>
<th>Categorical Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>25.82</td>
<td>27.11</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>17.04</td>
<td>28.29</td>
<td>B</td>
</tr>
<tr>
<td>Final</td>
<td>Nonsense</td>
<td>Posttreatment</td>
<td>11.40</td>
<td>31.86</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>6.57</td>
<td>29.21</td>
<td>B</td>
</tr>
<tr>
<td>Sentences</td>
<td></td>
<td>Posttreatment</td>
<td>26.74</td>
<td>27.18</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>17.37</td>
<td>28.96</td>
<td>B</td>
</tr>
<tr>
<td>Sentences</td>
<td></td>
<td>Posttreatment</td>
<td>39.62</td>
<td>23.30</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
<td>27.22</td>
<td>25.83</td>
<td>B</td>
</tr>
</tbody>
</table>

Note. *Categorical Rating: MB indicates > 50% to 100% change, B indicates > 1% to 50% change, S indicates > -1% to 1% change, W indicates -1% to -50% change, MW indicates < -50% to -100% change.

Discussion

The initial aim of this study was to examine if EPG assisted instruction can be used to assist native Japanese L2 English learners acquire the /r/-/l/ phonemic contrast in a more native-like manner. Auditory perceptual ratings from a group of native English speaking adult listeners found that after the EPG assisted pronunciation training sessions, overall improvement was noted in productions of both /r/ and /l/, with more improvement noted in productions of /l/ than in productions of /r/. The greater improvement noticed in the English phoneme /l/ may be partially explained by the fact that it shares a similar articulation contact pattern on the alveolar ridge with the Japanese apicoalveolar tap. This study corroborates the results of previous studies performed by Bright (1999), Schmidt and Beamer (1998), Gibbon et al. (1991), McDougal (2019), and Price (2019).
After completing an EPG assisted pronunciation training program focused on improving phonemic consonant contrasts, each of the three participants in the Bright (1999) study also demonstrated progress in pronunciation of target phonemes. Even though Bright’s EPG assisted pronunciation training program did not target the /r/-/l/ contrast, the targeted consonant contrasts showed demonstrable improvement across all participants. Like the current study, the methodology for assessing pronunciation changes in Bright’s study included elements of perceptual analysis. Both the current study and Bright’s study provide support for the idea that EPG assisted pronunciation training can make a functional difference in improving the articulatory accuracy of phoneme production as perceived by listeners.

Schmidt and Beamer (1998) also showed results indicating that the participants demonstrated changes in articulation in target phonemic contrasts (including the /r/-/l/ contrast) after EPG assisted pronunciation training. Although the improvement in Schmidt and Beamer’s study was mainly assessed through contact pattern analysis, it also included a perceptual component as did the current study. Like the Schmidt and Beamer study in which each of the participants showed changes in articulation applying to both previously-learned and new sounds, the participants in the current study showed changes in pronunciation accuracy which helped them increase the differentiation between a phonemic contrast not found in the speakers’ first language. Additionally, Schmidt and Beamer noted that even with additional visual biofeedback provided from the EPG, the speakers learned visual consonants (such as /θ/) more quickly than less visual consonants (such as /r/); the results of this study support the same conclusion.

Gibbon et al. (1991) showed that after treatment, participants were able to make consistent contrasts between the liquids /r/ and /l/, as did the current study. As both studies followed a similar training protocol including a focus on awareness, articulatory patterns,
establishing the sound in different vowel contexts and also in more complex phonetic contexts with support of EPG, it gives additional evidence to the idea that this protocol for EPG assisted pronunciation training can help Japanese speakers of English improve their productions of the /l/-/r/ contrast.

Examining the efficacy of EPG assisted pronunciation training through the lens of lingual-palatal contact pattern analysis, McDougal (2019) examined data from the same participants as the current study. Her results showed improvement in contact pattern accuracy across participants. However, in some instances, such as MK, the contact pattern analysis showed that the participant made the greatest changes in contact patterns for the target phoneme /l/, and fewer changes in the placement of the target phoneme /r/. In the current study, however, listeners perceived that MK made more improvement in the pronunciation of the /r/ phoneme than in the pronunciation of the /l/ phoneme. There could be several factors influencing this discrepancy; McDougal’s work examined and averaged results from all phonemic tokens, while the current study examined only eight randomly selected targets and thus the results may have been skewed by the smaller sample size. Additionally, it is possible that MK may have made improvements in /r/ that although heard perceptually, may not have been fully captured by the tongue contact pattern analysis. In addition to tongue contact pattern accuracy across participants, McDougal’s study also noticed an increase in differentiation between productions of /r/ and /l/.

Evaluating the efficacy of EPG assisted pronunciation training as measured by F3 formant analysis, Price (2019), also evaluated data from the same participants as the current study and as McDougal (2019). Results of Price’s study indicated that all subjects demonstrated an increase in the F3 contrast values at the posttreatment assessment period, and that this increase in F3 contrast was maintained at the follow-up assessment for the subjects NH and YM,
but that MK and TA did not maintain the increase in F3 contrast at the period of the follow-up assessment. These results support the observations noted in this study that as a whole, NH and YM showed continually increasing levels of performance from posttreatment to follow-up assessment periods, while MK and TA did not demonstrate full maintenance at follow-up of the gains obtained by the assessment period of posttreatment. This may indicate that F3 formants and auditory perceptual analysis may be correlated.

Another aim of the current study was to evaluate if the possible benefit of EPG assisted instruction varied as a function of linguistic task type. Overall results suggest that the benefit of EPG assisted pronunciation training may vary as a function of target phoneme and linguistic task type. Overall, the phoneme /l/ showed the most improvement across participants in the linguistic task type of nonsense syllables, while the phoneme /r/ showed the most improvement in the context of words in sentences. As research literature and clinical evidence suggest that simpler linguistic contexts (such as nonsense syllables) are typically acquired more quickly and with a higher degree of accuracy than more complex linguistic contexts (such as words in sentences), it would be expected that both phonemes would show greater improvement in the context of nonsense syllables, at least initially. In this case, additional information, such as the baseline levels of performance of each of the target phonemes, would be extremely valuable in helping to elucidate whether the greater comparative improvement of /r/ noted in the context of sentences was secondary to a lower baseline performance than seen in the context of nonsense syllables, or whether the performance in the context of words in sentences was actually better than in the context of nonsense syllables.

This study also sought to examine if the improvement in an L2 learner’s production of the American English /r/ and /l/ differed depending on whether the sounds were in a word initial
or word final position. Overall results indicated that the benefit of EPG assisted instruction may vary as a function of both word position and target phoneme. Speakers showed the greatest improvement with the target phoneme /l/ in the initial position, while the greatest improvements in the target phoneme /r/ were seen in the final position. As the Japanese language dictates that the only consonant allowed in the final position of the CVC syllable structure is an /n/, closing a syllable with a consonant such as an /r/ or an /l/ is an unfamiliar concept to beginning Japanese learners of English. It would be expected that the Japanese speakers might have showed the most improvement in final position of words, as it would technically have the lowest baseline and thus the greatest amount of room for improvement.

One limitation of the current study was that no baseline ratings were obtained; only comparative ratings between assessment periods were obtained. Thus, we do not have a way to compare whether the Japanese speakers initially showed greater accuracy in the initial or final positions of words. However, one benefit to the comparative rating system is that it may provide a more ecologically valid way to measure the true improvement across assessment periods. Providing listeners with an “anchor” sound helps reduce the cognitive complexity of the task to deciding if the second sound has improved in comparison to the first sound, and if so, to what degree. Conversely, if the participants were presented with one sound at a time and asked to rate it without a comparison, the cognitive complexity would increase and the likelihood of obtaining valid and reliable results would decrease. Even if one were to include the anchor sound of a native English speaker producing the target phoneme before the presentation of the L2 speaker’s target phoneme production, it would be problematic as any mismatch in fundamental frequency or other vocal qualities between the speakers could disrupt the overall results. Although it was ecologically very valuable to have the baseline token be compared directly and immediately with
the posttreatment or follow-up token, these types of comparisons limited the types of inferential statistics that could be run with this particular variable. In future studies, researchers may consider obtaining baseline ratings in addition to comparative ratings to expand the repertoire of available statistical analyses.

Additionally, as mentioned previously, the study may have been constrained by the selection of a limited number of targets per speaker for audio perceptual analysis, which may not provide a full context of the speakers’ true levels of performance. Furthermore, it is a possibility that there may have been lingering context effects from other productions of other sounds in the word. In future studies, it may be helpful for researchers to include a larger number of targets for audio-perceptual analysis. It may also be helpful to extract target phonemes from their linguistic contexts and compare them in isolation in addition to comparing them in context.

Furthermore, as this study included a high proportion of right-handed raters, there may be a slight positivity effect in these data, as it is less likely for right-handed speakers to drag the mouse as far to the left as left-handed raters. In the future, it may be helpful to implement measures to help control for the positivity effect, such as including more left-handed participants.

Perhaps the greatest limitation in the study is a lack of a control group to help determine whether the improvements noted in the Japanese L2 English learners’ productions of /r/ and /l/ may have been influenced by outside causes, such as their participation in classes at the English Language Center. Future studies may consider a multiple baseline design to control for threats to history and internal validity.

Notwithstanding the need for additional research in this area, this study provided further insight into how an EPG assisted method of instruction might facilitate L2 learning of English for native Japanese speakers. As this demographic of English language learners have historically
been resistant to traditional types of instruction targeting the /r/-/l/ contrast, it is significant to note that EPG assisted pronunciation training helped them make improvement. It is hoped that the auditory perceptual findings collected and analyzed in this study will improve our understanding in how this type of treatment might be used in future pronunciation training.
References


Retrieved from

https://open.library.ubc.ca/cIRcle/collections/ubcthesis/24/items/1.0073242
APPENDIX A

Annotated Bibliography


Objective: This study examined whether native Japanese learners of English would be more successful producing the English /r/ than the English /l/ through a longitudinal study of native Japanese children and adults. Methods: The authors based their hypothesis on the “Speech Learning Model” theory that L2 speech sounds are more learnable if they are more distant from the closest L1 speech sounds; and thus posited that it would be easier for Japanese learners of English to produce the English /r/ than the English /l/. To test this hypothesis, sixteen adults and sixteen children, all native Japanese learners of English, participated in the study. Sixteen adults and sixteen children, all native English speakers, also participated as controls. The participants listened to auditory consonant contrasts of /l/-/r/ in addition to several other consonant contrasts. In sets of three auditory productions, participants were instructed to identify whether or not one of the three consonants was different from the other two. In the second experiment, the production of English /r/, /l/, and /w/ was examined in the same group of participants who took part in the first experiment; to examine phonetic learning, data was collected one year apart, so that two sets of data points could be referenced. Results: The children’s production of /r/ showed more improvement than /l/, and their discrimination of contrasts was better during the second testing period. Conclusions: Based on their previous research and the results of their study, the authors concluded that when one first begins to learn an L2, it will be easiest to learn L2 speech sounds that are more similar to L1 speech sounds, as substituting the L1 speech sound can yield
relatively high intelligibility. However, in the long term, it is hypothesized that L2 speech sounds which are more phonetically dissimilar from the closest L1 sounds will actually result in a more advantageous effect: more accurate perception and segmental production, as it is more likely to result in the formation of a “new phonetic category.” Relevance to current study: The current study will also be focused on the /l/ and /r/ liquid contrast within the native Japanese learners of English population.


**Objective:** To determine whether EPG could help English L2 learners improve their pronunciation. **Methods:** Three native Spanish speaking English L2 learners received EPG pronunciation training, and 10 listeners assessed their quality and goodness of pronunciation of target phonemes before and after pronunciation training. First, each individual was interviewed, then, hearing was tested and the patients were recorded in producing a minimal pair phoneme list. Then, treatment began which included three 90-minute sessions per week. During training sessions, both the clinician and therapy participants used EPG palates, and the clinician provided training on lingual placement of the two phonemes. Additionally, each phoneme was presented to the participants to aid in the development of auditory discrimination. Then, the clinician provided feedback on each participant’s target phoneme productions, and providing models when needed. After establishing a contrastive pronunciation pattern, the L2 learners practiced self-correction with each other. While targeting phonemes, L2 learners practiced minimal pairs in isolation until reaching a 90 percent accuracy rate; then, L2 learners practiced at the word
level, and finally, they practiced at the sentence level. After completing twelve treatment sessions, the researchers readministered the Comp and recorded participants as they read the list of minimal pairs containing the target phonemes. **Results:** Each participant demonstrated progress in accent reduction in target phonemes at both the word level and sentence level, as measured by the Comp. On a least mean square comparison of composite scores from ten listeners who rated “goodness of pronunciation” (of words overall rather than target phonemes independently), overall improvement was shown in standardized scores. Although an ANOVA of aggregate baseline and posttreatment data across participants showed significant changes (p=0.008), two-tailed t-tests for individual participants did not show significance. **Conclusions:**

Based on these results, the authors concluded that EPG is an effective tool for helping L2 learners improve their discrimination and production of target phonemes. **Relevance to current study:** The current study also uses EPG as a tool to help L2 English learners improve their pronunciation of target phonemes.


**Objective:** To determine if children with speech difficulties would benefit from EPG treatment, and also to investigate whether predictive factors could accurately prognose which children would benefit the most from therapy. **Method:** Ten children were selected and given rankings which attempted to prognose degree of improvement of articulation errors based on age, severity of speech difficulty, duration of therapy, additional problems, and awareness and motivation. Each child then engaged in 10 EPG sessions with a therapist, and each of their articulation errors
were targeted. Baseline and post-therapy measures were obtained and compared for each subject. Results: The change in scores of percent consonants correct before and after therapy was found to be significant, and all children showed improvement in the speech sounds which were targeted during treatment. Conclusions: Based on the results, the authors concluded that EPG assisted treatment is beneficial for children with difficult to treat speech sound difficulties. Relevance to current study: This study is important as it shows that EPG can significantly improve pronunciation in individuals with non-organic etiology of speech difficulties. EPG is the same device that will be used in this study in an attempt to help non-native English speakers (with non-organic etiology) improve their pronunciation of consonants.


Objective: The authors sought to determine whether one’s ability to speak English impacts employment status, work status, and earnings; if those relationships hold across a variety of personal characteristics relating to earnings and employment; and if the relationship between employment status and one’s English-speaking abilities, earnings, and work status are different between language groups. Methods: The authors used data obtained from the long form of the 2000 census, as well as data from the internal restricted file of the US Census. Then, to study the relationships between the variables of interest, the authors sorted and analyzed the data by work status, thirty-nine languages, English language ability level, age groups, gender, race and ethnicity, levels of educational attainment, occupation categories, age of entry into the United
States, years spent in the United States, and median earnings. *Results:* Individuals who speak English as an L2 “well” earned, on average, $7,000 less per year than individuals who spoke English as an L2 “very well.” *Conclusions:* Work status, employment, and earnings have a direct variation with one’s English-speaking proficiency. The positive relationship between level of higher English-speaking ability and increased annual wages, higher rates of employment, and increased amount of full-time employment occurred even when individuals had a variety of different personal characteristics. The higher one’s level of English proficiency, the higher one’s income. *Relevance to current study:* The current study specifically researches the impact of a technological advance to help Asian learners of English improve their English-speaking proficiency, and the work of Day and Shin supports the idea that increasing one’s English proficiency will have a direct impact on improving one’s quality of life, specifically through increasing annual income.


*Objective:* This paper evaluates the role and function of biofeedback systems, such as EPG, in the treatment of speech pathologies, and attempts to provide clinical reasoning explaining why EPG is an efficacious tool. It also examines the various motor speech patterns acquired by individuals in a study, including the establishment of new sound patterns, elimination of incorrect patterns, and improved modification of approximated patterns. *Methods:* The authors of the article summarized and synthesized information from a variety of research articles on the topic of interest. Additionally, the authors analyzed a data set of 23 participants obtained by a
study funded by the British Medical Research Council. **Results:** Each of the 18 participants who successfully finished treatment showed perceptually normal productions of the sounds targeted during the EPG treatment therapy immediately following and three months after the conclusion of treatment. **Conclusions:** Additionally, EPG was an efficacious tool to help individuals with velopharyngeal inadequacy, speech sound disorders, repaired cleft palates, etc., and is an efficacious method to help individuals correctly learn lingual configurations required to produce accurate speech sounds and modify temporal or spatial aspects of production patterns. **Relevance to current study:** This study is important as it shows that EPG can be a useful biofeedback tool, which helps individuals with speech sound disorders learn correct patterns through providing immediate visual feedback. EPG is also used in this study in an effort to help Asian learners of English modify their production patterns of /l/ and /r/.


**Objective:** This paper is a descriptive multiple subject case study which uses electropalatography to examine the lingual contact patterns of native Japanese speakers as they attempt to produce the English /r/, /l/ contrast, and monitors their subsequent response to treatment. **Methods:** Participants included two native Japanese learners of English who both had difficulty producing the /r/, /l/ contrast and were approximately nineteen years old, as well as one of the authors of the study, all of whom were living in England. Each of the participants read a list of words containing all of the lingual consonant sounds, including a list of minimal pairs. Additionally, the participants read a list of words in Japanese with examples of /r/ in varying
phonetic contexts. As a control, a native English speaker was recorded reading the same list of words. Then, the participants participated in four treatment sessions with the support of visual biofeedback from the EPG device which focused on awareness, articulatory patterns, establishing the sound in different vowel contexts, and finally, complex phonetic contexts. Then, a final recording was made of the participants as they read the same word list that they read prior to initiation of the study. Results: Before intervention, many of the Japanese learners of English had many similarities between their productions of /r/ and /l/; additionally, they produced a variety of patterns for the Japanese /r/ phoneme, which the authors attributed to the fact that Japanese phonology only has one liquid phoneme, which is phonemically transcribed as /ɾ/.

After treatment, the participants were able to make consistent contrasts between the liquid sounds in controlled contexts. The students who participated in the program seemed to be able to better identify and then change their articulatory movements for the liquids /l/ and /ɾ/.

Conclusions: The authors concluded that although further research is needed, especially in researching the ability of EPG to help improve pronunciation in fluent speech, that further research is warranted into the impact of EPG as a tool for pronunciation training. Relevance to current study: This study also utilizes EPG as a visual biofeedback tool to help native Japanese speakers improve their /l/-/ɾ/ phonemic contrasts.


Objective: This study provides an overview on electropalatographical research on persistent sound system disorders; additionally, it describes the establishment of a system designed to
provide centralized locations for EPG access. A case study of an eight-year-old boy with persistent speech disorder who obtained an EPG through the network was presented.

Methods: The authors evaluated and summarized a plethora of various articles regarding electropalatography, the EPG patterns of typically-developing children, the EPG patterns of children with persistent sound system disorders and the role of EPG in assessment and diagnosis, and using EPG as a form of visual feedback therapy. The authors then performed a case study on an 8 year-old male with an articulatory backing process, measuring his lingua-palatal patterns both before and after treatment using EPG as a form of visual biofeedback. Results: The authors noted that although children with speech sound disorders produce phonemes that sound identical when measured by perceptual auditory methods, EPG technology has revealed that children actually consistently differentiate these sounds through “covert contrasts,” in which they produce lingual contact patterns that subtly differentiate between target categories. The data taken from the case study indicate that prior to treatment, the child produced sound patterns typical of children with persistent sound disorders, such as covert contrasts; after treatment, the child had learned to correctly articulate. Conclusions: The authors concluded that children without speech sound disorders have more consistent articulatory placement as measured by EPG technology than children with persistent speech sound disorders, who often presented with undifferentiated lingual gestures, in which lingual consonants were produced with excessive contact between the palate and the tongue. Overall, the authors conclude that EPG can play an important role in assessing, diagnosing, and also treating children who have speech sound disorders. Relevance to current study: This study will also use electropalatography as a tool to help individuals improve their speech sound production.

**Objective:** The researchers sought to gain greater clarity on a potential link between the accuracy levels in which native Japanese speakers correctly identify and produce the target English phonemes /l/ and /r/, and the extent to which the same speakers assimilate the target phonemes into the native Japanese apicoalveolar tap category. **Methods:** The results from 36 native Japanese speakers were used in the study, with ages ranging between 19 and 48 years old, all of whom had received English instruction for over 5 years and had lived in English-speaking countries between 1 month - 13 years. First, researchers measured the participants’ ability to identify English /r/ and /l/ by using a bilingual three-alternative identification measure. Then, trained listeners rated the audio perceptual quality of the native Japanese speakers’ spoken productions of the target phonemes /r/ and /l/. Finally, the listeners were presented with several synthetic representations of /r/, /l/, and the Japanese apicoalveolar tap which included modified F1, F2, F3, transition duration, and closure duration, and asked to the best exemplars of both L2 and L1 target phonemes. **Results:** The identification accuracy for the English phonemes /l/ and /r/ were correlated significantly with the /r/ production accuracy, but not with the /l/ production accuracy. **Conclusions:** Results suggest that for Japanese adults, the category assimilation for the /l/ - /r/ contrast is slightly asymmetrical; the English liquid phoneme /l/ assimilates more strongly into the category of the apicoalveolar tap than does the English /r/. However, the extent to which the native Japanese speakers assimilate the phonemes /l/ and /r/ into the apicoalveolar tap category does not predict one’s ability to correctly identify the target liquid phonemes. Instead, the ability of native Japanese speakers to correctly identify the target phonemes /l/ and /r/ is
thought to be caused by their phonetic and auditory sensitivities and the way in which they represent the F3 for these target phonemes. **Relevance to current study:** The current study also includes an auditory perceptual analysis of native Japanese speakers’ productions of the liquid phonemes /l/ and /ɾ/.


**Objective:** The researchers sought to analyze the learning style preferences of native Japanese learners of English. **Methods:** Sixty three native Japanese students, all of whom were English majors at a university in Japan, were given a questionnaire to probe their learning style preferences; after this, ten subjects were selected for a more in-depth interview. **Results:** The authors noted that from the gestalt, no single learning style emerged as a preference of all Japanese learners of English. However, researchers noted several of the top preferences of Japanese learners of English included having a preference for learning English by watching and listening native speakers of English, preference for practicing pronunciation of speech sounds, and preference for thorough explanations from the teacher. Their least frequent preferences included studying English alone, learning English with the whole class, talking to friends in English, studying grammar, teacher-directed problems, and reading. **Conclusions:** Based on the results, the author recommends that teachers should help Japanese students adapt to communicative learning styles by initially call on students individually to speak, to carefully structure and scaffold group activities, and to increase student participation by asking simple questions which initially only require one-word answers. **Relevance to current study:** The current
study sought to utilize the preferred learning styles of Japanese learners of English to help facilitate acquisition of the liquid phonemes /ɾ/ and /l/ through specialized intervention.


**Objective:** To research whether visual cues can help Japanese L2 learners of English improve perception of a new phonemic contrast, and to determine the efficacy of audiovisual perceptual training in improving the learner’s pronunciation of the novel contrast. **Methods:** Use of visual cues in the perception of speech for the English phonemic labiodental/bilabial contrasts of /v/, /b/, /p/ and the phonemic contrasts of /l/, /ɾ/ were separately evaluated. The first study of 39 participants evaluated the perceptual distinction of the labiodental/bilabial contrasts between audio, visual, and audiovisual modalities. Then, each participant engaged in 10 auditory or audiovisual training sessions and was subsequently retested. The second study included 62 participants, and measured the perceptual distinction of the /l/, /ɾ/ contrast between audio, visual, and audiovisual modalities. Then, each participant engaged in 10 auditory, natural audiovisual, or synthetic audiovisual training, and then was retested. **Results:** After /l/ and /ɾ/ training, pronunciation improved for all groups, and especially for those in the natural audiovisual training group. **Conclusions:** Audiovisual training was shown to be more efficacious than audio training for the labiodental/bilabial contrasts. The /l/-/ɾ/ perceptual training was efficacious in all trial groups, but the audiovisual training wasn't more efficacious than audio training. **Relevance to current study:** This study is relevant, as it shows how audiovisual biofeedback can be efficacious in improving the pronunciation of the /l/ /ɾ/ contrast in native Japanese speakers who
are learning English as an L2. Visual biofeedback will also be used in this study to attempt to help native Japanese speakers improve their pronunciation of the /l/ and /r/ contrast.


Objective: The researchers sought to learn the largest motivating factors which inspired native Japanese EFL learners to learn English. Additionally, the researchers sought to learn whether a student’s components of EFL motivation varied with one’s learning situation (i.e., level of schooling), and if and what types of differences in motivation might exist across grade levels and gender. Methods: Using motivational factors derived from social and education psychology, the researchers created a 50-question Likert-scale questionnaire which was given to a convenience sample of 1,027 participants, which included students studying English at the university, junior college, high school, and junior high school level. Exploratory factor analysis and MANOVA (multivariate analysis of variance) was then used to analyze the data. Results: The data indicated that students were motivated by a variety of complex factors. For example, intrinsic factors (internally rewarding), instrumental factors (desire to learn English for a very specific purpose, e.g., finding a good job), and integrative factors (a desire to integrate with target culture) were highly motivating to students at the junior high school level, as well as junior college English majors, English language school learners, and foreign language majors. Extrinsic (motivation secondary to expecting a reward from sources outside of oneself) and instrumental motives were highly motivating for junior high school students and college students majoring in engineering, but not for other groups of students. Conclusions: The authors suggested that since Japanese
learners of English who have had negative experiences with learning English might have high levels of language use anxiety, EFL teachers should try to help their students feel comfortable, build good working relationships with them, and offer extracurricular opportunities for learning English. *Relevance to current study:* The current study includes native Japanese learners of English, with whom the researchers sought to help improve their production of English speech sounds through facilitative techniques such as those recommended in the study of Kimura et. al.


**Objective:** The researchers examined the impact of the /l/-/r/ consonant vowel stimulus duration on the MMNm (magnetic mismatch negativity response) in native Japanese speakers who were not easily able to discriminate the /l/-/r/ consonant contrast. **Methods:** Nine right-handed native Japanese speakers, living in English-speaking countries for three months or less, first were assessed on their ability to discriminate between /l/ and /r/ minimal pair contrasts; their accuracy averaged 61%. Then, for the first portion of the study, the researchers used a cascade formant synthesizer to generate synthetic /la/ and /ra/ sounds. The duration of the stimuli was set to 110 ms (short-duration) and 150 ms (long-duration). Then, for the second portion of the study, the researchers synthetically removed formants 1 and 2 from the previously synthesized /la/ and /ra/ stimuli, so that only the F3 formant remained. Each of the participants listened to stimuli through binaural presentation via insert phones, while watching a silent film and being instructed to not pay attention to the auditory stimuli. The brain’s magnetic responses were recorded with gradiometers which were placed over the right and left temporal sites, and the MMNm was
calculated by subtracting the magnetic response to the frequently presented stimuli from the magnetic response produced in response to the deviant (less-frequently presented) stimuli. The magnetic strength of the MMNm was calculated in both the right and left hemispheres of the brain. Then, Brain Electric Source Analysis (BESA) was used to analyze MMNm equivalent current dipoles, and estimate source locations of neural activity. **Results:** The MMNm response for /ra/ was significantly enhanced for short-duration stimuli in comparison to longer-duration stimuli bilaterally, yet an enhanced MMNm response to short-duration /la/ stimuli was found unilaterally in the left hemisphere of the brain, and not in the right hemisphere. In the second experiment, when both formant 1 and formant 2 were removed from the stimuli, leaving only the F3 component, the brain demonstrated a bilateral MMNm response to both short-duration and long-duration /la/ and /ra/ stimuli. **Conclusions:** These findings support the authors’ conclusion that vowels have the potential to mask preceding consonants, as the results suggest that in non-vowel F3 stimuli, masking is absent. As MMNm occurs when the brain detects a difference between two stimuli, and vowels which are long in duration can have a “masking effect” on the consonants that come before them, researchers have observed that the studied Japanese individuals can not discriminate the difference between the /l/-/r/ contrast when a long-duration vowel comes right after the target phonemes. To avoid the masking effect, make the vowel shorter, and this can help Japanese detect difference by MMNm. **Relevance to current study:** The current study aims to help native Japanese learners of English to improve their speech production quality of the /l/-/r/ contrast, which is dependent on their ability to accurately discriminate between the target phonemes /l/ and /r/. The findings of the study conducted by Koyama et. all suggest that in order to help avoid the masking effect and help native Japanese
learners of English to experience the MMNm effect on a neural level, it could be helpful to make
the vowel shorter in duration, which is an important consideration for this study.

Lev-Ari, S., Keysar, B. (2010). Why don’t we believe non-native speakers? The influence of

**Objective:** The researchers sought to elucidate whether native English speakers evaluated
statements produced with foreign accented speech as less true than those statements produced
without a foreign accent. Additionally, the researchers sought to test whether native listeners
would rate statements produced with accented speech as more credible, if the listeners were first
primed with the knowledge that processing highly accented speech was difficult. **Methods:** Nine
English speakers (three native speakers, three mildly accented non-native English speakers, and
three heavily accented English speakers) were recorded while reading forty-five trivia
statements. Thirty native American English speakers rated the veracity of the statements on a
visual analogue scale ranging from “definitely false” to “definitely true.” A second group of
twenty-seven native American English speakers performed a similar task, after being told that
the experiment sought to research the impact of how the difficulty of processing speech would
impact their perception of the credibility of the speaker. **Results:** When listeners were made
aware that decreased processing fluency could negatively impact their perception of another
speaker’s credibility, they tried to compensate. While their increased awareness positively
impacted the credibility ratings given to mildly-accented speakers of English; the credibility of
individuals with heavily accented speech was rated equally poorly across the two groups.

**Conclusions:** The researchers concluded that native speakers of American English perceived
accented speech as being less credible than non-accented speech; this is posited to be due to the
fact that difficulty comprehending the statements led to reduced processing fluency. However, rather than filtering the statements as being simply harder to understand, listeners instead filtered the statements as being less truthful. **Relevance to current study:** The current study seeks to help accented speakers of English to reduce their accents, specifically within the /l/-/r/ contrast; the study by Lev-Ari and Keysar seeks to address key concerns experienced by speakers similar to those in the current study’s experimental group, and might lead one to conclude that decreasing the accentedness of one’s speech might lead to increased credibility ratings.


**Objective:** This study sought to gain a qualitative phenomenological understanding of how Asian learners of English perceive electropalatography, specifically their expectations regarding use of EPG, negative and positive comments about EPG when used in a classroom setting, what challenges the students experienced while using EPG in the classroom, and also what recommendations the ESL students had regarding improving EPG pronunciation classes.

**Methods:** Nine Asian ESL learners from South Korea, Taiwan, Japan, and China, who all had difficulty producing the [l, r, θ, s, ʃ, ɪ, and i:] phonemes in English, were selected as research participants. Each participant attended 7 weeks of a supplementary group pronunciation course utilizing EPG to facilitate the improved acquisition of English phonology. To collect qualitative and quantitative information about the participants’ experience while using EPG, researchers administered surveys, interviews, and also performed observations of the participants.

**Results:** Data obtained from participants indicated that participants believed EPG was socially acceptable EPG in a classroom setting, as long as their other classmates were also using the EPG.
Additionally, participants noted that while adapting to the sensor included initial discomfort and pronunciation distortion, that these side effects diminished with time and adaptation. Several participants also noted that they wished that the intervention sessions could have been longer.

Conclusions: It is recommended that one should realize that although EPG assisted pronunciation instruction helps improve phoneme production, it does not necessarily result in the acquisition of a native-like accent, as this is dependent to a great extent on suprasegmentals, rather than on phonemes alone. Future research should include more participants from a wide variety of cultural and linguistic backgrounds, longer EPG treatment, and should focus on improving pronunciation from single words to spontaneous speech. Relevance to current study: This study also utilizes electropalatography as an intervention to assist Asian learners of English to improve their English phonology.


**Objective:** This study sought to gain a greater understanding of the beliefs and opinions held by native Japanese students and native and non-native Japanese teachers regarding English language teaching and learning in Japan. **Methods:** Researchers administered a 36-item survey to 82 college teachers and over 300 college students, with questions addressing instructional areas, objectives and goals, methods and styles of instruction, materials used in teaching, and important cultural considerations. Then, the researchers performed statistical analysis to further visualize the difference in opinion between teachers and students. **Results:** The researchers noted that teachers and learners of Japanese have slightly different preferences for preferred teaching styles
and focus of English language classes. The teachers showed preference for a learner-centered approach and a focus on overall fluency, while the students showed preference for listening to lectures, learning pronunciation, and having a focus on accuracy. Conclusions: To help reduce student anxiety, teachers should be prepared to either adapt their teaching style, or explain to their students how they can best learn in response to the teacher’s teaching styles. Relevance to current study: The current study will be examining the impact of highly-specialized technology to improve the pronunciation of native Japanese speakers; this addresses the need indicated by native Japanese learners of English learning in Japan that a focus on pronunciation training would be helpful for them as they learned English as an L2.


Objective: This article seeks to describe the language learning styles commonly found in multicultural ESL classrooms; six case studies are presented, and management of optimizing teacher and student styles is discussed. Methods: This article summarizes descriptions of learning styles, learning strategies, techniques, information on style conflicts between teachers and students, and suggestions to optimize teaching style to best suit multicultural students. Results: The authors noted, citing Harshberger’s 1988 study, that Japanese students desire to receive correction rapidly and constantly; furthermore, they do not like several correct answers. The authors noted that in Reid’s 1987 study, Japanese students were the “least auditory” learners in comparison to other cultures studied. Conclusions: The article concludes that although the generalizations made about cultures may not apply to each individual of a specific cultural
heritage, that the descriptions given are rooted in observations made in a scientific way. Additionally, the authors hypothesized that Japanese students may be reticent to take part in speaking activities because of a traditional focus culturally on belonging to a group and saving face. **Relevance to current study:** The current study will be examining a specific method for teaching ESL to a multicultural group of individuals, native Japanese learners of English; as such, it is important to consider the preferred learning styles of Japanese individuals.


**Objective:** The authors sought to further elucidate the impact of a non-native accent on the career satisfaction and advancement of non-native speakers of second languages through creation of a conceptual model which assessed the affective, behavioral, and cognitive reactions of individuals with non-native accents. **Methods:** The authors built a conceptual model based upon a wide variety of research articles, to further clarify the ways in which managers responded to non-native accents of L2 speakers, and the ways in which L2 speakers responded to their own accented speaking. **Conclusions:** The authors provided support for several of their hypotheses, namely, the ideas that when managers at work listen to accented speech, they experience reduced perceptions of cognitive fluency. This leads the managers to have negative perceptions of the accented L2 speakers, develop lower expectations of job performance abilities of accented L2 speakers, and that these difficulties will lead to less positive regard (i.e., displays of acceptance, warmth, and value) towards the accented L2 speaker, and to assume a controlling management style of accented L2 speakers. The authors propose that when an accented L2 speaker recognizes
that their manager has low positive regard and low expectations for them, that they would feel excluded, devalued, and unappreciated, and perhaps to utilize avoidance strategies at work through defensiveness and disengaging from tasks that might seem threatening. Additionally, the authors suggest that the low expectations of the managers for accented L2 speakers in addition to the behavioral and affective reactions of managers would lead to the assignment of unimportant job responsibilities, leading to less career satisfaction and advancement. The authors also provided research evidence to support the idea that when an L2 speaker holds a prestige accent, that the negative impacts of the above study might be mitigated.

Relevance to current study: The current study seeks to help accented L2 speakers improve their pronunciation of speech sounds, and Russo’s study seeks to address key concerns impacting accented L2 speakers in the workplace.


**Objective:** The objective of this study was to study the impact of using EPG as a tool to aid in accent modification of native Thai speakers in developing the contrasts /s/-/ʃ/, /t/ - /θ/, and /l/ - /ɹ/.

**Methods:** Three native speakers of Thai volunteered to participate in the study. Each had studied English as taught by native Thai speakers, and their first exposure to being taught in the standard American English dialect occurred within 14 months of their participation in the study. Kay Elemetrics Corporation produced the palatometry system used to train the subjects. The treatment contrasts were chosen based off of an initial evaluation of the participants’ productions of phonemes in both Thai and English. Participants attended between 39-48 forty-five minute sessions, which were held twice per week. In each session, the participants practiced
conversation for 10 minutes (without EPG visual support), were tested on paired contrast productions, and then after testing participated in producing the phonemic contrasts with visual support from the EPG and clinician guidance. **Results:** Each of the participants showed changes in articulation which applied to both previously-learned sounds as well as to new sounds. **Conclusions:** The authors suggest that EPG provides feedback that helps in changing or modifying errored consonantal production, even if the errored production has been habituated over many years, and that progress can be made more quickly than with traditional types of therapy. Additionally, participants learned completely new sound articulatory patterns, such as the very visual “th” more quickly than they learned to change old articulatory patterns such as “r” or “ch.” **Relevance to current study:** This study also seeks to use EPG as an intervention technique to help those in the Asian population improve their articulatory contrasts between phonemes such as /r/ and /l/.


Retrieved from https://open.library.ubc.ca/cIRcle/collections/ubctheses/24/items/1.0073242

**Objective:** The objective of this study was to research the efficacy of using ultrasound as a tool to help teach the /l/ and /r/ contrast to native Japanese learners of English. **Methods:** Six native Japanese learners of English were recruited to participate in the study; they received four one-on-one sessions to practice perception and production of the /l/-/r/ contrast using ultrasound. Audio recordings of their voices were collected both before and after treatment.
Results: The study gave an in-depth overview of Japanese liquid phonology. Conclusions: The authors concluded that ultrasound can help play a role in training native Japanese speakers to increase their production accuracy of the liquids /r/ and /l/ in the contexts of both singletons and clusters, even if the auditory-perceptual analysis may not yet yield perfect ratings. Relevance to current study: The current study also attempts to use a form of visual biofeedback to help native Japanese speakers improve their productions of the liquid /l/-/r/ contrast.


Objective: The researchers sought to further elucidate whether phonetically variable speech, delivered via computerized discrimination training, could be effective in helping Japanese speakers in their /l/ and /r/ perception and production. Methods: The researchers compared discrimination training to traditional identification training by presenting 10 sessions of discrimination and identification training to adult Japanese speakers. Baseline, mid-teaching, and post-teaching tests of /r/-/l/ production, auditory discrimination, categorical discrimination, and identification were collected. Results: After discrimination and identification training, the native Japanese learners of English demonstrated improved accuracy of /l/ and /r/ production and perception. Conclusions: Although both discrimination and identification training helped native Japanese learners of English to improve the accuracy of their /l/ and /r/ production and perception, there was not much benefit gained by using both discrimination and identification training in tandem. Rather, as long as the method selected for instructive use includes high phonetic variability, both discrimination and identification training methods have similar impacts.
in helping L2 learners improve their speech sound production and perception. *Relevance to current study:* The current study also seeks to help native Japanese speakers of English to improve their production of the English phonemes /l/ and /ɾ/.
APPENDIX B

Informed Consent

Consent to be a Research Participant

Introduction
The purpose of this study is to investigate the use of visual information to assist learning English as a second language. This experiment is being conducted under the supervision of Shawn Nissen, Ph.D., an associate professor in the Department of Communication Disorders at Brigham Young University. You have been invited to participate because you are a native English speaker with no known history of a speech, language or hearing problems.

Procedures
Participation in this study will involve up to two visits of approximately one hour each, which will take place in a research laboratory in the John Taylor Building at BYU. Each session will involve four ten-minute blocks, with a rest period in between. First, you will receive a hearing screening. Then, you will receive training on how to use the audio perceptual rating system, and have the opportunity to select a comfortable listening volume for the study. Then, you will listen to a series of sounds, words, or sentences produced by second language learners and rate each audio sample on pronunciation clarity by using a computer mouse to move a slider button.

Risks/Discomforts
There are minimal risks for participation in this study. You may encounter some discomfort from wearing the over-the-ear headphones, but you will be allowed to customize the fit of the headphones at the beginning of the session so that they fit comfortably.

Benefits
There are no direct benefits to you. It is hoped this study will provide understanding in developing more effective approaches to assist the learning of English as a second language.

Confidentiality
All information provided will remain confidential and will be reported only as group data with no identifying information. All data, including records of your listening responses, will be kept on password-protected computers in a locked laboratory and/or on a password protected cloud and only those directly involved with the research will have access to them.

Compensation
You will be compensated $15 per hour for your participation.

Participation
Participation in this research study is voluntary. You have the right to withdraw at any time or refuse to participate without penalty.

Questions about the Research
If you have questions regarding this study, you may contact Shawn Nissen, Ph.D., at (801) 422-5056 or shawn_nissen@byu.edu.

Questions about your Rights as Research Participants
If you have questions regarding your rights as a research participant, you may contact the BYU IRB Administrator, A-285 ASB, Brigham Young University, Provo, UT, 84602 or at (801) 422-1461.

I have read and fully understand the consent form. Any questions have been answered to my satisfaction.

I give my consent to participate in this research.

Signature: __________________________ Date: __________
Printed Name: __________________________

IRB NUMBER: IRB2019-395
IRB APPROVAL DATE: 01/23/2020

BYU