Teaching English /ɹ/ to Korean Adult EFL Learners: The Role of Corrective Feedback Type in Phonological FFI

Reginald M. Gooch

Department of Integrated Studies in Education

McGill University, Montreal

April 2015

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Arts in Second Language Education

Acknowledgments

I would like to begin by thanking my amazing co-supervisors, Dr. Roy Lyster and Dr. Kazuya Saito. Even with their hectic schedules, both of my supervisors made sure to send me feedback quickly, time after time, draft after draft. Not only that, but my supervisors were both great motivators as well. Each of our meetings took different shapes, but one constant was that each and every time after I met with either Dr. Saito or Dr. Lyster I felt more refreshed ready to tackle my research than I had before the meeting. I want to especially thank Dr. Lyster for his help in guiding me through all of the logistics of conducting a study, as well as helping me to break it down step by step so that I did not get lost along the way. Dr. Lyster also aided me greatly, both in securing funding for my study, and in contacting and recruiting Dr. Saito to be a co-supervisor for me. Without the cooperation of Dr. Saito I could hardly imagine what the process would have been like to attempt a replication study. Dr. Saito was very helpful in allowing me use of classroom materials from his original study, and in walking me through the particular steps I would need to take to mount a successful replication. Dr. Saito was also particularly insightful in his recommendations about how to analyze my data, and I would not have devised my system for tracking feedback and responses without his advice, as well as the foundational work of Dr. Lyster. I cannot list all the ways my professors helped me, but lastly I want to acknowledge the huge amount of help Dr. Saito gave me with the statistical analyses for this study.

Secondly, I want to thank Professor Jin-ho Kang and Ms. Gwangmi Yoo from Kyungpook National University in Korea. I am very grateful to Prof. Kang for giving me permission to conduct my study at the university, and to Ms. Yoo for all of her help in organizing the logistics to make that happen. Prof. Kang and Ms. Yoo were very gracious hosts and made

me feel very welcome as I carried out my study. Both of them took time out of their schedules, and went beyond the scope of their normal workload to help me.

Thirdly, I would like to thank my friends, classmates, and teachers at McGill for all of their help throughout this study. I want to especially thank Mrs. Victoria Skipper for her work coding video data, Miss Misung Kwon for assistance translating documents, Mr. Jungkyu Park for help with statistical analysis, and Mr. Jeong Hwan Park for his help creating graphs and tables. I would also like to thank Professor Anila Asghar for her guidance as I created my thesis proposal at the early stages of this project. Too many friends, teachers and classmates provided me help and support on this project to mention them all, but before I close this paragraph, I want to acknowledge the help and assistance of Mr. Andrew Lee. Andrew gave me assistance and advice at nearly every point in the process of conducting my thesis, especially the parts pertaining to phonological learning in particular.

Lastly, I would like to acknowledge the love, help, and support of my family throughout this process. My parents have always given me unconditional support throughout all my ventures in life, and this thesis was no different. I would also like to thank my sister, Ivy, who helped me to prepare some of the testing material. I especially want to thank my fiancé, not only for listening to me talk about this thesis for months on end, but also for her very active support in helping me organize the particulars of conducting a study in Korea, from finding accommodation to understanding websites written in Korean.

This thesis received financial support from a Graduate Research Mobility Award, an EGSS Conference Masters Proposal Award, a McGill Graduate excellence award, and a SSHRC grant (#410-2011-0671) awarded to Dr. Roy Lyster.

Abstract

Ever since the decline of audio-lingual language pedagogy, phonological teaching has been a neglected area in applied linguistics. Building on the work of Saito (2013a), this study helped to address this neglect by integrating pronunciation teaching into a communicative classroom through the use of form-focused instruction (FFI). Additionally, the study investigated the effects of three different feedback conditions on perception and production of the problematic English sound, /ɹ/. Twenty-two Korean learners of English as a foreign language were divided into three groups: prompts (*n*=6), recasts (*n*=7) and control (*n*=9). Each group received four hours of instruction, distributed over two weeks, in which they were taught how to structure an argument in English through activities designed to simultaneously focus their attention on /ɹ/, including explicit articulatory instruction. Lessons were video-recorded, which allowed all instances of feedback and individual learner responses occurring during the instruction to be identified and coded.

Perception of /ɹ/ was measured using a two-alternative forced-choice identification test. Production was measured through both spontaneous and controlled tests and judged by 5 English native-speaker listeners. Results from this study suggest that recasts were helpful in the improvement of controlled production of /ɹ/ for untrained items, and that prompts were helpful not only in this area, but also in spontaneous production of /ɹ/. Patterns of learner response suggest that students were pushed by prompts to improve intelligibility through the adjustment of interlanguage strategies.

Résumé

Depuis le déclin de l'approche audio-linguale, l'enseignement phonologique se fait négliger en didactique des langues. Les recherches de Saito (2013a) ont toutefois mis fin à cette négligence et c'est en s'appuyant sur celles-ci que la présente enquête se propose d'intégrer l'enseignement de la prononciation dans une classe communicative. L'enquête avait pour but d'examiner les effets des reformulations et des incitations ainsi que leur absence sur la perception et la production du phonème anglais /ɹ/. Vingt-deux participants coréens étudiant l'anglais comme langue étrangère en Corée du Sud ont formé trois groupes : groupe-reformulation (n = 7), groupe-incitation (*n* = 6) et groupe-témoin (*n* = 9). Une intervention de quatre heures échelonnées sur deux semaines a servi à enseigner à chaque groupe comment structurer un argument en anglais au moyen d'activités conçues de manière à attirer, par la même occasion, leur attention sur le son /ɹ/. L'intervention comprenait également un enseignement explicite portant sur l'articulation. Comme toutes les leçons ont été filmées, il était alors possible d'identifier et de coder toutes les instances de rétroaction ainsi que les réponses immédiates des participants.

La perception de /ɹ/ a été mesurée en utilisant un test d'identification forcée à deux choix, alors que la production a été mesurée par des tests de production contrôlée et spontanée dont les résultats ont été jugés ensuite par 5 juges anglophones. Les résultats indiquent que les reformulations semblent efficaces pour aider les apprenants à améliorer leur production contrôlée des items ne paraissant pas durant l'intervention. De leur côté, les incitations se sont avérées efficaces non seulement pour la production contrôlée mais aussi pour la production spontanée du /ɹ/. L'analyse des réponses suite à la rétroaction suggère que les incitations ont permis aux apprenants de restructurer leur interlangue et donc d'améliorer leur compréhensibilité.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Corrective feedback (CF) plays a pivotal role in second language (L2) learning. Not only can it aid L2 learners in the automatization of an L2 form (Ranta & Lyster, 2007), but it can also help learners increase their accuracy in the L2, and has generally been shown to contribute to better overall performance than similar instruction without CF (Lyster, Saito, & Sato, 2013; Ellis, Loewen & Erlam, 2006). This study looks at CF within form-focused instruction (FFI), a form of language pedagogy designed to draw learners' attention to form within a primarily communicative environment. Over the past 20 years, many studies have investigated the role of CF and FFI on L2 learning, but their focus has been primarily on morphosyntactic targets. Recently, researchers have also begun to research the potential of FFI for teaching of phonological targets. This research has shown promise, but has yet to tease out the details of the effectiveness of various components of FFI for phonological learning. This study aims to contribute to the disentanglement of CF effects for the phonological target English /ɹ/.[1]

Research on the effects of CF for morphosyntactic targets has shown diverse effects on L2 learning for different types of CF, and there is no reason to assume that effects would be any less various for phonological targets. However, the effects of different types of CF on phonological learning have barely begun to be investigated. To address this, the current study builds on the initial investigations of Saito and Lyster (2012) and Saito (2013a), which investigated the role of recasts for L2 phonological learning. In order to provide a more nuanced

[1]In this study, the decision was made to use the notation /ɹ/ rather than [ɹ], to cover allophonic variation, as the study features /ɹ/ in various phonetic conditions. This is also in keeping with the norm of L2 phonetics and phonology studies of this kind (e.g., Flege, 1995; Saito, 2013).

account of the roles different types of CF might have in L2 phonological learning, the current study tests three different feedback conditions to determine each one's effect on perception and production of English /ɹ/. The feedback conditions were recast (a reformulation of the student error), prompt (CF which provides negative, but not positive evidence), and no CF.

Not only is it important to investigate the role of CF in learner acquisition of L2 phonological targets, but it is also important to test the generalizability of FFI across different learning contexts. With this in mind, the current study brings the teaching techniques of the precursor study to a new environment. Whereas the precursor study involved beginner-level Japanese learners of English, the current study involved intermediate- to advanced-level Korean learners of English. Because of phonological similarities between Korean and Japanese, and the similar difficulties both learner groups have demonstrated with the target sound, /ɹ/, it is hypothesized that the FFI treatment will remain effective, even for a new learner group, and for more advanced learners.

This thesis includes a total of six chapters, including the introductory chapter presented here. In Chapter 2, the literature surrounding FFI, CF, sound learning and English /ɹ/ are presented. Through an examination of the relevant literature, this chapter lays out the gaps and unanswered questions, which need to be, and are addressed by this thesis.

Chapter 3 lays out the methodology of this study. While the current study is largely a replication of Saito (2013a), it also differs in a number of important respects, including the first language (L1) background of the participants, the context in which the study took place, the CF conditions of the groups in this study, and the method of analysis of CF and learner responses employed.

Chapter 4 presents the results of this study and is followed by a discussion of those results in Chapter 5. Specifically, Chapter 5 will address differences found among the learning outcomes of the groups in this study, and among response patterns to CF given throughout the study. The study will close with a brief conclusion in Chapter 6, and suggestions for the direction of future research.

Chapter 2

Pronunciation Instruction and FFI

In the field of Second Language Acquisition (SLA), few studies have focused on finding effective means to teach pronunciation in the classroom. Yet good phonological development is important for any language learner; not only does it impact the ability of learners to effectively communicate, it has also been linked to learners' orthographical development (Wang, Park, & Lee, 2006), feelings of belonging (Gluszek, 2010), confidence (Tang, Zhang, Li, & Zhao, 2013), willingness to communicate (Thomson & Derwing 2014), and overall language development (French, 2006). This study focuses on developing effective pedagogical techniques for pronunciation teaching in the context of a Korean EFL classroom.

Korean learners of English often face difficulty with English /ɹ/ (Cho & Park, 2006), and this difficulty can persist even after years of living in an English-speaking environment (Tench, 2003). Indeed, I encountered this issue frequently in my own experience as an EFL teacher in Korea, but like many other teachers, I was at a loss as to how to best help my students. My confusion as a teacher is unsurprising when the research literature is considered; ever since the decline of audio-lingual teaching, pronunciation teaching has largely been ignored in the field of Second Language Education (Derwing & Munro, 2005). This is not to say that theories of L2 phonological learning do not exist, but that there is not, nor has there ever been, a clear path for teachers to apply this theoretical knowledge to the real-world classroom. Through this research, I examine whether a specific set of pedagogical techniques, namely a combination of FFI and CF, is effective for applying what we know about phonological development to help learners in the L2 classroom. In the following sections I will explore the literature around L2 phonological

development, FFI and Korean learners of English /ɹ/, ultimately laying out the rationale for this study.

Second Language Phonological Learning

While there is no shortage of theories about how language learners acquire new sounds in an L2, there has been, until recently, a dearth of research into how these theories can be effectively tested and applied in a language classroom setting. In examining this issue, I will start by giving a brief overview of some of the most important theories of L2 phonological learning. I will then discuss relevant laboratory research that has been conducted in this area and, following this, I will discuss what has been done in a classroom-based setting. I will conclude this section with a discussion about some of the gaps that exist in this research literature.

Theories of L2 Phonological Development

In this subsection I will discuss three of the most important theories of L2 phonological development. The first will be the Linguistic Perception Model (Boersma, Escudero, & Hayes, 2003; Escudero & Boersma, 2004), the second will be the Perceptual Assimilation Model (Best, 1995) and the third will be the Speech Learning Model (Flege, 1995, 2003). I will give a brief overview of each theory, followed by a short analysis of pedagogical implications.

Linguistic Perception Model. Boersma, Escudero, and Hayes' (2003) Linguistic Perception Model (LPM) was initially developed as a model of L1 learning. According to this theory, infants use a gradual learning algorithm to perceptually warp sound input and eventually develop distinct phonetic categories. That is to say that while the infant may initially perceive all sound input as distinct, over the course of time, the infant is able to lump sounds that are phonemically distinct for his/her L1 into their relevant categories. It should be noted that this

model centers only on the perception of different sound categories, rather than on accurate production of those sounds by the infant.

The LPM has also been adapted for L2 learners. To explain how the LPM applies to L2 learners, Escudero and Boersma (2004) lean heavily on the concept of Universal Grammar (Chomsky, 1965). They propose that, when learning an L2, a learner initially has full transfer of L1 sound categories and constraints, but that the learner also has full access to all Universal Grammar features. In other words, when a person begins learning an L2, they will hear all phonetic input in terms of their previously constructed L1 phonemic categories. However, as time goes by, that person can adapt better perceptual strategies (that is, to pay attention to things they did not heed before) and gradually develop new phonemic categories in the L2 through a gradual learning algorithm. Effectively, once the learners have developed a given L2 phonemic category, they have acquired that sound, and from that point on it is a matter of the learner gradually fine-tuning the constraints on that category to make it increasingly accurate.

Perceptual Assimilation Model. One main difference between the Perceptual Assimilation Model (PAM: Best, 1995) and Boersma et al.'s LPM is that, unlike the LPM, the PAM is not based on the idea of Universal Grammar. Best suggests that infants learn new sounds from experience, and based on the input they receive, acquire the articulatory gestures required to produce those sounds. According to the PAM, once a person has developed L1-specific phonemic categories in this way, his/her ability to discern new sounds in an L2 is impeded (Lee, 2013). It is not, however, impeded based on Universal Grammatical constraints. Rather, in order to create new L2 phonological categories, a listener must be able to perceive the difference between an L2 sound and the nearest L1 sound. The PAM posits that the greater the perceived phonetic distance between the L2 sound and the L1 sound, the easier that L2 sound will be to

acquire (Guion, Flege, Akahane-Yamada, & Pruitt, 2000). One benefit of the PAM is that it provides for predicting which L2 sounds will be difficult for speakers of a given L1 to acquire.

Speech Learning Model. The Speech Learning Model (SLM: Flege, 1995, 2003) is similar to the PAM in many respects. Like the PAM, the SLM hypothesizes that "learners can establish new L2 phonemic categories if they detect phonetic differences between an L2 sound and the nearest L1 sound" (Guion et al., 2000). Because of this similarity, the SLM is also useful as a tool for predicting which L2 sounds will be difficult to acquire for learners of a given L1. One difference between the theories is that the SLM concentrates on how experienced L2 learners acquire sounds, rather than on which L2 sounds may be difficult to distinguish for a naïve listener, as is the case for the PAM (Lee, 2013). It should be noted that for this theory, learners develop new phonemic categories if they are able to *perceive* differences between their L1 sound and the target L2 sound, and not based on their ability to produce those sounds. This model could then be said to take a perception before production approach.

Discussion of Theories of Sound Learning. While there are significant differences between the theories discussed here (e.g., the role of Universal Grammar), what are more interesting are the striking similarities between them. In each model, learning novel L2 sounds consists in the creation of new phonetic categories. Furthermore, it should be noted that each model places speech perception at the center of learning. It may be, as in the case of the LPM, that L2 production is not even considered, or it may be as in the case of the SLM, that production is considered, but is thought to follow from accurate L2 speech perception. This has implications for pronunciation teaching. If these theories are valid, good pronunciation teaching practices should center on helping learners to create new phonemic categories through a perception-first approach. In that sense, pronunciation teaching should not focus solely on production, but rather

on overall phonological development. Through the following review of laboratory and classroom research, I will show that this conclusion has gathered some support, but that further research is needed to determine the best way to put these theories into practice in the classroom.

Laboratory-Based Phonological Research

Hardison (2010) provides a useful summary of laboratory studies that have looked into phonological perception over the last half century. Many of these studies, including Lively, Logan, and Pisoni (1993) demonstrate the possibility of teaching new L2 sounds through perceptual training. This supports the perception-first approach taken by the theories discussed in the previous subsection. In other studies (Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Hardison, 2003), perceptual training in an L2 sound was even shown to lead to more accurate production of that sound. Laboratory studies also indicate that perceptual learning may be enhanced if learners are exposed to both visual and auditory cues, if L2 sound tokens are provided in a variety of different positions, and if those tokens are provided by a number of different L1 native speakers (Hardison, 2010; Lively et al., 1993). In short, variety in the input seems to help L2 learners discern new sounds. These findings suggest that in learning new phonetic categories, a variety of input may help the learner to construct a new perceptual space (Boersma et al., 2003) or that it may help learners notice the gap between the L2 sound and the nearest L1 sound (Best, 1995; Flege, 1995, 2003; Guion et al., 2000). Of course, there are a number of issues with the findings in these studies as well.

Firstly, it should be pointed out that not every laboratory study that has undertaken to teach perception has been successful. For example, Borden (1983) was able to show short-term perceptual gains for Korean learners of /l/ and /ɹ/, but those gains were not maintained over time. Perhaps one of the reasons for this shortfall was the decontextualized nature of the teaching

techniques used (e.g., differentiation of nonsense words), which raises another, larger issue. Even though many laboratory studies show support for the theories of L2 phonological development reviewed in this thesis, these laboratory studies leave us with few teaching implications. The isolated and artificial nature of laboratory testing is very different from L2 classroom settings, which means it is difficult to tell if results would translate well. This is especially true as some of these laboratory studies (e.g., Bradlow et al., 1997) did not test whether learners were able to generalize their gains from a controlled to a spontaneous level. If a given student improves his/her pronunciation of a given word in a controlled setting, it is not necessarily an indication that the participant will be more intelligible in natural speech. While it may be more difficult to design tests that accurately measure pronunciation in a more free and communicative context, this type of test has greater authenticity and ecological validity than controlled tests (Lee, Jang, & Plonsky, 2014). An additional issue is that due to length of instruction and narrow focus on a few sounds, many of the teaching techniques used in these laboratory studies are impractical to apply in a classroom setting where other aspects of language besides pronunciation need to be covered (Saito, 2013a). For these reasons, it would be more useful to see future research carried out in more realistic settings and focused on practical teaching techniques.

Classroom-Based Pronunciation Teaching Research

Until recently, few researchers had made serious attempts at classroom-based research on pronunciation teaching. In fact, research into this area has been paid so little heed that Derwing and Munro (2005) described it as “marginalized” within the field of SLA. Fortunately, in recent years researchers have begun to delve into this forgotten realm. Early results show additional support for perception-first theories of phonological development (Lee, 2013), but what is perhaps even more exciting is that they have begun to point to what may be effective classroom

practices for pronunciation teaching. Kazuya Saito has led the field in this regard. In his studies on Japanese learners of English, Saito effectively used FFI in combination with CF to help learners acquire new English L2 sounds (Saito, 2013a, 2013b, 2014a, 2014b; Saito & Lyster, 2012). Given that recent studies have begun to provide some evidence that FFI may be more effective for pronunciation teaching than decontextualized (Focus on FormS) method (see Saito, 2012 for review), the current study aims to further examine in depth how one crucial component of FFI techniques—CF—can impact L2 pronunciation development.

Discussion of L2 Sound Learning

Throughout this section, I have reviewed the state of the fields of phonological development and pronunciation teaching, and analyzed to what degree knowledge of phonological development has been applied toward finding effective classroom teaching practices. The three theories of L2 phonological development reviewed in this chapter differ in some respects, but seem to come together on their emphasis of the role of perception in phonological learning, which they define as the establishment of new L2 phonemic categories in the mind of the learner. Research in the laboratory and the classroom up to this point seems to support this theoretical lean toward a perception-first approach. Up to this point, however, there have been few studies with direct implications for effective classroom pronunciation teaching. This trend has begun to change through recent classroom-based studies, which have effectively used a combination of FFI and CF to help learners acquire new L2 segmental and suprasegmental features. While initial results from this research are promising, more studies are needed to determine whether results will hold true across different contexts and with different groups of learners. The small number of studies looking at using FFI and CF for pronunciation

teaching and the promise held in that area combine to make it ideal ground to research at this time.

Form-Focused Instruction and its Components

Research into FFI as a pedagogical technique sprang from an analysis of the inadequacies of purely naturalistic communicative teaching (e.g., Harley, Cummins, Swain, & Allen, 1990). Since its inception, FFI has been used to teach a number of different language targets, but it is predominantly FFI's effect on learning morphosyntactic targets which has been measured (Derwing & Munro, 2005; Saito, 2013b). In this section, I give a brief background of FFI and look at some elements that have been combined with FFI in pronunciation teaching: CF and explicit instruction (EI). I look at some initial results of studies using FFI to teach pronunciation, and lay out how the current study expands upon existing knowledge in this area.

FFI

Form-focused instruction has been described as a method of language pedagogy that "draw[s] the learners' attention to form either explicitly or implicitly" (Spada, 1997). Research suggests FFI is more effective within a communicative environment (Saito, 2013b), and for the purposes of this thesis, the term "FFI" refers only to FFI that is conducted within a primarily communicative environment. The goal of FFI is to build on the effectiveness of naturalistic communicative teaching "in enabling children to attain high levels of communicative proficiency in the L2" (Day & Shapson, 2001, p. 49) while addressing its weaknesses. FFI accomplishes this by enriching input, increasing student output, increasing saliency of difficult L2 features, and providing learners with unambiguous CF (Ranta & Lyster, 2007).

Corrective Feedback

Oral CF may be given by a teacher (e.g., Ammar & Spada, 2006; Lyster & Ranta, 1997; Ranta & Lyster, 2007; Saito, 2013a) or by peers (e.g., Sato & Lyster, 2012) in response to an error. Different types of CF include recasts, metalinguistic feedback, explicit correction, repetition, clarification requests, elicitation and others (Lyster & Ranta, 1997). CF has been analyzed for its effectiveness in different contexts, with different linguistic targets, according to its degree of explicitness (Ellis, Loewen, & Erlam, 2006), the presence or absence of explicit instruction (Saito, 2014a) and whether it provides a reformulation of the student's response (Ranta & Lyster, 2007), to name just a few examples. The current study will investigate two types of CF in particular: recasts and prompts.

Application of FFI for Pronunciation Teaching

In the limited number of studies in which it has been applied to pronunciation teaching, FFI has shown great promise. In comparison studies, it has been found to be more effective than decontextualized Focus on FormS instruction and naturalistic communicative teaching (for review see Saito, 2012). FFI has been particularly effective for pronunciation teaching when it includes explicit explanations of articulation (Saito, 2013b) and when it has been combined with CF (Lee et al., 2014). One reason for FFI's effectiveness may be that it aids proceduralization of declarative knowledge (Ranta & Lyster, 2007), in this case meaning new phonetic categories. Until now, recasts have been the best-researched CF type within FFI. The nearly exclusive focus on recasts in classroom-based pronunciation instruction research may, however, be an oversight. The literature suggests that other types of CF maybe be effective when applied to the setting of phonological FFI; in the subsections below I will therefore examine not only the recast, but also another form of CF that shows promise for application in phonological teaching, the prompt.

Phonological recasts and prompts. Recasts are defined as "reformulation…of a student's utterance minus the error" (Lyster & Ranta, 1997). When designed for maximum saliency (see Sheen, 2006), phonological recasts have been theorized to be effective due to their ability to "provide students with an explicit signal of errors and with teacher pronunciation models, while at the same time, eliciting self-modified output" (Saito, 2013a, p. 520). Phonological recasts have been shown to be particularly salient when used with stress, a declarative intonation, and for one change at a time (Loewen & Philp, 2006). This kind of recast is hypothesized to provide students not only with positive evidence in the form of the teacher's model pronunciation, but also with indirect negative evidence that can help students realize their errors and self-correct.

One critique of recasts has been that students may not perceive their corrective purpose (e.g., Lyster & Ranta, 1997; Ellis et al., 2006). If this is the case, then any negative evidence they provide might be lost to the student. However, several studies have found that recasts for phonological errors are more salient than those for morphosyntactic errors (e.g., Carpenter, Jeon, MacGregor, & Mackey, 2006; Lyster, 1998), and so this problem may not be as large of an issue as in other contexts. In fact, in an observational study, Lyster (1998) found that teachers tended to recast phonological errors and that these recasts led to a substantial number of repairs.

All this being said, it is too early to conclude that recasts are the ideal type of CF to use in conjunction with phonological FFI. While phonological recasts have shown early indications of effectiveness, to my knowledge there are no specific studies systematically comparing with any other kind of phonological CF, or even investigating an alternative form of phonological CF such as prompts in isolation. Indeed, there is reason to believe that other types of CF could be even more effective than recasts.

The efficacy of recasts versus other CF types for non-phonological linguistic targets has been a matter of some debate; in fact, in a 2010 meta-analysis (Lyster & Saito, 2010) found recasts to be less effective than prompts (types of CF that do not provide error reformulation) in promoting L2 acquisition. One advantage of prompts over recasts is their more salient provision of negative evidence. While recasts have been demonstrated to be more salient for phonological errors than for grammatical errors, it is possible that prompts may prove to be even more salient. Prompts also have the advantage of placing a greater demand on the student to produce modified output, and may be more effective in aiding learners in proceduralizing their knowledge (Lyster, 2004). In terms of phonological development, it is possible that prompts may be effective at helping learners to perceive the difference between their production and the target sound, and thus refine their L2 phonetic category (as in the PAM), or to gradually refine the constraints of their developing L2 phonetic category (as in the LPM). Prompts may also be as effective or more effective than recasts in helping learners to adjust their articulatory configurations for the L2 sound because of their ability to draw learners' attention to form (or in this case, pronunciation) in a way that allows them to re-analyze and modify their output as they test new hypothesis about the L2 (Lyster, 1998, p. 191). Because of the promise that prompts show for phonological CF, the lack of research that has been conducted on their effects for phonological teaching up to this point, and the evidence of their effectiveness for other linguistic targets vis-à-vis recasts, it is important at this time to investigate their effectiveness in pronunciation teaching.

Explicit instruction. The role of EI in phonological FFI has just begun to be investigated. While this term has been used in more general ways within the field of SLA, for the purpose of phonological FFI, EI has been used to refer to an explanation of place and manner of articulation for a target sound. These types of explanations have been investigated in conjunction

with decontextualized practice as well, but have just begun to be examined in conjunction with FFI (Saito, 2013b). EI is theorized to help orient students' attention to phonetic information (Thomson & Derwing, 2014), thereby giving students valuable declarative knowledge in order to speed their creation of new phonetic categories. FFI with EI has been shown to be more effective than FFI without EI (Saito, 2013b); that being said, there have been too few studies conducted involving FFI and EI to be certain in this conclusion.

The aim of the current study. In light of the ongoing debate over the effectiveness of CF type, it is prudent at this time to investigate the relative effects of recasts and other forms of CF in pronunciation teaching. Additionally, the application of FFI for pronunciation teaching has yet to be explored within a variety of different contexts, and it is important to see how these new conditions will affect not only FFI, but also FFI with EI. This is especially important, as both learner background and context have been shown to greatly influence the impact of pronunciation interventions (Lee et al., 2014). With this in mind, the present study investigates the efficacy of different CF types within FFI with a new group of learners (Korean learners of English) in a different language-learning environment. In the following section, I detail the sound target and learner group that is investigated in this study.

Korean EFL Learners and English /ɹ/

Korean L1 speakers tend to have difficulty with English /ɹ/. According to the Perceptual Assimilation Model (Best, 1995), Korean perception of English /ɹ/ and /l/ should fall into either the "single-category" (two L2 sounds are perceived as one L1 sound) or "category-goodness" (one of the two L2 sounds is more similar to the L1 sound) pattern, indicating that Koreans will face perceptual difficulties with these sounds. This theoretical prediction is echoed in empirical data. For Koreans, issues with pronunciation and perception of /ɹ/ are persistent, often continuing

even after many years living in Anglophone countries (Tench, 2003). While there are many potential groups of learners and language targets which should be investigated, it is especially interesting to investigate Korean learners of /ɹ/ not only because of the well-documented difficulty this target presents, but also because Koreans constitute such a large population of EFL students.

In addition the need for effective instruction of /ɹ/ for Korean learners demonstrated above, another factor makes this group ideal for future studies. Many of the studies that have investigated the efficacy of FFI for pronunciation teaching to date have looked at Japanese learners of English /ɹ/ (e.g., Saito, 2013a, 2014b; Saito & Lyster, 2012). Because the Korean and Japanese L1 phonemes /ɾ/ are similar, and are each language's closest L1 phoneme to English /ɹ/, (Hardison, 2003; Sakow, 1990), it is reasonable to expect that FFI techniques used with Japanese learners might also be effective with Korean learners. This hypothesis is further supported by evidence that phonological recasts can be highly salient for Korean EFL students (Sheen, 2006). Altogether, the degree and persistence of the difficulty Korean learners face with English /ɹ/ and the expectation that FFI treatment might be effective for this group combine to justify it as the focus of the current study.

Summary

Theories of phonological development point to the importance of aiding students in developing speech perception, and thus in the creation of new phonetic categories. These theories have been supported through laboratory research, and more recently have begun to be investigated through classroom-based research. Further research is needed, however, before it can be definitively determined whether a perception-based pedagogical approach is indeed the most effective method of pronunciation teaching, and if it is, which specific teaching techniques

are most effective in implementing that approach. One approach that shows promise in aiding phonological development is FFI, especially when combined with CF and EI. Once again, however, there is insufficient research to conclude that this is indeed the most effective approach for pronunciation teaching. Additionally, it bears investigation whether one particular type of CF is best combined with FFI for aiding phonological development, or if perhaps a mixture of CF types would be ideal. A fruitful group to investigate is Korean learners of English /ɹ/. The present study on this group aims to address a number of research gaps raised in this introduction, such as the replicability of favorable FFI+EI+ CF outcomes to different (non-Japanese) learner groups, the viability of a perception-first based teaching approach, and the relative efficacy of different CF types (recasts vs. prompts).

Chapter 3

Methodology

The present study followed a quasi-experimental pre- and posttest design in a simulated classroom setting modeled on the design found in Saito (2013a). This design allowed for greater experimental control than an intact classroom, but offered a more realistic setting than a laboratory study. Participants in this study were randomly assigned to either a control group, or one of two experimental groups. Each group followed the same procedure of pretest, treatment and posttest. Treatment consisted of a four-hour language course, which was split into four separate one-hour sessions over the course of two weeks. The variable that differentiated these groups was the type of corrective feedback they received during class sessions. Groups one received recasts, group two received prompts and group three (control) received no corrective feedback. Each class session in this study was video-recorded for later analysis of corrective feedback (teacher-initiated or otherwise) and student responses or self-corrections. For a visual representation of the study design, see Figure 1.

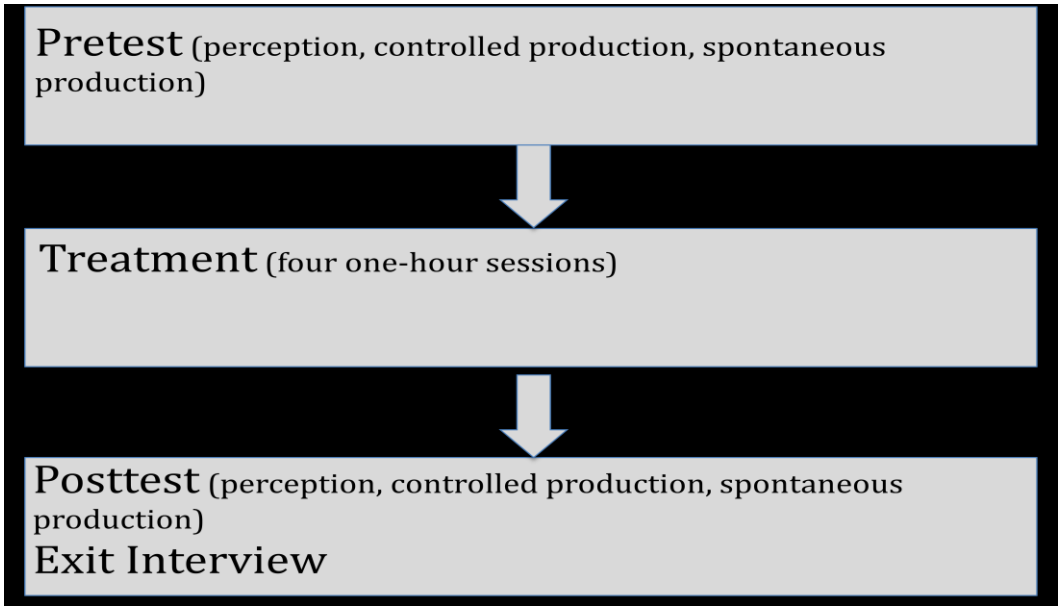


Figure 1. Study design.

Setting

This study took place at a public university in South Korea. Pretests, posttests and classes were conducted in classrooms at the university. Judgment sessions (see p. 34) were conducted in a language laboratory at a public university in Canada.

Participants

The participants in this study were 22 undergraduate Korean learners of English as a foreign language (10 male, 12 female) recruited from a public university in South Korea. These participants were recruited through an advertisement posted on the school's English department's webpage, and through flyers posted in the English department building and Humanities building. To help put students at ease and to ensure that they fully understood what was being asked of them, information about the study was provided in Korean. Participants were randomly assigned to one of three groups (see Table 1) and each participant was compensated 20,000 Korean Won, equivalent to about $20 CAD. Each group underwent the same tests and received the same instruction (FFI with explicit phonetic instruction and identical content); the differentiating feature of each group was the type of corrective feedback received. More detail will be given in the instruction section.

Table 1

Groups and instructional conditions

Groups	Argumentative skills content	FFI on /ɹ/	CF type
No CF (*n* = 9)	Yes	Yes	None
Recast (*n* = *7*)	Yes	Yes	Recast
Prompt (*n* = *6*)	Yes	Yes	Prompt

English NS Judges

Five English NS listeners (2 male, 3 female) were recruited from a public university in Canada to rate 704 speech tokens recorded during pretests and posttests. All judges were native speakers of English under the age of 50, and had at least two years of EFL or ESL teaching experience and normal hearing. Listener judges were compensated $30 CAD for their time.

Instructor

The instructor for these classes was the author. The author is a male NS of American English who has two years experience as an EFL instructor at private language institutes in South Korea.

Instruction

For this study, the classroom instruction consisted of three aspects. The first was explicit instruction (EI) on the articulation of /ɹ/, the second was form-focused instruction (FFI), and the third was corrective feedback. All three groups received both EI and FFI. The only differentiation between the three groups was the type of corrective feedback received, and for this reason I refer to each group by its corrective feedback condition. Group one received recasts as corrective feedback for its pronunciation of /ɹ/, group two received prompts as corrective feedback for its pronunciation of /ɹ/, and group three received no corrective feedback, but was otherwise identical to groups one and two in terms of instruction, content, length of instruction etc.

Explicit Instruction

EI has been demonstrated to be more helpful for improving production of /ɹ/ when combined with FFI than when FFI alone is provided (Saito, 2013b). One explanation for this effectiveness is that EI helps students to see, hear and understand the difference between their

nearest L1 phoneme (in this case, Korean /ɾ/) and the target L2 sound, /ɹ/. In this way, EI may aid in the development of new phonological categories. Additionally, participants may be able to use articulatory explanations to help produce more targetlike sounds when making repairs in response to corrective feedback. Overall, it is hoped that this kind of explicit phonological knowledge can enhance later implicit phonological learning (Saito, 2013b).

EI for /ɹ/ was given for 10 minutes at the beginning of class on the first day for all groups, and a short review of the EI was given in the first 3-5 minutes of class on all other days for all groups. EI on Day 1 was given as follows: An introduction of the /ɹ/ sound, with an exaggerated model of pronunciation from the instructor (1 min.), followed by an introduction of the different aspects of articulation of /ɹ/. For each articulatory aspect, the instructor explained what was happening inside the mouth using cross-sectional charts, and provided exaggerated demonstrations for the students of each aspect. During this process, students tried following along and making the sounds and mouth movements for themselves (8 min.). Finally, students practiced making the /ɹ/ sound by reading example sentences (1 min.). These sentences were typographically enhanced (in this case, bolded) to draw students' attention to /ɹ/ in the sentence.

Form-Focused Instruction

For comparability purposes, the same materials used in the precursor research were used as instructional materials. The content of the course was teaching English argumentation skills, but by giving the course through FFI, learner attention was focused on /ɹ/ in a number of ways throughout the course. Very minor modifications were made to take into account differences in cultural knowledge between Japanese and Korean students (e.g., changed the name of an example river from one in Japan to one in Korea). Two aspects of FFI are explained separately in

this chapter (EI and corrective feedback), but the other aspects designed to bring learner attention to the target sound /ɹ/ were as follows:

- ***Structured input*** (Van Patten, 2004): In accordance with perception-first theories of phonological acquisition (e.g., Best, 1995; Boersma, Escudero, & Hayes, 2003; Flege, 1995), participants were exposed to content requiring them to process /ɹ/ words (including /ɹ/ and /l/ contrasts) for meaning without initial pressure to produce output.
- ***Typographically enhanced input*** (Han, Park, & Combs, 2008): Throughout classroom slides and materials, /ɹ/ words were written in red and bolded and/or italicized to help participants notice the target sound.
- ***Focused tasks*** (Ellis, 2006): Participants were asked to use accurate pronunciation to complete communicative tasks (e.g., a debate about whether it is better to go running outdoors or in a gym).

Target minimal /ɹ/ and /l/ pairs (those referred to as "trained words" in the measurements section of this chapter) appeared frequently throughout the class material. These words were typographically enhanced and were embedded within both the structured input and the focused tasks.

Classroom activities. As mentioned above, the focus of the content of the course was English argumentation. Over the course of four days, students learned to critique an argument, to support their points during a debate, to structure an argument, and ultimately, to create an argument and present it orally in front of the class. Throughout these activities, students encountered and used many target /ɹ/ words, which were typographically enhanced (bolded, red, and in italics) in the class materials. For example, students critiqued the argument "Whenever I eat Japanese ***rice*** I have digestion problems. So, eating Japanese ***rice*** causes digestion problems,"

and debated the topic "Is ***read***ing comic books good for children or not?" In addition to these main activities, students also played a variety of warm-up games designed to help them practice distinguishing English /ɹ/ in their perception and production. Below is a short description of the games.

English Karuta. Thirty-six cards featuring pictures of minimally paired /ɹ/ or /l/ words were placed on the table. When the teacher said a word, students had to compete to grab the correct card as quickly as possible. At the end of the game, the student with the most cards was the winner. This game was designed to draw students' attention to the perceptual differences between /l/ and /ɹ/.

English card game. Each pair of students was given a stack of cards (36 cards in total). On each card, there were two sentences written, each of which had an accompanying picture or illustration. Each sentence was identical except for one word, which was a member of an /ɹ/ /l/ minimal pair (e.g., Josh is a good ***reader*** vs. Josh is a good ***leader***). Students took turns reading one of the sentences from the card; the other student had to guess which of the two sentences had been read. This activity was designed to promote students' awareness of their production of the /ɹ/ /l/ contrast.

Guessing game. Students formed pairs, and each pair was given a stack of cards. On each card was a target word. One student gave clues to his/her partner and attempted to get the partner to guess the word. Students who were able to guess the word after fewer clues received more points.

Corrective Feedback

As mentioned previously, there were three groups that took part in this study. Two experimental groups received different types of corrective feedback from the instructor, and the

third, control, group received no corrective feedback. Here I will provide details of the types and implementation of corrective feedback for the experimental groups.

Recast group. While there are a number of different forms a recast can take, for the purpose of this study the instructor provided recasts in the following way: Whenever a student produced a non-targetlike /ɹ/ word, the instructor would repeat the word back to the student with falling intonation. This combination of partial (one-word) recast and falling intonation has been shown to be more salient than other types of phonological recasts (Sheen, 2006), and even to predict posttest accuracy, vs. other formulations of recasts (Loewen & Philp, 2006).

Prompt group. The most easily identifiable difference between prompts and recasts as phonological corrective feedback is that while recasts provide positive evidence (i.e., the participants can hear the instructor's model pronunciation of the word) prompts do not. Like recasts, there are a wide variety of prompt types which are possible as corrective feedback. For the purpose of this study, elicitations (e.g., Pardon?; Could you say that again please?) were used. These elicitations can be considered prompts as they "[withheld] correct forms and instead [offered] learners an opportunity to self-repair by generating their own self-modified response" (Lyster, 2004 p. 405), and will be referred to as prompts throughout the rest of this study.

Measures

This study involved a pretest and a posttest, each of which was broken down into three components: a perception test, a controlled production test, and a spontaneous production test, with each test measuring a different facet of phonological learning. Test items were identical to those used in Saito (2013a), and included both trained and untrained items in a variety of phonetic contexts. Trained items were defined as those words that appeared in instructional (classroom) materials and untrained as those that did not. Using a mixture of trained an untrained

items is assumed to test the generalizability of student gains made at a segmental level. That is to say, to see whether students would be able to improve merely for "r" words, or whether they would be able to isolate their knowledge of the sound /ɹ/ and transfer it to new words. Untrained words were included at both the pretest and at the posttest in order to ensure that any differences between trained and untrained items were not due to test-retest effects. In addition to the pre- and posttest measures, video recordings of both the recast and prompt classes were analyzed for the type and quantity of corrections given and for the student responses to those corrections.

Spontaneous Production Test

In response to calls for pronunciation intervention studies to include more ecologically valid measures of production (e.g., Lee et al., 2014) this study, like its precursor, included a picture description task to measure spontaneous production. This type of task is theorized to contribute more information about the linguistic target within global speech patterns than a controlled production task alone would do (Thomson & Derwing, 2014). The spontaneous production test measured student production of 8 target r-initial words (four trained and four untrained), with a variety of vowel sounds following the initial /ɹ/. The initial /ɹ/ context was chosen for purposes of comparability with the precursor study. The words used were as follows: *read, roof, rain, road* (trained), *ring, red, route, rope* (untrained). This spontaneous production test was designed to measure student production in a communicative context, which can often be more cognitively challenging than production under controlled circumstances wherein the speaker does not have to attend to meaning and syntax. In order to create an environment for communicative language use, a number of steps were taken. To start, the spontaneous production test was administered before the controlled production and perception tests. The rationale for this was to minimize participants' focus on form during the spontaneous section (Saito, 2013a). To

further ensure a communicative language opportunity, the spontaneous production test was administered via a timed picture description task, which is theorized to be cognitively demanding (see Rau, Chang, & Tarone, 2009) in that it requires students to complete a task using target language while paying equal attention to morphosyntactic, lexical and phonological aspects of language.

The spontaneous production task was administered as follows: First, participants were given two sets of two words, some of which were target words, and some of which were distracter items. Next, the participants were given a short time (<10 seconds) to memorize those words. Third, the words were taken away and the participants were given a set of two pictures (one at a time). Students then spontaneously described the picture using the corresponding set of two words (that is to say, the first two words for the first picture, and the second two words for the second picture). In sum, the learners described sixteen pictures using a total of eight target words and various distracter words (see Appendix A). As mentioned previously, all student responses were recorded via a Roland 05 Wave recorder.

Controlled Production Test

Like the spontaneous production test, the controlled production test featured eight target words. The words used were as follows: *rink, rule, race, road (trained), reach, rude, rate, roll (untrained)*. The words selected feature word initial /ɹ/ followed by vowels of varying height and backness. For the task itself, participants were asked to read a list of 40 words, throughout which the 8 target words were interspersed. In this way, students were able to read the words without having to attend to morphosyntactic, semantic or other concerns. Responses were recorded with the Roland 05 recorder.

NS Judgment of Production Tests

During the spontaneous and controlled production tests a total of 704 items were recorded (22 students x 2 test sessions x 16 items = 704). These items were extracted and isolated using the computer software *Praat*. This was necessary especially in the case of the items from the spontaneous production test, which were produced within communicative sentences. In order to isolate these words, the author followed the process outlined in Saito (2013a), namely listening multiple times to find the onset of the target word, and including any inflected endings (e.g., roads, raining) to reduce distortion.

After all 704 items had been prepared in this way, the items were randomized for judgment by NS judges. Judgment took place in a quiet room at a Canadian university. During judgment, tokens were presented one-by-one to the NS judge on a computer screen as he/she listened to sound file through headphones. Tokens were judged on the same 9-point scale originally elaborated in Flege, Takagi, and Mann (1995) and subsequently adapted in Saito (2013a), that is *1 (Very good /ɹ/), 2 (Good /ɹ/), 3 (Probably /ɹ/), 4 (Possibly /ɹ/), 5 (Neutral), 6 (Possibly /l/), 7 (Probably /l/), 8 (Good /l/), 9 (Very good /l/).* To make a judgment along this scale, the judge simply clicked the corresponding button onscreen. There was also a repeat button available if a judge wanted to listen again to a file before making judgment. The repeat button could be clicked a maximum of three times per token. While making judgments, NS judges were asked as much as possible to base their judgments on the initial sound for each word they heard, rather than to take other factors into account, such as the mispronunciations of other sounds within the words. To judge the entire dataset, each judge took between 60-90 minutes. Judges were allowed to take a short break at any time during judgment if they began to feel fatigued.

Perception Test

The last test to be administered during the pretest and posttest was the perception test. For this test, students listened through headphones to audio files played from the author's laptop. There were two versions of this test: one for which the speaker on the sound files was the instructor (27-year-old male NS of North American English), and one for which the speaker was a female NS (24-year-old female NS of North American English). Both versions of the test featured identical items, but in a different order. All participants took both versions of the test at both pretest and posttest. Students were asked to take two versions of the perception test in order to examine whether any gains made on the perception test were specific to the instructor's voice, or whether they would be generalizable to the voice of an unfamiliar speaker.

Each test consisted of 70 items total (Appendix B). For each item, the participant listened to a sound file (either the instructor's voice for Version A or the female NS's voice for Version B), then circled the word from a minimal or near-minimal pair that they felt matched the sound they had just heard. Students were allowed to ask for any given sound file to be replayed up to a maximum of three times. Out of 70 items, 50 were part of /ɹ/ and /l/ minimal pairs (that is, 25 minimal r/l pairs) and 20 were part of distracter minimal pairs. Although some distracter pairs (e.g., berry, very) featured the /ɹ/ sound in other positions, only those with /ɹ/ or /l/ in the initial position were used to calculate student's perception scores. Of the 25 /ɹ/ and /l/ minimal pairs, 15 pairs were trained items and 10 pairs were untrained items. Because some test items were infrequently encountered words (e.g., roan, lewd), participants were told that all tokens were real English words, and to concentrate on contrasting sounds, rather than word meaning. Participant scores for the perception test were calculated as a percent accuracy.

Exit Interview

After finishing the posttest, students completed a short exit interview. Questions were posed in English, but students were encouraged to respond either in English or Korean, whichever was more comfortable for them. The interview covered their thoughts on the course and their usual habits in English class, and also gave students an opportunity to give any additional comments they wanted, or to ask any questions they had. For a full list of interview questions, see Appendix C.

Measurement of Corrective Feedback and Responses

To track corrective feedback and student responses, the author analyzed video recordings of both the recast and prompt groups. The author used a modified version of Lyster and Ranta's (1997) coding scheme of "repair" and "needs repair." Student responses and/or corrections were tracked and coded on two levels. Because not all student repair or attempted repair was initiated from teacher feedback, the first level tracked was whether the student response was:

1. Teacher-initiated (that is, a response or lack of response to corrective feedback)
2. Peer-initiated (a result of peer correction, or a breakdown of communication between peers)
3. Self-initiated (student attempted to self correct without any external push).

Additionally, in order to add sensitivity to the "repair/needs repair" coding system, a four-point scale was adopted. The scale was as follows:

0. Student does not attempt repair (used for peer- and teacher-initiated instances only).
1. Student does not repair (responds with Korean /ɾ/).
2. Student attempts change, but does not achieve targetlike /ɹ/ (interlanguage response, neither /ɾ/ nor /ɹ/).
3. Student repairs (Targetlike /ɹ/).

By accounting for both initiation and response, the researcher was able to code each instance in a total of eleven different ways (see Table 2). This coding scheme classifies interlanguage responses as all student responses that could be classified as neither /ɹ/, /ɾ/, nor /l/. This classification accounts for a wide variety of possible student responses; this thesis does not necessarily take the stance that all responses in this range represent equal progress in students' L2 phonological development, but rather used this classification in order to be able to note the instances in which students attempted to repair from /ɾ/, but were unable to produce targetlike /ɹ/. To ensure the accuracy of data coding, both the researcher and another coder (an experienced second-language instructor and MA candidate in Second Language Education) independently coded all classroom hours for both the recast and prompt groups. After conducting independent analyses, the researcher and coder met together and went over every instance of feedback or attempted repair together until they were able to reach a consensus. In this way, the researcher and coder eventually reached 100% agreement for each coded instance. Instances of feedback and repair were tracked not only for each group of students, but for each individual student, in order to provide an opportunity to look at the individual progression of each student in detail. This tracking of individual patterns is important not only because it provides another angle from which to view the data, but also because mean learning trajectories for a given set of data may

not be representative of even a single learner from within that sample (Thomson & Derwing, 2014, p. 8).

Table 2

Coding system for CF and learner-attempted repair

Initiation Type	Repair
Teacher	None attempted
	Fails to repair
	Repairs to interlanguage form
	Repairs to targetlike form
Peer	None attempted
	Fails to repair
	Repairs to interlanguage form
	Repairs to targetlike form
Self	Fails to repair
	Repairs to interlanguage form
	Repairs to targetlike form

Chapter 4

Results

Results are divided into two main sections, each of which will be divided into two subsections. The first main section will detail the overall results from the pretests and posttests, and present a short statistical analysis. The second main section will look at instances of CF and/or learner repair for the two experimental groups. It will first report overall trends among the groups, and then analyze instances of CF and/or learner repair at an individual level for selected participants.

Pre- and Posttest Results

In this subsection, we examine whether and to what degree FFI with and without CF impacted learners' phonological improvement in perception and production over time. Raw perception and production test scores are summarized in Tables 3 and 4. According to the descriptive statistics, the participants demonstrated relatively advanced phonological skills in perception (M = 82.34%, SD = 15.57, range = 53% to 100%), controlled production (M = 2.71, SD = 1.00, range = 1.15 to 5.7) and spontaneous production (M = 3.34, SD = 1.04, range = 1.1 to 5.7).

Because the visual inspection of histograms noted positive skewness, the data were transformed via log transformation to approximate log-normal distributions. To ensure the validity of any inferential statistics based on the small sample size (n = 22), we also report statistical power for each significant effect.

Table 3

Descriptive Results of the Perception Tests

Group		Trained Lexical Items				Untrained Lexical Items			
		Familiar Talker		Novel Talker		Familiar Talker		Novel Talker	
		Pre	Post	Pre	Post	Pre	Post	Pre	Post
Recast	*M*	87.29	86.57	82.86	84.71	82.14	81.43	82.86	83.57
(*n* = 7)	*SD*	15.25	20.07	21.98	20.30	17.53	19.94	20.38	23.40
Prompt	*M*	78.33	80.00	73.83	82.83	75.00	81.67	77.50	77.50
(*n* = 6)	*SD*	19.09	19.78	20.78	12.64	19.49	15.71	16.36	19.94
Control	*M*	83.56	89.00	88.89	89.67	82.78	84.44	83.89	82.22
(*n* = 9)	*SD*	12.25	10.22	11.62	10.89	11.49	15.09	10.83	14.17
Total	*M*	83.32	85.77	82.86	86.23	80.46	82.73	81.82	81.36
(*n* = 22)	*SD*	14.92	16.19	18.20	14.49	15.50	16.16	15.32	18.26

Table 4

Descriptive Results of the Production Tests

Group		Controlled Test				Spontaneous Test			
		Trained Lexical Items		Untrained Lexical Items		Trained Lexical Items		Untrained Lexical Items	
		Pre	Post	Pre	Post	Pre	Post	Pre	Post
Recast	*M*	2.17	2.31	2.31	1.89	2.71	2.37	3.17	2.80
(*n* = 7)	*SD*	0.66	0.76	0.75	0.72	0.96	1.08	0.73	0.84
Prompt	*M*	3.30	2.52	3.48	2.52	3.68	3.00	3.98	2.92
(*n = 6)*	*SD*	1.18	0.62	1.53	0.64	1.21	0.57	1.37	0.76
Control	*M*	2.44	2.33	2.82	2.57	3.10	2.90	3.56	3.38
(*n* = 9)	*SD*	0.62	0.59	0.86	0.60	0.74	0.85	1.12	0.83
Total	*M*	2.59	2.38	2.84	2.34	3.13	2.76	3.55	3.07
(*n* = 22)	*SD*	0.90	0.63	1.10	0.69	0.98	0.87	1.08	0.82

The participants' perception test scores were submitted to a 4-way ANOVA with Group (recasts, prompts, control) as a between-subject factor and Talker (trained, novel talker), Lexis (trained, untrained items), and Time (pre-, post-tests) as within-subject factors. The results did not demonstrate significant interaction effects in any contexts ($p > .05$), suggesting that most of the students did not have much room for improvement, probably due their ceiling-level performance at the beginning of the project.

Their production test scores were also submitted to a 4-way ANOVA with Group (recasts, prompts, control) as a between-subject factor and Task (controlled, spontaneous), Lexis (trained, untrained items), and Time (pre-, post-tests) as within-subject factors. Although no significant main or interaction effects were found for Group, Talker, or Lexis ($p > .05$), a main effect for Task and Time reached statistical significance, $F(1, 19) = 36.577$, $p < .001$, with large statistical power (1.00), and $F(1, 19) = 8.916$, $p = .008$, with large statistical power (0.81). According to Bonferroni multiple comparisons, whereas the recast group demonstrated significant improvement only in the context of untrained lexical items in the controlled task ($p = .046$), the prompt group demonstrated similar generalizable gains both in the controlled ($p = .017$) and spontaneous tasks ($p = .036$).

Taken together, these results are suggestive of the following patterns, which remain tentative due to the small sample size. First, all groups significantly improved their /ɹ/ performance between pre- and post-tests as a result of the instruction. Second, the recast group's gain was evident in their controlled /ɹ/ production of untrained lexical items. Third, the prompt group's improvement was apparent not only in the case of untrained lexical items but also in more spontaneous levels of speech processing.

Corrective Feedback and Learner Responses

As noted in the methodology section of this thesis, all instances of CF and/or attempted student repair were tracked and coded throughout this study. Each instance was coded via an eleven point system (presented in Table 2). In this subsection, I will present the results of this coding for the two experimental groups (recast and prompt).

Recast Group

Initiation. For the recast group, a total of 99 instances of either CF or attempted repair were recorded. Of these 99 instances, 68 (69%) were teacher-initiated (that is, the teacher gave a recast in response to a student error), 5 (5%) were peer-initiated (CF given by a classmate in response to a student error), and 26 (26%) were self-initiated (student attempted self-repair without CF from teacher or peers). From the data available, it appears that these proportions held fairly steady across the four days of this study (though, due to video file corruption, there was a substantial amount of data missing from Day 2 especially).

Response. When CF was initiated by the teacher (68 total instances), students attempted repair 54 times, or at a rate of about 80%. Of the times in which students attempted repair, 6 responses (11%) were Korean /ɾ/, 15 responses (28%) were interlanguage, and 33 responses (61%) were English /ɹ/. There were a total of only 5 instances of peer-initiated CF recorded, but among those responses, repair was attempted 3 times, or at a rate of 60%. Of the three instances in which repair was attempted, participants repaired to interlanguage twice and to English /ɹ/ once. Over the course of the study, students attempted to self-repair without receiving external CF a total of 26 times. Five of those times (19%), students repaired to interlanguage, and 21 of those times (81%) students repaired to English /ɹ/. An analysis of student responses over time reveals that even at the beginning of the study, this group rarely repaired to Korean /ɾ/, and was often able to produce English /ɹ/, suggesting that these participants could be considered to be

relatively advanced learners. One additional trend of note is that students attempted repair after CF at a higher rate as the study progressed, suggesting that over the course of the study, phonological recasts may have become increasingly salient to the learners. The results are visually presented in Table 5.

Table 5

CF and learner repair, descriptive data for recast group

	Day 1	Day 2	Day 3	Day 4	Total
Korean /ɾ/	1 (5%)	1 (33%)	4 (7%)	0 (0%)	6 (7%)
Interlanguage	5 (25%)	1 (33%)	14 (27%)	2 (22%)	22 (26%)
English /ɹ/	14 (70%)	1 (33%)	33 (65%)	7 (78%)	55 (66%)
Total No. CF	26	3	60	10	109
Total No. attempted repair	20 (77%)	3 (100%)	51 (85%)	9 (90%)	83 (76%)

Prompt Group

Initiation. For the prompt group, a total of 142 instances of either CF or attempted repair were recorded. Of these 142 instances, 97 (68%) were teacher-initiated, 12 (9%) were peer-initiated, and 33 (23%) were self-initiated (student attempted self-repair without CF from teacher or peers). In a similar pattern to the recast group, these proportions held fairly steady across class days, though it is worth noting that no peer-initiated CF was given in Day 1.

Response. After teacher-initiated prompts (97 instances), students attempted repair a total of 90 times, or at a rate of about 93%. Of those 90 times, students responded with Korean /ɾ/ a total of 9 times (10%). Students repaired to interlanguage 41 times (46%) and to English /ɹ/ 40 times (44%). There were 12 total instances of peer-initiated CF, after which students attempted repair 10 times, or at a rate of about 83%. Of those 10 instances, students repaired to

interlanguage 2 times (20%) and to English /ɹ/ 8 times (80%). Over the course of the class, there were also a total of 33 instances of self-initiated attempts to repair. In 3 of these instances (9%), students produced Korean /ɾ/. In 7 instances (21%), students produced interlanguage, and in 23 instances (70%), students produced English /ɹ/.

It is also interesting to note how learner responses changed over the course of the four-day program (Figure 4). The most noteworthy trend that emerges when the data are analyzed this way is how dramatically instances of correction to Korean /ɾ/ drop off after Day 1. While repairs to /ɾ/ accounted for 25% of the total student responses on Day 1, this number dropped to below 10% for Days 2 and 3, and 0% by Day 4 of the study. Correspondingly, the number of responses which repaired to English /ɹ/ increased after Day 1. While students repaired to English /ɹ/ in 33% of responses on Day 1, this rate was no lower than 43% on any of the following days (for raw numbers, see Table 6). The patterns for repairs to interlanguage represent a trend that is less clear from these data, as the percentage of interlanguage repairs decreased on Days 2 and 3 of the study, but increased to their highest level on Day 4.

Table 6

CF and learner response, descriptive data for prompt group

	Day 1	Day 2	Day 3	Day 4	Total
Korean /ɾ/	6 (26%)	2 (5%)	4 (10%)	0 (0%)	12 (9%)
Interlanguage	9 (39%)	14 (33%)	12 (29%)	15 (56%)	50 (37%)
English /ɹ/	8 (35%)	26 (62%)	25 (61%)	12 (44%)	71 (53%)
Total no. CF	24	45	45	28	142
Total no. attempted repair	23 (96%)	42 (93%)	41 (91%)	27 (96%)	133 (93%)

CF and Responses for Selected Individual Learners

Throughout this study, instances of CF and response were tracked and coded not only at the group level, but also at the level of six individual participants. In this section, CF and responses from three individual students from the prompt group and three from the recast group are analyzed in an attempt to identify patterns in CF or responses among students who made large or minimal gains from the pretest to the posttest. These students were selected based on having shown a range of responses to treatment in their pre- and posttest results, and based on having had at least some room to improve from their pretest. Because this analysis is conducted at the level of the individual learner, rather than over a group of learners, these results should be interpreted with caution.

Recast Group Participant 17

Participant 17 demonstrated large gains in his controlled untrained production (3.6 "possibly /ɹ/" to 1.55 "good /ɹ/"), and slight gains in his controlled trained production (3.2 to 2.9), but was unable to show improvement at the spontaneous level of production (trained 3.6 to 3.3 and untrained 3.1 to 4.15). Moreover, this student demonstrated no gains in perception despite having relatively low pretest scores (61% trained, 55% untrained perception). So while this student was able to make gains in one area of production, he was not able to make large overall gains from the instruction.

This learner had a relatively high number of recorded episodes of CF or attempted repair (40 total vs. an average of only 14 for the recast group). Of these 40 episodes, 27 (68%) were teacher-initiated, 11 (28%) were self-initiated and 2 (5%) were peer-initiated. These averages for teacher/peer/self initiation fall roughly in line with averages for the recast group.

In terms of response, this learner attempted repair after about 78% of teacher-initiated CF, which again falls in line with group averages. Out of the remaining 34 episodes of CF or

attempted repair, this participant was able to repair to /ɹ/ in a total of 17 instances (50%), and to interlanguage in 11 instances (32%), but was unable to repair (Korean /ɾ/) in a total of 6 instances (18%). Thus this student repaired to English /ɹ/ about 10% less frequently than the recast group average. Also of note, this student repaired to Korean /ɾ/ only after teacher-initiated CF, and had a higher rate of repair to English /ɹ/ after self-initiated repair than after teacher- or peer-initiated CF (Figure 2).

As noted above, participant 17 failed to show improvement in either perception or spontaneous production. Interestingly, these trends are mirrored in his responses to CF. When his responses to CF from the first half of the course are compared to his responses from the second half of the course, it can be seen that not only did his rates of repair to interlanguage or English /ɹ/ decrease in the second half of the course, but his rate of attempted repair decreased, and his instances of production of Korean /ɾ/ increased. These results are visually presented in Figure 3.

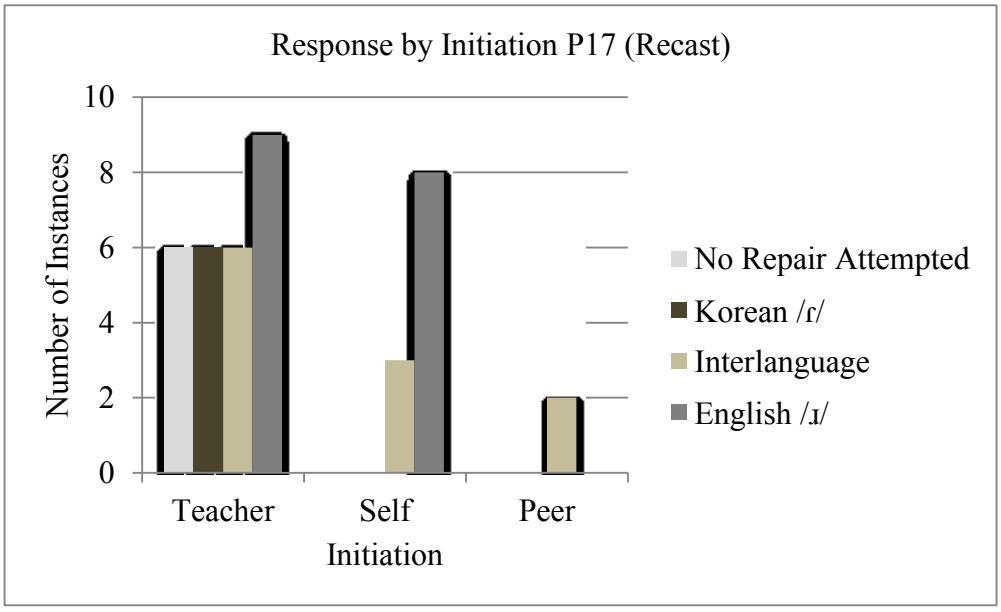


Figure 2. Learner response by initiation type, participant 17, recast group.

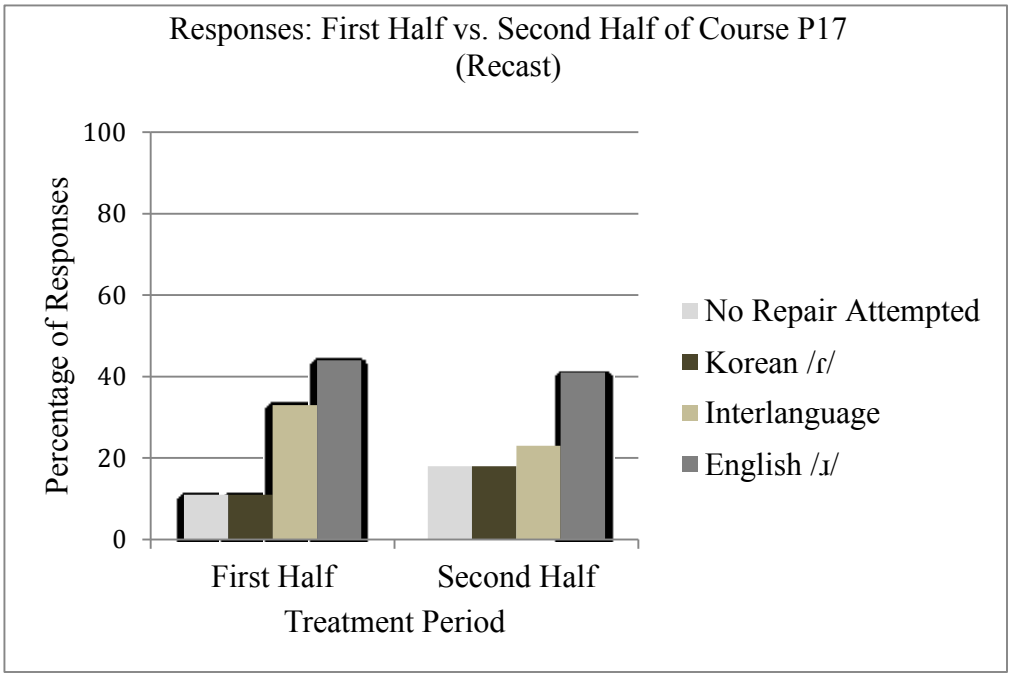


Figure 3. Participant 17's responses over time.

Recast Group Participant 14

Participant 14 demonstrated no gains in production, and even scored worse in spontaneous trained production (2.55 "probably /ɹ/" to 4.2 "possibly /ɹ/"). Participant 14 did, however, make some gains in perception, for both trained items and a familiar speaker (77% to 90%) and untrained items, with an unfamiliar speaker (70% to 95%). From these numbers, it appears that this student benefitted from this study in terms of perceptual training, but not in terms of production.

Participant 14 was involved in a slightly below average number of CF or attempted repair episodes (11 vs. an average of 14 for the recast group). This student attempted repair after 75% of the CF, which is relatively close to the group average of 80%. Of the 9 total instances in which this student attempted repair, he repaired to English /ɹ/ a total of 6 times (67%), and to interlanguage a total of 3 times (33%). Notably, this student never repaired to Korean /ɾ/. Also, this student did not receive any peer-initiated CF in the classtime that was recorded.

Recast Group Participant 16

Participant 16 had high scores for perception at the pretest (97% trained, 90% untrained), and due to ceiling effects was not able to demonstrate any gains in those areas. Likewise, this participant's pretest scores for controlled production were already fairly advanced (trained 1.9, untrained 2.25. Both "good /ɹ/"). However, participant 16's spontaneous production for the pretest had more room for improvement (3.85 trained, 3.8 untrained. Both "possibly /ɹ/"). Participant 16 demonstrated strong gains in both trained and untrained spontaneous production (posttest 1.85 trained and 2.25 untrained; both "good /ɹ/"). Although participant 16 had strong initial scores for controlled production, he also made some small gains there (trained 1.9 to 1.6, untrained 2.25 to 1.6). Overall, this student can be said to have made noteworthy gains in production, and to have had an already advanced level of perception for /ɹ/.

Participant 16 had an above average number of episodes of CF or attempted repair (24 vs. average of 14). The student attempted repair after 82% of the teacher-initiated CF, which once again falls about in line with the group average. This student received one episode of peer-initiated feedback, after which he did not attempt repair. In those instances in which this student did attempt repair, he repaired to English /ɹ/ 13 times (65%), and to interlanguage 7 times (35%).

One interesting trend in the data for this student is that even though he made considerable gains in production, his rate of repair to English /ɹ/ actually decreased from the first half to the second half of the study. In contrast, participant 16's instances of interlanguage response actually increased from the first to the second half of the course. Participant 16's rate of attempted repair held fairly steady over the course of the study (Figure 4).

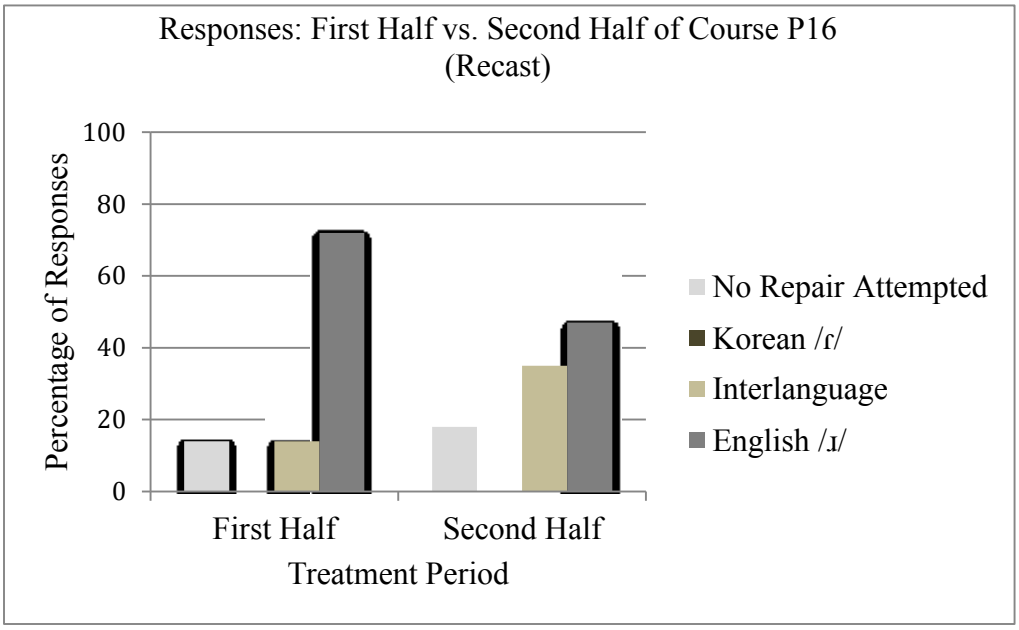


Figure 4. Participant 16's responses over time.

Prompt Group Participant 22

Participant 22 demonstrated gains in all areas of perception and production. In perception, this participant improved his score by at least 10 percentage points in all areas (familiar speaker/trained from 60% to 73%; familiar speaker/untrained from 60% to 75%; unfamiliar speaker/trained from 53% to 70%; unfamiliar speaker/untrained from 60% to 70%). Participant 22 demonstrated large gains in all facets of production as well (controlled/trained from 4.65 "neutral" to 2.35 "good /ɹ/"; controlled/untrained from 5.70 "possibly /l/" to 3.00 "probably /ɹ/"; spontaneous trained from 5.55 "possibly /l/" to 3.75 "possibly /ɹ/"; spontaneous untrained from 5.60 "possibly /l/" to 2.45 "good /ɹ/"). Of all of the participants in the study, this student showed the largest individual gains.

Participant 22 was involved in substantially more episodes of CF or attempted repair than the average for the prompt group (46 vs. average of 24). In total, this participant received teacher-initiated CF 34 times, and attempted repair in 32 of those instances, or at a rate of 94%, which is in line with the group average for prompts. In 15 of those instances (47%) the student

repaired to English /ɹ/, in 13 instances (41%) the student repaired to interlanguage, and in 4 instances (13%) he produced Korean /ɾ/. Participant 22 also attempted self-repair 9 times, after which he repaired to English /ɹ/ 5 times (56%), to interlanguage 3 times (33%), and after which he produced Korean /ɾ/ one time (11%). He also received 3 instances of peer-initiated CF; in one case he did not attempt repair, and in the other two instances he repaired to interlanguage. When participant 22's responses are viewed over time, it becomes clear that his patterns of response changed significantly after Day 1 of the study (Figure 5). On Day 1, 4 out of 9 (44%) of his repair moves were Korean /ɾ/, but that number dropped to 0 out of 12 for Day 2, and his rate of production of Korean /ɾ/ did not exceed 7% of his repair moves on any day thereafter. By Day 4 of the study, 100% (8/8) of participant 22's attempts at repair resulted in either interlanguage or English /ɹ/, but his most dramatic change in response patterns occurred between Days 1 and 2.

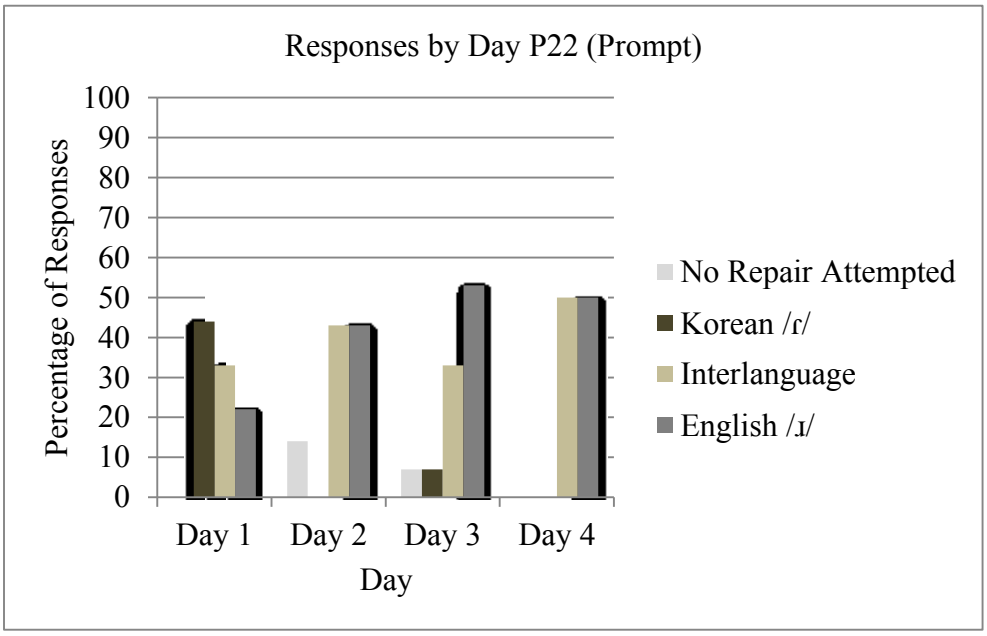


Figure 5. Learner response by day, participant 22, prompt group.

Prompt Group Participant 24

Participant 24 demonstrated gains at all levels of production (controlled, trained 3.4 to 2.55; controlled untrained from 2.85 to 2.20; spontaneous trained from 3.55 to 2.80; spontaneous

untrained from 5.35 to 2.80) but these gains were especially evident at the spontaneous untrained level. At the pretest, only 20% of this participant's spontaneous untrained tokens were successfully identified by NS judges as examples of /ɹ/ (a rating of 4, 3, 2 or 1), whereas on the posttest 75% of those tokens were successfully identified. This participant's results on the perception tests were more ambiguous. Whereas participant 24 demonstrated no gains in perception for the familiar speaker, she was able to show gains on both trained (67% to 80%) and untrained (65% to 80%) tokens for the unfamiliar speaker. Overall, this participant can be said to have shown strong improvement in her production of /ɹ/, with more ambiguous results in terms of her perception of /ɹ/.

Participant 24 was involved in a roughly average number of episodes of CF or attempted repair (26 vs. prompt group average of 24). Thirteen of these episodes were teacher-initiated and 2 were peer initiated. Participant 24 attempted repair after every one of those instances of CF, or at a rate of 100%. In 9 instances (38% of total), this participant attempted self-repair. This is at a higher rate than the group average, despite participant 24 not having attempted self-initiated repair at all in Day 1. The participant repaired to English /ɹ/ a total of 19 times, or at a rate of 73%, substantially higher than the average for the prompt group. She repaired to interlanguage 4 times (15%) and produced Korean /ɾ/ after attempted repair 3 times (12%). Participant 24 did not show large changes in her patterns of attempted repair over the course of the study.

Prompt Group Participant 26

While participant 26 was able to make some gains in controlled production (trained 3.1 to 2.45; untrained 4.05 to 3.25), her results in spontaneous production were less convincing. She made slight improvement on trained tokens, but was unable to generalize to untrained tokens (trained 4.2 to 3.7; untrained 4.00 to 4.05). For perception, this participant was able to make large gains on both trained and untrained items for a familiar speaker (57% to 77% trained; 50%

to 80% untrained), but had difficulty generalizing her gains to an unfamiliar speaker (for trained tokens 60% to 77%, but for untrained tokens 75% to 65%). Thus, this participant can be said to have greatly improved her perception of /ɹ/ for a familiar speaker, but to have had difficulty generalizing those gains. This participant did garner some benefits from this study in her controlled production of /ɹ/, but was unable to greatly improve her spontaneous production of /ɹ/.

Participant 26 was involved in a slightly below average number of episodes of CF or attempted repair (20 vs. group average of 24). She had a rate of attempted repair after teacher CF of 100%, but of only 67% after peer feedback (that is, 2 out of a total of 3 instances). Out of the instances in which she attempted repair (19 total), participant 26 repaired to English /ɹ/ 11 times (58%) and to interlanguage 8 times (42%). Noteworthy is that this student had a much higher rate of repair to interlanguage after teacher-initiated CF and a higher rate of repair to English /ɹ/ after self-initiated repair (see Figure 6).

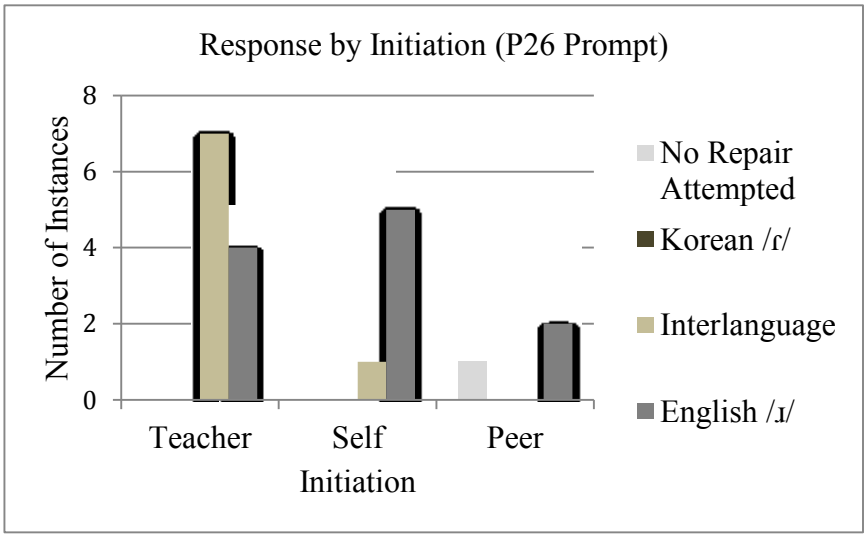


Figure 6. Learner response by initiation type, participant 26, prompt group.

There are a number of interesting trends in this student's response patterns over time. First, it can be seen that between the first half and second half of the course, participant 26 began to produce a higher ratio of interlanguage responses relative to English /ɹ/ responses; in the first

half of the course, participant 26 produced interlanguage responses at a rate of 1:2 English /ɹ/ responses. In the second half of the course, this changed to a 1:1 ration (Figure 7). It is also noteworthy that this participant never produced Korean /ɾ/ after attempted repair.

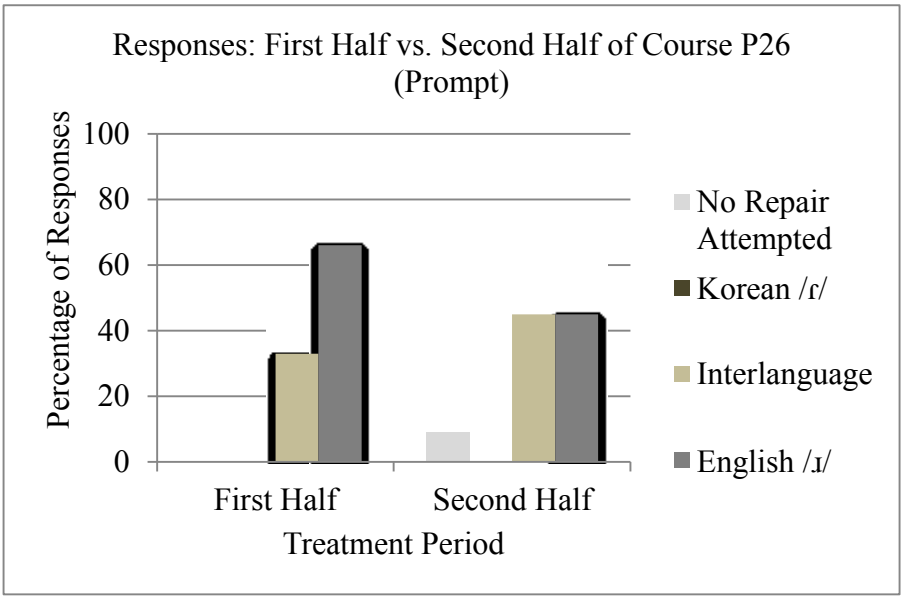


Figure 7. Participant 26's responses over time.

Summary of Individual Response Patterns

Analysis of individual responses revealed that students changed their response patterns over the course of the study, often in ways that did not follow the overall trends for their group. Moreover, these patterns reveal that students may have taken more than one route to improved production. The two students from the prompt group who made the largest gains in production were P22 and P24. Each student seems to have made their gains in a slightly different fashion, but a couple of factors also held true for both. For example, both students increased their rate of repair to /ɹ/ by Day 4 of the study, but whereas P22 also increased his rate of repair to interlanguage, P24 actually decreased her rate of repair to interlanguage. Likely, this difference was due to the different initial levels of expertise between these learners. P22 was a beginner-level learner and so producing either interlanguage or English /ɹ/ was an improvement over

production of Korean /ɾ/ (which he produced in 44% of instances on Day 1). P24, on the other hand, was already producing /ɹ/ at an intermediate level, and so was able to transition from interlanguage to targetlike utterances.

In the case of the less successful student in the prompt group, P26, an opposite pattern to that of P24 was evident. From the first to the second half of the class, P26 actually decreased her rate of repair to /ɹ/ relative to her rate of repair to interlanguage. It is important to note, however, that this does not necessarily mean that P26 did not benefit from the CF. In fact, this student benefitted in both controlled production and perception of /ɹ/. One possibility is that in the case of this individual student, the lack of positive evidence in prompts may have been somewhat confusing. This supposition is supported by her patterns of repair following prompts vs. her patterns of self-repair. After prompts, P26 repaired to interlanguage at a rate of nearly 2:1 to targetlike /ɹ/. In contrast, when P26 self-repaired, this ratio changed to 1:5: a substantial swing.

Interestingly, P26's response patterns were paralleled by P16, who was the most successful learner in the recast group. Like P26, P16 also decreased his rate of repair to /ɹ/ in favor of an increased frequency of repair to interlanguage. Counter-intuitively, this student showed large gains in production. One possible explanation for these gains is that the student may have merely been mimicking the teacher's pronunciation at the beginning of the course without consciously attending to his own pronunciation, but later may have paid more conscious attention to his production of /ɹ/, resulting in temporarily untargetlike interlanguage responses as he began to break down his old habits. While tentative, this conclusion is supported by the student's exit interview (see following subsection), in which he reported watching, and attempting to copy the teacher's mouth shape during provision of recasts, suggesting that he began to consciously attend to his articulation of /ɹ/.

The less successful participants from the recast group, P17 and P14, showed different patterns of feedback from P16. P17, for example, decreased his rate of repair to both /ɹ/ and interlanguage, and actually increased his rate of utterance of Korean /ɾ/. In the case of P14, the participant was involved in so few episodes of CF and response that it is difficult to gauge how those patterns may have changed over time. It may be that P14 was simply involved in too few episodes of CF or attempted repair to make a large difference in his production. P14 did, however, still manage to make gains in his perception of /ɹ/, though it is unclear whether those gains were due to CF or to another element of the FFI treatment.

Exit Interviews

After the completion of posttests, participants completed short exit interviews. These interviews revealed some interesting supplementary information, including participants' strategies for producing /ɹ/ and differentiating it from /l/, and their views on CF. In these interviews, all students were asked their views on CF. An interesting difference in response could be seen between the experimental groups (recast and prompt) and the no CF group. In the experimental groups, 100% of students said that they thought the CF they received in the class was helpful. This can be contrasted with the responses of the no CF group, in which only 6 students (out of 9) said they thought teachers should provide CF in class some or all of the time, and 3 students said that teachers should not provide CF in class.

These interviews also provided some insight into student strategies for improving their pronunciation of /ɹ/. Students from all groups reported using the techniques described in the EI portion of this study to help improve their pronunciation of /ɹ/. In the prompt group, 100% of participants reported consciously thinking about at least one element of the articulatory EI (either tongue position or lip rounding) when pronouncing /ɹ/ in the class. Five out of the seven recast group participants also reported using this strategy, as did five out of nine participants in the no

CF group. It is likely that students in the CF groups thought about EI more during their pronunciation of /ɹ/ because their attention was continually focused on that sound through the CF. Students in the prompt group may have relied even more heavily on EI than students in the recast group because they could not rely on positive evidence. Besides thinking about elements of EI when attempting to improve their pronunciation of /ɹ/, some students in the recast group reported using an additional strategy. Two out of seven students in the recast group said that during recasts, they watched the teacher's lips, and tried to mimic what he did as they attempted repair. Because the recast group was the only group for which CF with positive evidence was provided, no students in the prompt or no CF groups were able to adopt this strategy.

Chapter 5

Discussion

This chapter discusses implications from the results of this study. It begins by looking at the production data from this study, with a specific focus on the differences and similarities which were observed between the recast and prompt groups. Following this is a discussion and analysis of the perception data, including some thoughts on the factors that contributed to the ambiguity of those data.

Production

Overall, this study demonstrated that phonological FFI with EI can be effective for Korean learners in an EFL context. All groups in this study significantly improved their performance of /ɹ/ between pretest and posttest. It should be noted at this time, that this study included no true control group (but compare to true control group in precursor study). The no CF group in this study was a control group in the sense that it allowed for the isolation of the CF variable for analysis, but this group nonetheless did receive FFI and EI, and improved /ɹ/ performance as a result. These results support the hypothesis that the treatment used in the precursor study could be effectively transferred to a novel context.

Recasts and Prompts

The results of inferential statistical measures hinted at a continuum of effectiveness among the different CF conditions. To be precise, the no CF group did not show significant improvement at any particular level of production, the recast group showed significant improvement at the controlled/untrained level of production, and the prompt group showed significant improvement at the controlled/untrained and the spontaneous/untrained levels of production. These results should be interpreted with caution, but they are very interesting in light of the fact that prompts, which had not been investigated as a form of phonological CF up to this

point, seemed to outperform recasts. It will be useful to keep these results in mind as we take a closer look at the information gleaned from the in-depth analysis of CF and learner response in this study.

Trends in the recast data. The in-depth examination of recasts and learner responses undertaken in this study has revealed a number of interesting trends. Firstly, this study provides additional evidence that phonological recasts can be salient for English L2 learners. The isolated declarative recasts used in this study led learners to attempt repair in 80% of instances. Furthermore, the rate of learner repair attempted after recasts increased over the course of the four-day study, suggesting that as students became habituated to these phonological recasts, the salience of the recasts was further increased. The high rates of learner response and attempted repair for this kind of phonological recast are consistent with the predictions of Sheen (2006), and the results of descriptive studies such as Lyster (1998) that have also documented the saliency of phonological recasts. Nevertheless, even the relatively high rate of attempted repair after recasts recorded in this study was lower than the rate of repair after prompts.

The recast group's fairly advanced pretest level of production of /ɹ/ was apparent in the CF data. Even on the first day of the study, recast group participants corrected to /ɹ/ in over 50% of instances and to /ɾ/ in less than 5% of all instances. Despite an already high rate of repair to /ɹ/ (54%) on Day 1 of the study, the recast group actually increased its rate of repair to 70% by Day 4. Much of this increase, however, can be attributed to the fact that learners were attempting repair at a greater rate in general. If repair to /ɹ/ is instead calculated as a percent of the total instances in which learners attempted repair, the increase is smaller, from 70% on Day 1 to 78% on Day 4. This increase in the rate of repair to /ɹ/, while interesting in that it represents a change in learner response patterns, should not be taken as an indication that learning took place, as the relationship between repair and L2 learning remains unclear (Ellis, Loewen, & Erlam, 2006). For

example, in some instances students may have simply echoed the teacher's pronunciation without processing it beyond the level of short-term memory. In other instances, students may have been attending more closely to their own production processes; indeed, during posttest interviews, two students from the recast group reported watching the teacher's lip movement during provision of recasts in order to better self-correct. Whatever the exact processes involved, the recast group was able to not only increase its rate of repair to /ɹ/, but to demonstrate gains in controlled production despite its advanced levels of production at the start of the study.

Trends in the prompt data. There are a number of interesting trends that emerged from the analysis of the CF and learner responses of the prompt group, especially when compared with the patterns found in the recast group. Like the recast group, the prompt group had very high rates of repair following CF. Over the course of the study, students attempted repair after 93% of all prompts. This high rate of repair held true over the course of the four-day study, suggesting that prompts were highly salient as a form of phonological CF even before students had received much FFI.

On Day 1 of the study, students in the prompt group produced Korean /ɾ/ in 25% of all instances of CF or attempted repair. By Day 4 of the study, this rate had dropped to 0%. Based on these numbers alone it is possible to see that students in the prompt group changed their habits of repair over the course of the study. Unfortunately, it is not possible to say whether this change was caused by the CF received by the group, or by some other element of the treatment. This is because the recast group had very low rates of production of /ɾ/ on Day 1; thus the changes in production of /ɾ/ cannot be compared. Nonetheless, this observation indicates that prompts may be an effective way to help L2 learners move away from L1 pronunciation strategies, especially considering the prompt group's strong performance on production posttests. It would certainly be worth investigating in future studies with larger participant pools whether

prompts can remain effective across a variety of different phonological features, and whether there is indeed a difference between prompts and recasts in this respect. This is especially true when one considers that most of the change from production of /ɾ/ to interlanguage or /ɹ/ occurred after just one hour of treatment (25% /ɾ/ on Day 1 to 5% on Day 2). It cannot be said from the data collected in this study whether students in the prompt group had improved their intelligibility or not after one day of class, but it is certainly worth noting that it was at this point that their patterns of response changed most dramatically. After all, a treatment that requires less class time to take effect leaves more class time for other linguistic targets and educational objectives.

One further item of note is that the prompt group produced significantly more interlanguage responses after CF than the recast group (42% of responses to teacher-initiated CF vs. 22%). While part of this difference may be attributed to slightly better initial production scores for the recast group, another explanation is also possible. Prompts do not provide positive evidence, whereas recasts do. Because of this difference, students receiving a recast have the option of mimicking the instructor's pronunciation but students receiving prompts are pushed to use their own resources to try to make a targetlike utterance, which may result in the use of more interlanguage strategies. This study demonstrated slightly stronger results for students who were pushed in this way through prompts, suggesting that students improved the intelligibility of their pronunciation of /ɹ/ by improving their own interlanguage strategies.

Recasts or prompts? The initial findings presented in this study indicate that, as in other realms of SLA, prompts may be effective at pushing learners to increase the accuracy of their output (Lyster, 2004). Because students in the prompt group, but not the recast group were able to transfer their gains to spontaneous production, it is the tentative conclusion of this thesis that prompts were more effective at pushing learners to improve their production of /ɹ/ than recasts.

This conclusion comes with a number of caveats, and what is perhaps more interesting is not whether prompts or recasts are better for phonological CF, but rather in what ways they were effective, and how they might be best applied in the future in classroom situations.

One possible reason the prompt group was better able to transfer its gains to spontaneous production is that, because prompts do not provide positive evidence, students were pushed to draw on their own resources to self-correct. When responding to a recast, in contrast, it is possible for a student to simply parrot the teacher's recast using only short-term memory, without drawing on the student's own resources (Lyster, 2004). When a student echoes in this fashion, rather than drawing on his/her own resources to respond, it may be less effective for proceduralizing knowledge of the production of /ɹ/, which could explain how students in the recast group were able to show improvement at controlled tasks, in which they were able to concentrate on their pronunciation, but not in spontaneous tasks, which are more demanding on linguistic processing (Saito & Munro, 2014).

As mentioned previously, the results of this study suggest that learners may adjust their interlanguage strategies in response to prompts in order to make their production of /ɹ/ more intelligible. It remains to be seen what the limits of this process are. For example, learners in this study could draw on information given in the EI portion of the study as a basis for trying to improve their interlanguage strategies. If students did not have the resource of EI about /ɹ/, results from the prompt group may well have been very different. This supposition is supported by posttest interviews conducted with participants in this study, wherein all six of the participants in the prompt group specifically mentioned attending to the manners of articulation described in the EI treatment when they attempted to make corrections to their production of /ɹ/. Without the structure provided by the EI, students may have had more difficulty increasing the accuracy of their production of /ɹ/.

Despite the prompt group's strong improvements in production, one potential weakness of prompts as a type of phonological feedback was also made apparent. After Day 1 of the study, the prompt group saw a dramatic decrease in its production of Korean /ɾ/ (25% to 4%) and a concurrent rise in its rate of repair to /ɹ/ (33% to 58%), but these levels plateaued after Day 2. This raises a question as to whether provision of positive evidence in CF might be needed at some point to help students reach higher levels of attainment. This argument is made by Saito (2014b), who sees recasts as an effective way to draw learners' attention away from interlanguage strategies. Anecdotally, it might have been useful to provide positive evidence to students in the prompt group when it was clear that they were unsure how to modify their production of a difficult word after prompting. The following excerpt is from a prompt class on Day 3.

P21: /ɹuɹaɹ/ (rural)

Instructor: One more time.

P21: /ɹuɹaɹ/? Is it right? No?

In this case, the student has overgeneralized English /ɹ/ to the /l/ sound at the end of the word "rural," but even after prompting, is clearly unsure how to repair her utterance. She is looking to the teacher for help. In instances such as this, the opportunity provided by the prompt for self-repair may not be realized because the student is unsure how to make that repair (Loewen & Philp, 2006). If this is the case, it might be beneficial to either provide the student with a model for the pronunciation of this word (that is, to provide a recast) or perhaps to remind student of relevant metalinguistic information so that the student is able to repair successfully. Loewen (2002) proposes using a recast in this type of situation, in which the learner is unsuccessful in self-generating the target form, as a way to resolve the knowledge gap. A recast would have the additional benefit of resolving the situation quickly, in the case that the student is becoming

frustrated or embarrassed from his/her inability to repair. In some cases, however, if the instructor can recognize the nature of the student's error, it could be possible to provide metalinguistic information instead of positive evidence. Examples include articulatory information (e.g., round your lips), or in this case something like "you are pronouncing the final 'l' as an 'r.'" This approach has the benefit of continuing to push the student to self-correct, while providing additional guidance to help the student succeed in that correction.

In short, this study found prompts to be more effective than recasts in helping learners improve production of /ɹ/, but has by no means settled the matter. Prompts need to be examined further to determine if they are effective at helping learners move beyond interlanguage strategies. Additionally, the effects of other types of CF on phonological learning should be examined. Although it would be difficult to measure experimentally, a mixture of CF types may ultimately prove to be the most effective for classroom application. Based on the observations of this study, a promising combination would be using prompts initially to push learners to draw on their own resources to self-correct, and supplementing those prompts with other types of CF such as recasts in instances in which students are unable to self-correct, or need additional help to move beyond interlanguage strategies to targetlike production.

Perception

Perception test results for this study reveal that participants were already able to distinguish /ɹ/ at a high rate (>80%) at the pretest. Based on this finding, the participants in this study can be considered to be advanced learners in terms of their perception of /ɹ/, and can also be considered perceptually ready to benefit from feedback on their production. High pretest scores mean that, because of ceiling effects, it is difficult to determine what effects the different elements of FFI examined in this study may have had on students' perceptual development.

Nevertheless, the inclusion of perceptually advanced learners in the study allows for a number of other interesting implications to be drawn.

The most exciting of these implications is that phonological FFI can remain effective even for advanced learners. Previous studies by Saito (e.g., 2013a) investigated learners with much more initial difficulty with /ɹ/. His studies successfully applied FFI including CF to help these learners develop their production and perception of /ɹ/. While lower level learners have more room to improve, and thus the possibility to make more dramatic gains in production, this study demonstrated that even advanced learners can improve their production of /ɹ/, perhaps because they are perceptually ready to benefit from the CF they receive. This is consistent with studies such as Sheldon and Strange (1982), and Sheldon (1985), which found that higher-level Japanese and Korean learners of English (but not less advanced students) could demonstrate more advanced levels of production than perception for /ɹ/. Thus, while students should be able to perceive the difference between L2 sounds in order to best benefit from production practice, it may be the case that once learners have reached a certain level of perceptual ability it becomes easier to further fine-tune their production than their perception (Sheldon, 1985). This conclusion is encouraging considering the lower level of ultimate attainment adult L2 learners typically face in their L2 production compared to young learners. This study has demonstrated the possibility for even advanced adult learners to improve their production of an L2 target sound with the help of FFI and CF. In sum, this successful replication of Saito (2013a) affirms the efficacy of FFI as a viable option for pronunciation teaching.

Chapter 6

Conclusion

Although there is still much work to be done in establishing best practices for pronunciation instruction, this study has taken a step towards establishing what those best practices might be. This study built on previous investigations into the efficacy of FFI for pronunciation teaching, and provided evidence that those practices can be effective in a new teaching context. Whereas previous studies focused on beginner-level students, this study demonstrated the feasibility of FFI with CF for improving L2 production even among advanced learners. Additionally, this study points towards new possibilities in pronunciation-centered FFI, which could potentially increase its effectiveness. Specifically, this study has provided evidence that prompts, when used as phonological CF within FFI, show promise for encouraging phonological development by pushing learners to test new linguistic hypotheses through modifying their output. This study also uncovered a potential drawback of prompts; they may help learners to improve interlanguage strategies, but not necessarily to move beyond interlanguage strategies. At this point, the learners may require additional help, possibly in the form of CF with positive evidence, in order to improve further. By demonstrating the promise of prompts as a form of phonological CF, this study has also raised the possibility that other, as yet untested forms of CF might also be shown to be effective for pronunciation instruction.

Besides specific findings, this study has also contributed a new method for analyzing classroom data that could prove useful in future studies. By tracking the CF and responses of individual learners in detail and over time, this study was able to reveal interesting trends that would otherwise have been invisible. This method of tracking the results of individual learners across time in a quantitative manner allows the researcher to see at what point during a treatment participants' behavior changes, and could thus be useful in future studies investigating the effects

of time of instruction. This method for tracking individual student's response patterns, if combined with testing throughout a study, might also provide insight into the relationship between uptake, repair and learning, because it allows the researcher to examine some of the ways in which patterns of uptake and repair relate to learning.

Limitations

One limitation of this study was that it did not include a control group that received neither FFI nor CF. Although this limitation is mitigated by the existence of a true control group in the precursor study, that group cannot be directly compared to the groups in this study because of differences in L1 (Japanese vs. Korean) and context. Nonetheless, it is very unlikely that students would have made significant gains in their production or perception of /ɹ/ due to factors other than the treatment in the short time in which this study took place (less than one month between initial pretest and end of study), especially considering that it took place in an EFL setting.

One other limitation more serious than the lack of a true control group was the small size of groups in this study. As was mentioned in the results portion of this study, statistical results were less robust than they would have been due to the small number of students who were recruited for this study. Additionally, many students had very high pretest scores for perception of /ɹ/, which made it difficult to judge whether or not participants progressed in their perceptual ability. These shortcomings could be overcome in future studies through the recruitment of more participants and through either a more careful screening of participants to ensure that their level was appropriate for the study, or through the use of a more nuanced perceptual test able to detect subtler differences in perceptual ability, perhaps through the inclusion of the target sound in linguistic environments that have been shown to be more challenging than initial /ɹ/, such as codas (Cardoso, 2011) or consonant clusters (Lee, Joh, & Cho, 2002).

Finally, a couple of limitations in terms of measurements of CF responses and production should be mentioned. As noted in the results section, some of the video data for this study became corrupted and was thus unable to be analyzed. Future studies that use a similar method for tracking feedback and learner response may reveal additional patterns that were not revealed in this study due to the missing data. In terms of production measurement, it was important in terms of ecological validity to include a measure of spontaneous production. This type of production is, however, more difficult to measure than controlled production. Because students were speaking freely, their target tokens were influenced by surrounding words. This kind of environmental influence is natural in spontaneous speech, but may have affected NS judges' perceptions of the /ɹ/ sound when that single word was extracted from the context of the participant's utterance. This effect was unavoidable because to include surrounding words would have provided context clues to the judges and influenced their judgment as well. Perhaps to control for the effect of extracting the token from spontaneous speech, future studies could include a NS baseline. Doing so would allow the researcher to see to what degree (if any) NS spontaneous production scores also suffer compared to controlled production tokens, and thus adjust scores for non-native speakers to more accurately reflect production of /ɹ/ without this confounding factor.

Future Studies

In the limitations section above, several suggestions were made for improving the techniques used in this study. Beyond these relatively minor changes, however, this study also suggests a large number of courses of future inquiry that would be interesting and informative for the as yet developing field of pronunciation instruction. Perhaps the most urgent need suggested by the results of this study is for researchers in pronunciation instruction to investigate further the role of different types of CF. The promising results shown by the prompt group in this

study especially suggest that the best kind of CF for pronunciation instruction has yet to be determined. Because of its provision of both positive and negative evidence, explicit correction would be a particularly interesting type of CF to investigate next.

Not only should the effect of different types of CF on learner acquisition of phonological targets be investigated, but how those different types of CF can be best used, and in what combinations also needs to be examined. Recasts, for example, might still be improved upon. One of the potential drawbacks of recasts discussed in this study is that students may merely echo the teacher's pronunciation without accessing their own long term memory; some students ,however, reported watching the instructor's mouth during recasts in order to copy his lip movement. If the instructor encourages this kind of focus (e.g., pointing to lips during CF provision), could this shortcoming of recasts be overcome? Merely improving recasts or prompts, however, may not ultimately uncover the best way to provide phonological CF. Indeed, there is no reason to assume that a single type of CF would be inherently superior to all other types for pronunciation instruction. For example, Lyster (2004) suggests that recasts are helpful in the encoding of declarative knowledge and that prompts are helpful in the transformation of declarative knowledge to procedural knowledge. Could the phases of acquisition of a phonological target parallel this pattern? If so, a combination of multiple CF types might ultimately prove to be the most beneficial. Another interesting direction of study would be to investigate the role of perception-focused CF vs. production-focused CF. Theories of phonological development reviewed in this thesis suggest a "perception first" approach; might training which focuses initially on perception-based CF (as in Lee, 2013) and later moves towards production practice and CF be more effective than training that pushes learners to produce before they are perceptually ready?

The questions posed above beg another question "when are learners ready for a given treatment?" In order to address this question, more research should be conducted to determine the appropriate length of instruction for a given phonological target. In a recent meta-analysis of pronunciation intervention studies, Lee et al. (2014) found that length of instruction was generally positively correlated with strength of results. But in terms of applicability to a real-world classroom, it is essential to find treatments that can be effective in shorter time periods. Research such as the current study that tracks CF in detail and across time may reveal patterns that help determine at what point in the classroom intervention changes are taking place. Not only that, but by tracking individual CF and responses, future studies may help to finally reveal the relationship between uptake, repair, and attainment of a given linguistic target—not only in pronunciation instruction, but across many other subfields of SLA.

Lastly, future studies should continue to investigate a variety of phonological targets. While, for purposes of comparability, the current study followed the lead of its precursor to investigate English /ɹ/, other studies need to continue to investigate a diverse array of consonants and vowels, segmental and suprasegmental targets, not only for ESL/EFL, but for a variety of target languages. Only when techniques for pronunciation instruction are investigated in a variety of environments can we begin to confirm the elements of instruction that are beneficial in a given situation.

References

Ammar, A., & Spada, N. (2006). One size fits all?: Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, *28*(4), 543-574.

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

Boersma, P., Escudero, P., & Hayes, R. (2003). Learning abstract phonological from auditory phonetic categories: An integrated model for the acquisition of language-specific sound categories. In Sole, M.J., D. Recansens and J. Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 1013-1016). Barcelona.

Borden, G. (1983). Production and perception of the /r/ - /l/ contrast in Korean adults learning English (pp. 499-526).

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

Cardoso, W. (2011). The development of coda perception in second language phonology: A variationist perspective. *Second Language Research, 27*(4), 433-465. doi: 10.1177/0267658311413540

Carpenter, H., Jeon, K., MacGregor, D., & Mackey, A. (2006). Learners' interpretations of recasts. *Studies in Second Language Acquisition, 28*(2), 209-236.

Cho, J., & Park, H. (2006). A comparative analysis of Korean-English phonological structures and processes for pronunciation pedagogy in interpretation training. *Meta: Translators' Journal, 51*(2), 229-246.

Chomsky, N. (1965). *Aspects of the theory of syntax* (Vol. 11): The MIT press.

Day, E., & Shapson, M. (2001). Integrating formal and functional approaches to language teaching in French immersion: An experimental study. *Language Learning, 51*(1), 47-80.

Derwing, T., & Munro, M. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39*, 379-397.

Ellis, R. (2006). Researching the effects of form-focused instruction on L2 acquisition. *AILA Review, 19,* 18-41.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition, 28*(2), 339-368.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition, 26*(4), 551-585.

Flege, J. (1995). Second language speech learning. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium, MD: York Press.

Flege, J. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer & N. Schiller (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 319-355). Berlin: Mouton de Gruyter.

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

French, L. (2006). *Phonological working memory and second language acquisition: A developmental study of francophone children learning English in Quebec*. Lewiston, New York: The Edwin Mellen Press.

Gluszek, A. (2010). Speaking with a nonnative accent: Perceptions of bias, communication difficulties, and belonging in the United States. *Journal of Language and Social Psychology, 29*(2), 224-234.

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

Han, Z., Park, E., & Combs, C. (2008). Textual enhancement of input: Issues and possibility. *Applied Linguistics, 29,* 597–618.

Hardison, D. (2010). Visual and auditory input in second-language speech processing. *Language Teaching, 43*, 84-95.

Hardison, D. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics, 24*(4), 495-522.

Harley, B., Cummins, J., Swain, M., & Allen, P. (1990). The nature of language proficiency. In C. Doughty & J. Williams (Eds.), *The development of second language proficiency* (pp. 7-25).

Lee, A. (2013). *Effects of Corrective Feedback on L2 Speech Perception: Perceptual decisions, linguistic hypotheses, and negative evidence* (Master's thesis). McGill University, Montreal.

Lee, J., Jang, J., & Plonsky, L. (2014). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*. Advanced online publication. *doi:10.1093/applin/amu040*

Lee, S., Joh, J., & Cho, M. (2002). Acquisition of English consonant clusters among Korean EFL learners. *언어,* [language] *27*(3), 439-472.

Lively, S., Logan, J., & Pisoni, D. (1993). Training Japanese listeners to identify English /r/ and /l/: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America, 94*(3), 1242-1255.

Loewen, S., & Philp, J. (2006). Recasts in the adult English L2 classroom: Characteristics, explicitness, and effectiveness. *The Modern Language Journal, 90*(4), 536-556.

Loewen, S. (2002). *The occurrence and effectiveness of incidental focus on form in meaning-focused ESL lessons* (Doctoral dissertation). Retrieved from ResearchSpace@Auckland.

Lyster, R. (1998). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, *48*(2), 183-218.

Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, *26*(3), 399-432.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in Second Language Acquisition, 19*, 37-66.

Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA. *Studies in Second Language Acquisition, 32*(2), 265-302.

Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, *46*(1), 1-40.

Ranta, L., & Lyster, R. (2007). A cognitive approach to improving immersion students' oral language abilities: The awareness-practice-feedback sequence. *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*, 141-160.

Rau, D., Chang, A., & Tarone, E. (2009). Think or sink: Chinese learners' acquisition of the voiceless interdental fricative. *Language Learning, 59*, 581–621.

Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly: A Journal for Teachers of English to Speakers of Other Languages and of Standard English as a Second Dialect, 46*(4), 842-854.

Saito, K. (2013a). The acquisitional value of recasts in instructed second language speech learning: Teaching the perception and production of English /ɹ/ to adult Japanese learners. *Language Learning, 63*(3), 499-529. doi: 10.1111/lang.12015

Saito, K. (2013b). Reexamining effects of form-focused instruction on L2 pronunciation development. *Studies in Second Language Acquisition, 35*(1), 1-29. doi:10.1017/S0272263112000666

Saito, K. (2014a). Communicative focus on second language phonetic form: Teaching Japanese learners to perceive and produce English /ɹ/ without explicit instruction. *Applied Psycholinguistics*, 1-33. doi: 10.1017/S0142716413000271

Saito, K. (2014b). Variables affecting the effects of recasts on L2 pronunciation development. *Language Teaching Research*. Advanced online publication. doi: 10.1177/1362168814541753.

Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. *Language Learning, 62*(2), 595-633. doi: 10.1111/j.1467-9922.2011.00639.x

Saito, K., & Munro, M. (2014). The early phase of /ɹ/ production development in adult Japanese learners of English. *Language and Speech*, *57*(4), 451-469.

Sakow, M. (1990). *Perception of /r/ and /l/ by Japanese and Korean adults learning English as a second language*.

Sato, M., & Lyster, R. (2012). Peer interaction and corrective feedback for accuracy and fluency development. *Studies in Second Language Acquisition, 34*, 591-626.

Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research, 10*, 361-392.

Sheldon, A. (1985). The relationship between production and perception of the /r/–/l/ contrast in

Korean adults learning English: A reply to Borden, Gerber, and Milsark. *Language Learning, 35*, 107–113.

Sheldon, A. & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics, 3*, 243–261.

Spada, N. (1997). Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching, 29*, 73-87.

Tang, X., Zhang, S., Li, Y., & Zhao, M. (2013). Study on correlation of English pronunciation self-concept to English learning. *English Language Teaching, 6*(4), 74-79.

Tench, P. (2003). Non-native speakers misperceptions of English vowels and consonants: Evidence from Korean adults in the UK. *International Review of Applied Linguistics in Language Teaching, 41*(2), 145-173.

Thomson, R. & Derwing, T. (2014). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics.* Advanced online publication. doi: 10.1093/applin/amu076

VanPatten, B. (2004). *Input processing and grammar instruction in second language acquisition.* Norwood, NJ: Ablex.

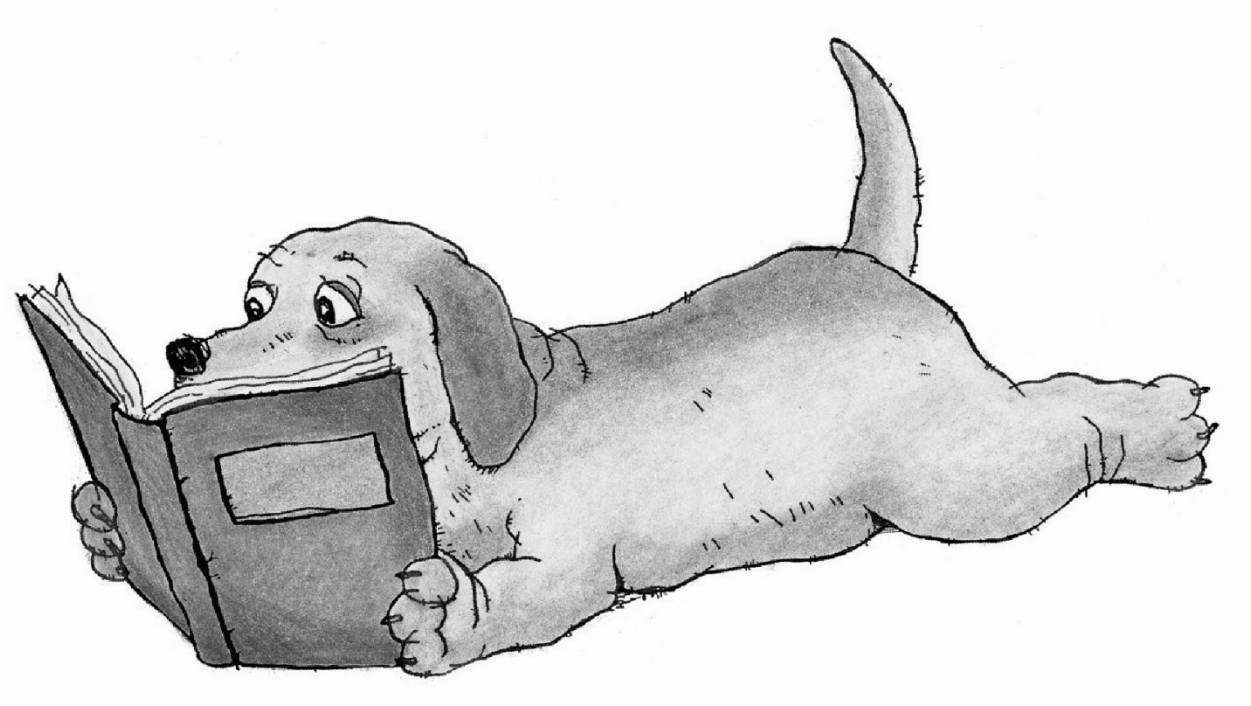
Wang, M., Park, Y., & Lee, K. (2006). Korean-English biliteracy acquisition: Cross-language phonological and orthographic transfer. *Journal of Educational Psychology*, *98*(1), 148-158.

Appendix A

Picture Description Task

Word List 1-2	
Picture 1. **read**	**lying**
Picture 2. **rain**	**table**
Word List 3-4	
Picture 3. **rice**	**a bowl**
Picture 4. **roof**	**violin**
Word List 5-6	
Picture 5. **clear sky**	**road**
Picture 6. **wedding ring**	**put on**
Word List 7-8	
Picture 7. **red rose**	**kiss**
Picture 8. **rush hour**	**passenger**
Word List 9-10	
Picture 9. **route 66**	**sign**
Picture 10. **rope**	**bite**
Word List 11-12	
Picture 11. **old man**	**nap**
Picture 12. **pack**	**clothes**
Word List 13-14	
Picture 13. **two girls**	**shopping bags**
Picture 14. **train**	**fast**
Word List 15-16	
Picture 15. **dad**	**daughter**
Picture 16. **rock music**	**guitar**

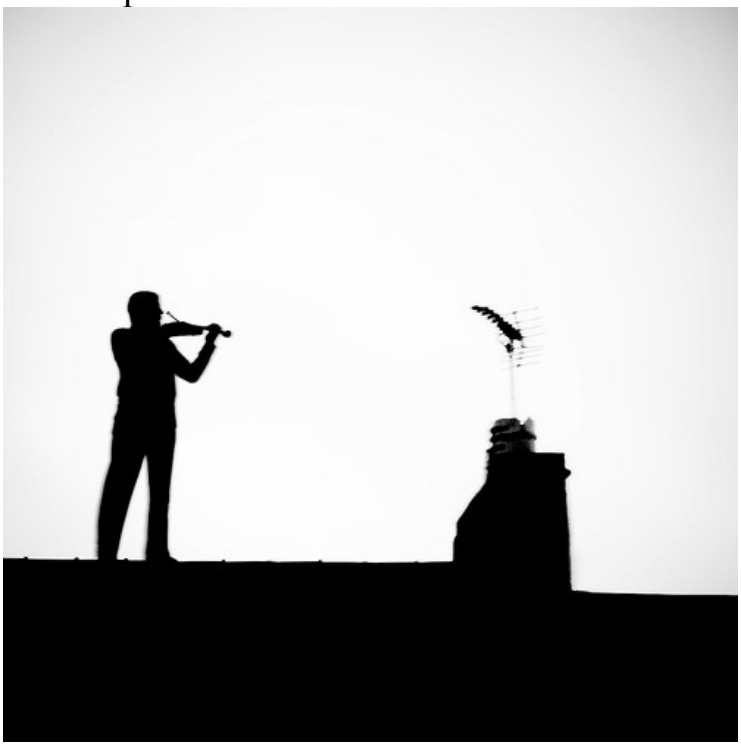
1. Please describe the picture below.



2. Please describe the picture below.

3. Please describe the picture below.

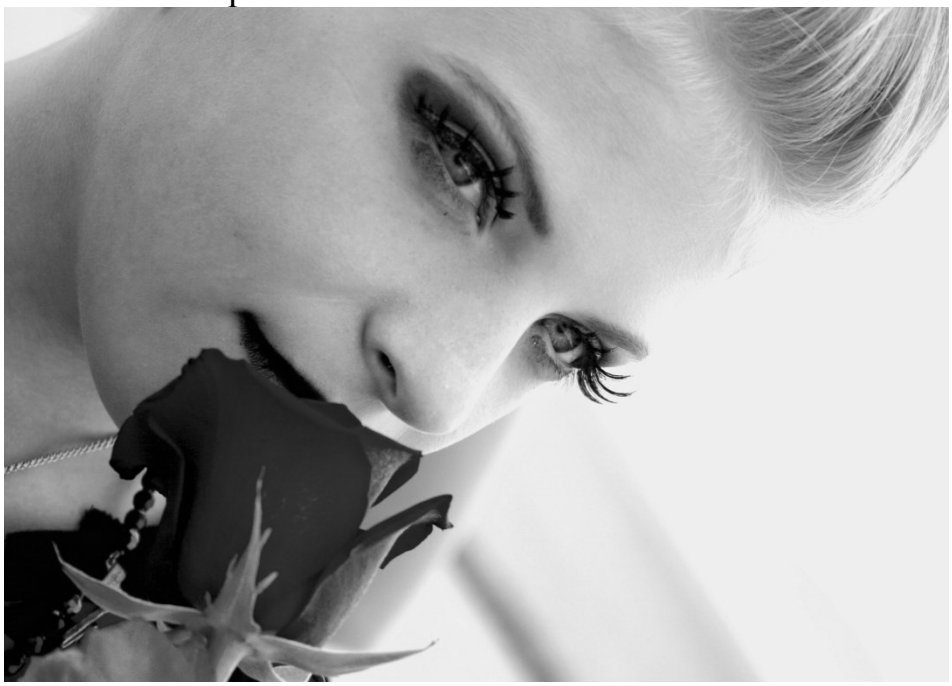
4. Please describe the picture below.



5. Please describe the picture below.

6. Please describe the picture below.

7. Please describe the picture below.



8. Please describe the picture below.

9. Please describe the picture below.

10. Please describe the picture below.

11. Please describe the picture below.

12. Please describe the picture below

13. Please describe the picture below

14. Please describe the picture below

15. Please describe the picture below

16. Please describe the picture below

Appendix B

Perception Test, Versions 1 and 2

Listening Test 1

1. f ↔ h
   fit · hit
2. r ↔ l
   road · load
3. th ↔ s
   thumb · some
4. r ↔ l
   race · lace
5. r ↔ l
   rink · link
6. r ↔ l
   rule · lure
7. r ↔ l
   read · lead
8. r ↔ l
   Ride · lied
9. r ↔ l
   roan · loan
10. r ↔ l
    rain · lane
11. r ↔ l
    rink · link
12. a ↔ u
    cat ↔ cut
13. r ↔ l
    reef · leaf
14. a ↔ u
    dad · dud
15. r ↔ l
    roof · loof
16. r ↔ l
    right · light
17. r ↔ l
    read · lead
18. b ↔ v
    best · vest
19. r ↔ l
    race · lace
20. r ↔ l
    roan · loan
21. r ↔ l
    reef · leaf
22. r ↔ l
    room · loom

23. r ↔ l
    rice · lice
24. r ↔ l
    rent · lent
25. r ↔ l
    rule · lure
26. b ↔ v
    boat · vote
27. r ↔ l
    roof · loof
28. th ↔ s
    think · sink
29. r ↔ l
    wrong · long
30. r ↔ l
    ride · lied
31. r ↔ l
    right · light
32. r ↔ l
    room · loom
33. r ↔ l
    rain · lane
34. r ↔ l
    road · load
35. a ↔ u
    lab · love
36. a ↔ u
    cap · cup
37. r ↔ l
    wrong · long
38. a ↔ o
    map · mop
39. r ↔ l
    rice · lice
40. r ↔ l
    rent · lent
41. s ↔ th
    sin · thin
42. f ↔ h
    five · hive
43. r ↔ l
    rush · lush
44. r ↔ l
    ring · ling
45. r ↔ l
    rude · lewd
46. r ↔ l
    rate · late

47. r ↔ l
    red · led
48. a ↔ u
    bat · but
49. sh ↔ s
    she · sea
50. r ↔ l
    ring · ling
51. b ↔ v
    berry · very
52. r ↔ l
    reach · leach
53. r ↔ l
    rough · laugh
54. a ↔ u
    tap · top
55. r ↔ l
    rude · lewd
56. r ↔ l
    roll · loll
57. r ↔ l
    root · loot
58. r ↔ l
    red · led
59. a ↔ u
    mad · mud
60. r ↔ l
    roll · loll
61. r ↔ l
    rope · lope
62. r ↔ l
    rate · late
63. r ↔ l
    rush · lush
64. r ↔ l
    rough · laugh
65. a ↔ u
    fan · fun
66. r ↔ l
    reach · leach
67. r ↔ l
    root · loot
68. a ↔ u
    bag · bug
69. f ↔ h
    fall · hall
70. r ↔ l
    rope · lope

Listening Test 2

1. a ↔ o
   map · mop
2. r ↔ l
   ride · lied
3. r ↔ l
   right · light
4. r ↔ l
   read · lead
5. r ↔ l
   wrong · long
6. r ↔ l
   rice · lice
7. v ↔ b
   vote · boat
8. th ↔ s
   think · sink
9. a ↔ u
   dad · dud
10. r ↔ l
    roof · loof
11. f ↔ h
    fit · hit
12. r ↔ l
    road ↔ load
13. r ↔ l
    race · lace
14. r ↔ l
    right · light
15. v ↔ b
    vest · best
16. r ↔ l
    rain · lane
17. r ↔ l
    race · lace
18. r ↔ l
    roof · loof
19. r ↔ l
    roan · loan
20. r ↔ l
    rule · lure
21. r ↔ l
    rink · link
22. r ↔ l
    reef · leaf
23. r ↔ l
    wrong · long

24. r ↔ l
    road · load
25. r ↔ l
    room · loom
26. r ↔ l
    rice · lice
27. th ↔ s
    thumb · some
28. r ↔ l
    rent · lent
29. r ↔ l
    read · lead
30. r ↔ l
    ride · lied
31. r ↔ l
    roan · loan
32. r ↔ l
    rule · lure
33. r ↔ l
    lab · love
34. r ↔ l
    room · loom
35. r ↔ l
    rain · lane
36. r ↔ l
    rink · link
37. a ↔ u
    cat · cut
38. r ↔ l
    reef · leaf
39. a ↔ u
    cap · cup
40. r ↔ l
    rent · lent
41. r ↔ l
    rope · lope
42. a ↔ u
    fan · fun
43. r ↔ l
    rough · laugh
44. r ↔ l
    root · loot
45. r ↔ l
    red · led
46. r ↔ l
    rate · late
47. r ↔ l

reach · leach

48. a ↔ u
    mad · mud
49. th ↔ s
    thin · sin
50. f ↔ h
    five · hive
51. r ↔ l
    rush · lush
52. r ↔ l
    roll · loll
53. r ↔ l
    rope · lope
54. r ↔ l
    ring · ling
55. r ↔ l
    rude · lewd
56. r ↔ l
    rate · late
57. s ↔ sh
    sea · she
58. r ↔ l
    ring · ling
59. v ↔ b
    very · berry
60. a ↔ u
    bag · bug
61. r ↔ l
    rough · laugh
62. r ↔ l
    roll · loll
63. a ↔ u
    tap · top
64. r ↔ l
    root · loot
65. r ↔ l
    rush · lush
66. f ↔ h
    fall · hall
67. r ↔ l
    rude · lewd
68. r ↔ l
    red · led
69. a ↔ u
    bat · but
70. r ↔ l
    reach · leach

Appendix C

Retrospective Interview Questions

Same for all groups, except as marked a. (control) b. (prompt group) or c. (recast group)

1. In English class, what percent do you concentrate on using correct language, and what percent do you concentrate on just showing your meaning?

2. a. Do you think the teacher should correct students in class? If so, what is the best way?
b. What did you think when the teacher said "could you repeat that?" Was it helpful?
c. What did you think when the teacher repeated your /r/ word? Was it helpful?

3. During this class, what did you do to try to improve your pronunciation of /r/? What was the most helpful part of the class for you?

4. Was it more helpful when the teacher explained how to make the /r/ sound with your mouth, or when the teacher gave you examples of the /r/ sound?

5. When you listen to hear /r/, do you pay attention to the length of the sound, or do you pay attention to something else? If so, what is it?

6. Normally, how comfortable do you feel speaking English? Did you feel comfortable practicing pronunciation of /r/ during this class?

7. Do you have any other comments about the class, or anything else you would like to tell me?