Effects of Corrective Feedback on L2 Speech Perception:

Perceptual decisions, linguistic hypotheses, and negative evidence

Andrew Hyunmin Lee

The Department of Integrated Studies in Education

(Major: Second Language Education)

McGill University, Montreal

September 2013

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

ACKNOWLEDGMENTS

I would like to extend my deepest gratitude to the beloved people who have shared my sweet and sometimes bitter moments during this thesis journey. Without them, this thesis would have never seen the light of day.

It is tremendously hard to describe in words my heartfelt appreciation for my supervisor, Dr. Roy Lyster. I at first encountered him through his seminal paper ‘Lyster & Ranta, 1997’ and then I felt so lucky to have such a great scholar as my supervisor. I indeed appreciate his sincere and extraordinary supervision and financial support for this study. He was always there for me whenever I asked for help.

I would also like to acknowledge my professor, Dr. Anila Asghar. Along with Dr. Lyster, she always welcomed me with her bright smile. Whenever I had concerns, she provided me with exceptional advice and endless encouragement to pursue my interests. I am also grateful to her for giving me an opportunity to join her meaningful research study in which I was able to reflect my work and move it forward. I owe her a lot.

My friends indeed helped me a lot to conduct this study. Particularly, I am thankful to Alan Hall, Lavinia Oncica, and Sarah-Eve Philippe-Beauchamp. When I needed them, they gave me active support without hesitation. In addition, I am indebted to Hye-Young Bang and

Shuwei Liu. Without them, I could not have managed to conduct this study by myself. I would also like to thank my participants for taking an active part in this study.

Had it not been for my friends, my life in Montreal would have been full of loneliness and boredom. They saved me. I would like to express my warm thankfulness to my dear friends being with me: Carolina Carpanzano, Cindy Chiang, Halley Silversides, Hye-Young Bang, Jade Seo, Lingwei Qian, Mao Ding, and Shuwei Liu. My students and colleagues in Korea also cheered me up whenever I had a tough time. Special thanks go to my dear colleague and friend, Sangyeon Mun.

Last but not least, I miss my family very much. My family always showed me unconditional support and love in my life. From the bottom of my heart, I would like to express my deepest love to my mom, dad, and Hyuna. They are all reason for my living.

This thesis was financially supported by a SSHRC grant awarded to Dr. Roy Lyster, a Graduate Research Mobility Award in Education, a Graduate Excellence Award granted by the Department of Integrated Studies in Education, and the Ellen Edith Grubb Stansfield Award.

ABSTRACT

Corrective feedback (CF) has been identified as an effective means for improving pronunciation in second language (L2) speech production, but there has been a general lack of relevant empirical studies regarding its effects on L2 speech perception. The current study was designed to investigate (a) the extent to which the perceptual accuracy of L2 learners is influenced by instruction with and without CF, and (b) the reasons for differential treatment effects.

Thirty-two Korean participants comprising Instruction-only and Instruction + CF groups were exposed to five one-hour pronunciation lessons that drew their attention to the English phonemic contrast /i/-/ɪ/, while only the Instruction + CF group was given relevant feedback. Forced identification tasks were provided to participants in a pretest, an immediate posttest, and a delayed posttest.

Both groups showed improvement on natural and duration-synthesized /i/, while only the Instruction + CF group showed an increase in perceptual accuracy vis-à-vis /ɪ/. Both groups retained the acquired knowledge after two weeks, but only the feedback group was able to transfer it to new words. These findings are discussed in terms of the pivotal role played by CF in developing accuracy in L2 speech perception.

RÉSUMÉ

Bien que l'on ait observé que la rétroaction corrective (RC) s'avère un moyen efficace d'améliorer la prononciation dans la production orale en langue seconde (L2), il existe peu d'études empiriques portant sur ses effets sur la perception de la parole en L2. La présente étude a été conçue pour explorer (a) dans quelle mesure un enseignement avec et sans RC influe sur l'exactitude de la perception chez les apprenants de L2 par et (b) les raisons des effets des traitements différents.

Trente-deux participants coréens, répartis en deux groupes, l'un avec enseignement sans RC et le second avec RC, ont suivi cinq cours de prononciation de 60 minutes qui attiraient leur attention sur le contraste phonémique /i/-/ɪ/ en anglais alors que seuls les membres du second groupe recevaient une rétroaction pertinente. Les participants ont eu à effectuer des tâches d'identification forcée dans le cadre d'un pré-test, d'un post-test immédiat et d'un post-test différé.

Les membres des deux groupes ont amélioré leur rendement relatif à /i/ naturel et à /i/ synthétisé d'une durée modifiée, tandis que seuls ceux du groupe avec RC ont accru l'exactitude de leur perception de /ɪ/. Les participants des deux groupes ont maintenu les connaissances acquises après deux semaines, mais seuls ceux du groupe avec RC étaient en mesure de les transférer à des mots nouveaux. L'analyse de ces résultats montre le rôle essentiel de la RC dans le développement de l'exactitude de la perception de la parole en L2.

TABLE OF CONTENTS

ACKNOWLEDGMENTS ii
ABSTRACT iv
RÉSUMÉ v
TABLE OF CONTENTS vii
LIST OF TABLES ix
LIST OF FIGURES x

CHAPTER 1. INTRODUCTION 1

CHAPTER 2. LITERATURE REVIEW 7

2.1. Linguistic targets: /i/ vs. /ɪ/ 8
2.2. Speech perception models 9
2.2.1. The Perceptual Assimilation Model 10
2.2.2. The Speech Learning Model 13
2.2.3. The Natural Referent Vowel Framework 14
2.2.4. The Linguistic Perception Model 16
2.3. Perceptual cue weighting 20
2.4. Effects of perceptual training 22
2.5. CF and L2 speech perception 26
2.6. CF and learning mechanisms 30
2.7. Motivation of the current study 32
2.8. Research questions 34

CHAPTER 3. METHODOLOGY 35

3.1. Participants 36
3.2. Testing materials 37
3.3. Procedure 40
3.3.1. Synopsis 40

3.3.2. Pretest 43
3.3.3. Instruction 43
3.3.3.1. Setting 43
3.3.3.2. Feedback sequences 47
3.3.4. Immediate posttest 51
3.3.5. Delayed posttest 51
3.4. Analyses 52
3.5. Summary 53

CHAPTER 4. RESULTS 55
4.1. Pretest–immediate posttest comparison 56
4.2. Generalizability test 60
4.3. Delayed posttest 63
4.4. The use of spectral vs. durational information 68
4.5. Summary 73

CHAPTER 5. DISCUSSION 75
5.1. Perceptual default 76
5.2. Perceptual decisions 78
5.3. Linguistic hypotheses and CF 81
5.3.1. Linguistic hypotheses and CF with /ɪ/ 81
5.3.2. Linguistic hypotheses and CF with /i/ 86
5.3.3. Ultimate attainment of the target vowels 89
5.4. Focused tasks and types of CF 90
5.5. Summary 92

CHAPTER 6. CONCLUSION 95
6.1. Implications of the present study 95
6.2. Limitations and future studies 97
6.2.1. Limitations 97
6.2.2. Future studies 97

REFERENCES 101

LIST OF TABLES

Table 1. Patterns and prediction of non-native phonemic perception in the PAM 11
Table 2. Prediction of perceptual assimilation of English /i/ and /ɪ/ by Korean learners of English 12
Table 3. The perception of acoustic input (340 Hz and 430 Hz) by a native speaker of English 19
Table 4. The perception of acoustic input (340 Hz and 430 Hz) by a native speaker of Korean 19
Table 5. Key research studies of perceptual training 23
Table 6. Target stimuli (cf. A-(9) ~ (12) for the generalizability test) 38
Table 7. Design of this study 42
Table 8. Participant information by group 44
Table 9. Means and standard deviations of pretest and immediate posttest scores by group and sound type 56
Table 10. Means and standard deviations of generalizability test scores by group and sound type 61
Table 11. Means and standard deviations of pretest, immediate posttest, and delayed posttest scores by group and sound type 64
Table 12. Perceptual preference by Korean learners of English in the pretest 76
Table 13. (Mis)match-based perceptual decisions by Korean learners of English 79
Table 14. A summary of results in the framework of (mis)match-based perceptual decisions (a = Instruction-only group; b = Instruction + CF group) 82

LIST OF FIGURES

Figure 1. Patterns of directional asymmetries in the infant vowel perception (Polka & Bohn, 2010, p. 2) 15
Figure 2. The Linguistic Perception Model (Escudero, 2005, p. 43; 2006, p. 125).. 16
Figure 3. The feedback sequences..48
Figure 4. Means of pretest and immediate posttest scores of the Instruction + CF group for each sound type 58
Figure 5. Means of pretest and immediate posttest scores of the Instruction-only group for each sound type 58
Figure 6. Means of generalizability test scores of the Instruction-only and Instruction + CF groups for each sound type 62
Figure 7. Means of pretest, immediate posttest, and delayed posttest scores of the Instruction + CF group for each sound type 66
Figure 8. Means of pretest, immediate posttest, and delayed posttest scores of the Instruction-only group for each sound type 67
Figure 9. The usage of durational information and effects of instruction (and corrective feedback) on the /i/-synthesized sounds (left: the Instruction + CF group; right: the Instruction-only group) 71
Figure 10. The usage of durational information and effects of instruction (and corrective feedback) on the /ɪ/-synthesized sounds (left: the Instruction + CF group; right: the Instruction-only group)... 72

CHAPTER 1

INTRODUCTION

Second language (henceforth, L2) learners confront a great deal of decision making. In the case of learning a new language that is seldom spoken in their communities, most of them acquire target sounds, lexical items, and syntactic knowledge in instructional contexts. While they practice what they have learned with instructors or peers, they are pushed to make countless decisions such as how to order a sentence and pronounce words in the target language. Specifically, in cases where the target language is typologically different from their native language, L2 learners are expected to encounter more complex and various decision points.

When L2 learners make linguistic decisions, they build linguistic hypotheses based on learning experiences that include instructional input and metalinguistic information. The hypotheses are sometimes sufficient to result in correct forms or sometimes non-target-like and thus lead to ill-formed utterances. In this respect, in order to facilitate successful decision making, instructors play an important role in mediating learners' challenges and target-like linguistic forms. For example, once learners make errors that are unacceptable in the target language, instructors tend to give correction (i.e., recasts) or push them to produce correct reformulations (i.e., prompts); by doing so, learners have opportunities to test their linguistic hypotheses and, if wrong, have opportunities to generate new hypotheses that reveal more target-like linguistic decisions. In this sense, corrective feedback has been highlighted as an important impetus for successful L2 acquisition.

Effects of corrective feedback have focused almost exclusively on morphosyntatic targets; in recent years, SLA researchers have converged on the need for research into the roles of corrective feedback in phonological development. Among few previous findings, Saito and Lyster (2012a, 2012b) revealed that L2 learners benefited from instruction and recasts on L2 speech production. However, it is hard to deny that speech production and perception are highly associated, and it is also well known that perception precedes production (Escudero, 2006; Flege, 1995; Leather, 1999; Llisterri, 1995). In this regard, it is inadvisable to discuss the effects of instructional treatments including corrective feedback on L2 phonological development while appealing only to the realm of L2 speech production. Moreover, many empirical studies found that L2 learners benefited solely from perceptual training without explicit pronunciation instruction on their L2 production (Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Jamieson & Rvachew, 1992; Rochet, 1995; Rvachew & Jamieson, 1995; Wang, Jongman, & Sereno, 2003). Taking previous research into account, it is postulated that corrective feedback may affect L2 learners' perception first, and then it triggers successful articulation of L2 sounds.

L2 speech perception arguably results from an interaction between physical information and the perceptual system — listeners are exposed to a number of auditory properties such as duration, static and dynamic spectral features, periodicity, noise, and intensity (Escudero & Boersma, 2004). That being said, it is the perceptual system that funnels this information and induces listeners to recognize it as a distinct phonemic segment. For instance, in terms of physical information, English vowels /i/ and /ɪ/ are discerned by vowel tenseness and length; hence, /i/ and /ɪ/, defined as a high front tense and lax vowel respectively, have different spectral information (Peterson & Barney, 1952) as well as different durational information (Peterson & Lehiste, 1960). According to previous empirical studies (e.g., Baker, Trofimovich, Mack, & Flege, 2002; Flege, Bohn, & Jang, 1997; Ingram & Park, 1997; Tsukada, Birdsong, Bialystok, Mack, Sung, & Flege, 2005), Korean learners of English have difficulty categorizing /i/ and /ɪ/, whereas native speakers of English show nearly perfect discrimination. Korean learners of English use primarily durational information when categorizing these vowels while native speakers of English are known to exploit both spectral and durational information (Bohn & Flege, 1990). In sum, in spite of the same physical information, Korean learners of English and native speakers of English show different patterns and strategies in categorizing and perceiving /i/ and /ɪ/. In this sense, the perceptual system — intrinsically linked to L1 background — is predicted to play a key role in L2 speech perception.

In this study, the effects of instructional treatment including corrective feedback on the perceptual system are investigated. So far, a great number of laboratory-based studies have found that L2 learners benefited from perceptual training in perceiving non-native phonemes. Most studies have focused on non-native consonant contrasts exclusively concerning the /r/-/l/ contrast (e.g., Bradlow et al., 1997; Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994; Strange & Dittman, 1984). There are only a few studies delving into the effects of perceptual training on L2 vowels. For instance, Wang and Munro (2004) found that perceptual training helped Mandarin and Cantonese learners of English to improve perceptual accuracy on /i/ and /ɪ/. Lee (2008) also revealed similar results with Korean learners of English. However, considering that most English as a foreign language (henceforth, EFL) learning virtually takes place in the classroom setting and that there is a huge disparity between the laboratory and classroom in terms of EFL learning, it is unclear whether the same degree of effects observed in laboratory-based perceptual training could be expected with EFL learners in the classroom setting. Nevertheless, to the best of my knowledge, there are no studies probing the effects of classroom-based perceptual training on L2 vowel perception.

As a key component of classroom-based perceptual training, this study specifically sheds light on the role of corrective feedback. Intriguingly enough, even many laboratory-based studies adopted feedback procedures during the perceptual training and also emphasized the impact of feedback on the training sessions (Hardison, 2003; Logan, Lively, & Pisoni, 1991; Wang & Munro, 2004). This gives rise to the main question driving the current study: What is the role of corrective feedback in L2 speech perception and training?

This thesis consists of a total of six chapters. Following this introductory chapter, Chapter 2 presents several speech perception models to address the question of why Korean learners of English have difficulty perceiving /i/ and /ɪ/. In addition, while presenting previous laboratory-based studies of perceptual training and its effects, the roles of corrective feedback in L2 learning mechanisms are summarized. Finally, research questions of the present study are presented.

Chapter 3 introduces the methodology designed to delve into the research questions. Specifically, the instructional treatments including focused tasks and corrective feedback are presented. Measurement and data analysis procedures are also addressed. In Chapter 4, the results are revealed.

In Chapter 5, the research questions are teased apart by integrating the empirical results and theoretical background. I specifically propose a framework of (mis)match-based perceptual decisions to explain how Korean learners of English benefit from corrective feedback in perceiving the target sounds. Lastly, in Chapter 6, a brief conclusion is presented along with limitations of the present study and suggestions for future research.

Ultimately, this study aims to convey the importance of L2 speech perception on L2 phonological development. In addition, it suggests a pivotal role for corrective feedback in speech perceptual training. The current study is thus expected to provide useful support to the fields of L2 education and speech perception research.

CHAPTER 2

LITERATURE REVIEW

In this chapter, the theoretical foundation of the current study and previous research findings are presented. Based on several empirical studies, two questions are addressed: what makes Korean learners of English (a) have difficulty perceiving /i/ and /ɪ/ correctly and (b) use durational information while categorizing these sounds? Specifically, four speech perception models — Best's (1995) Perceptual Assimilation Model, Flege's (1995) Speech Learning Model, Polka and Bohn's (2011) Natural Referent Vowel Framework, and Boersma, Escudero, and Hayes's (2003) Linguistic Perception Model — will be presented with a view to addressing the first question. The second question will be addressed through reference to Bohn's (1995) Desensitization Hypothesis and Escudero and Boersma's (2004) evidence of Full Access (Schwartz & Sprouse, 1996).

Although L2 learners are known to experience perceptual difficulties related to the new language they are learning, previous empirical studies have confirmed that these difficulties can be reduced through laboratory-based perceptual training (e.g., Thomson, 2007; Wang & Munro, 2004). However, given that most EFL learning in fact occurs in classroom settings with instructors and that there is a disparity between laboratory-based and classroom-based learning, the pedagogical relevance of laboratory-based studies remains unclear. In addition, considering there are no studies investigating the effects of classroom-based perception training[1], it is timely to delve into this research topic in the L2 education and L2 speech perception research. Specifically, this study aims to shed light on the role of corrective feedback in perceptual training. Indeed, many laboratory-based studies also acknowledged the effects of feedback in perceptual training (e.g., Hardison, 2003; Logan et al., 1991; Thomson, 2007; Wang & Munro, 2004). In this regard, the role of corrective feedback in L2 learning mechanisms and the development of L2 phonology comprise the main focus of this chapter. Finally, in light of the previous research studies, the motivation for the current study will be illustrated, followed by the research questions.

2.1. Linguistic targets: /i/ vs. /ɪ/

It is well known that L2 learners have difficulty perceiving non-native phonemes. Especially when it comes to non-native phonemes whose difference from existing native phonemes remains blurred, L2 learners are likely to be unaware of the non-native phonemic contrast. For example, English /i/ and /ɪ/ have different spectral information,

[1] Saito (2013) recently found effects for instruction treatment on the perception of English /ɹ/ by Japanese learners of English. However, the treatment including corrective feedback was designed to target L2 learners' production errors, not their perceptual errors. Saito (2013) is thus not directly addressing the effects of classroom-based perceptual training, but rather addressing transfer effects of production training on perceptual ability.

F_1 342 ~ 437 Hz for /i/; 427 ~ 483 Hz for /ɪ/, and durational information, 243 ~ 306 ms for /i/; 192 ~ 237 ms for /ɪ/ (Hillenbrand, Getty, Clark, & Wheeler, 1995). Regarding native speakers of English, they use both spectral and durational information to categorize these vowels (Bohn & Flege, 1990; Escudero & Boersma, 2004). For instance, when native speakers of English are asked to categorize /ʃip/ and /ʃɪp/ between them, they are unlikely to be confused about this task because these sounds are phonemically different in their native language. However, in an oft-cited study, Flege et al. (1997) found that Korean learners of English could not discern /i/ and /ɪ/ compared to native speakers of English. In addition, a number of previous studies (Baker et al., 2002; Ingram & Park, 1997; Tsukada et al., 2005) also revealed that Korean learners of English have difficulty categorizing and perceiving /i/ and /ɪ/ correctly. Baker et al. (2002) found that Korean learners of English tended to categorize /i/ and /ɪ/ as a single phoneme /i/, and Tsukada et al. (2005) also showed that Korean learners of English perceived /ɪ/ as /i/, /e/, or /ɛ/ while /i/ was always categorized as /i/. In sum, many empirical studies have indicated that Korean learners of English have difficulty categorizing /i/ and /ɪ/ and usually classify these sounds as one phoneme /i/. Thus, the following question grabs our attention: Why?

2.2. Speech perception models

Kuhl (2000) provides us with a clue to answer the above question: “No speaker of any language perceives acoustic reality; in each case, perception is altered in the service of language" (p. 11,852). However, this statement results in another conundrum: What is meant by 'the service of language'? Considering that L2 learners start to acquire the L2 with previous language knowledge, that is, their first language (henceforth, L1) (Ortega, 2009), it might be through the L1 that L2 speech perception is mediated. However, whether L2 speech perception is affected by the L1 and, if so, how and how much are all questions of ongoing theoretical and empirical interest in the realm of speech perception (Cook, 2002). In this section, four speech perception models are examined in order to address the question of what makes Korean learners of English have difficulty perceiving /i/ and /ɪ/ correctly.

2.2.1. The Perceptual Assimilation Model

Best's Perceptual Assimilation Model (PAM; Best, 1995) allows us to predict the initial state of L2 learners in perceiving non-native phonemes. The PAM was derived from the frameworks of Articulatory Phonology (Browman & Goldstein, 1989) and direct realism (Fowler, 1986). According to the PAM, rather than a set of innate linguistic features from Universal Grammar (UG), it is linguistic experience that leads infants to successful speech perception in their native language. When human beings listen to a particular phoneme, they recognize and extract patterns of articulatory gestures from the speech signal; in this way, infants build the L1-specific speech perception system. However, with regard to L2 learners, the L1-specific perception system impedes them from detecting the patterns of articulatory gestures in non-native sounds in that the non-native sounds are assimilated into phonemes pervading their L1-specific speech perception system. Consequently, depending on patterns of perceptual assimilation, L2 learners are good at categorizing L2 phonemes or not, as illustrated in Table 1.

Table 1. Patterns and prediction of non-native phonemic perception in the PAM

Pattern	L2 phonemes	L1 inventory	Prediction
Single-category	/A/ /B/	/Phoneme 1/	Poor
Two-category	/A/ /B/	/Phoneme 1/ /Phoneme 2/	Excellent
Category-goodness	/A/ /B/	/Phoneme 1/	Moderate to very good
Non-categorized	/A/ /B/	No exist	Poor to very good
Categorized-non-categorized	/A/ /B/	/Phoneme 1/	Very good

According to the PAM, if two non-native segments are assimilated to one L1 category, and both are equally compatible to the L1 category (i.e., single-category type), L2 listeners are highly likely to be poor at discerning those non-native sounds. If both sounds happen to be assimilated to the same L1 segment in that the L1 segment is more perceptually similar to one or the other (i.e., category-goodness type), L2 listeners are less likely to have difficulty discriminating them — in this circumstance, a non-native segment which is closer to the L1 category is prioritized. If each of two non-native segments is assimilated to a different L1 segment individually, L2 listeners show excellent categorization (i.e., two-category type).

Given that /i/ and /ɪ/ exist in the English vowel inventory while only /i/ exists in the Korean vowel inventory (Lee, 1993), Korean learners of English are predicted to show a pattern of single-category or category-goodness (see Table 2) in that they are unlikely to perform excellent discrimination with respect to these L2 vowels.

Table 2. Prediction of perceptual assimilation of English /i/ and /ɪ/ by Korean learners of English

Pattern	English phonemes	The Korean vowel inventory
Single-category	/i/ → /ɪ/ →	/i/
Category-goodness	/i/ → /ɪ/ →	/i/

2.2.2. The Speech Learning Model

Flege's Speech Learning Model (SLM) attempts to elucidate how L2 learners acquire L2 phonemes. Compared to the PAM, it pays more attention to L2 learners, not naïve listeners. The SLM primarily dealt with development of L2 speech production (Flege, 1995); however, more recently, it started to focus on L2 speech perception (Flege, 2003) while attributing successful L2 speech production to accurate L2 speech perception. One of the intriguing postulates of this model is that "The mechanisms and processes used in learning the L1 sound system, including category formation, remain intact over the life span, and can be applied to L2 learning" (Flege, 1995, p. 239). However, according to SLM, being applied to L2 learning is constrained by several factors, one of which is 'perceived cross-linguistic similarity'. When L2 learners acquire non-native phonemes, the phonemes are classified into 'new' or 'similar'. For example, non-native segments that do not prevail in the L1 sound system are defined as 'new' sounds while 'similar' sounds are those perceived as somewhat similar in the L1 system; in this regard, L2 learners have more difficulty perceiving similar sounds than new sounds, and they need more time and intervention to acquire similar sounds compared to new sounds.

Considering English /i/ and /ɪ/ within the framework of SLM, these sounds are rather similar to /i/ existing in the Korean vowel inventory such that Korean learners of English are expected to have difficulty perceiving these sounds correctly. It is also predicted that a great deal of time and learning are required for them to acquire this phonemic contrast.

2.2.3. The Natural Referent Vowel Framework

Polka and Bohn's (2011) Natural Referent Vowel (NRV) Framework suggests an acoustic perspective of why Korean learners of English have difficulty perceiving English /ɪ/ compared to /i/. With a number of speech perception studies (e.g., Bohn & Polka, 2001; Polka & Bohn, 1996, 2003; Polka & Werker, 1994), Polka and Bohn found directional asymmetries occurring in infants' vowel perception. Simply put, infants easily detect a vowel change from a more central (e.g., /ɪ/) to a more peripheral vowel (e.g., /i/), whereas they are unlikely do so in the reverse situation (i.e., from /i/ to /ɪ/) as illustrated in Figure 1. For example, monolingual English infants preferred /hid/ between /hid/ and /hɪd/. The NRV explains that such directional asymmetries result from high articulatory-acoustic properties of peripheral vowels, which could be defined as natural referent vowels. The peripheral vowels are in detail on formant frequency convergence or focalization wherein they are perceptually salient and stable (Polka & Bohn, 2011, p. 8). However, in cases where a phonemic contrast is crucially important (e.g., /i/ and /ɪ/ in L1 English), such early vowel bias is gradually reduced with L1 experience.

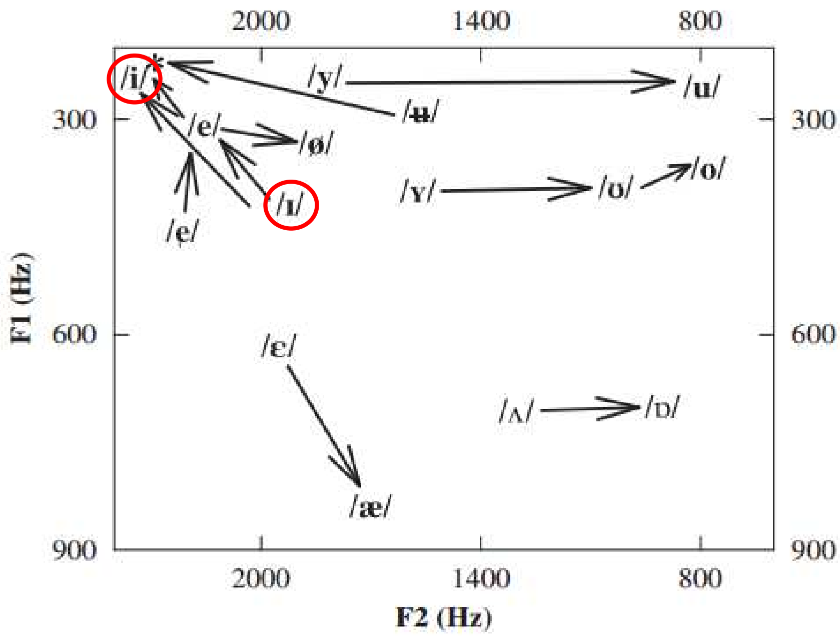

Figure 1. Patterns of directional asymmetries in the infant vowel perception (Polka & Bohn, 2011, p. 2)

The most interesting prediction of this model is that "This bias may also re-appear when perceivers are mapping out a new vowel system to learn a second language, possibly when new L2 vowel categories fall in regions of the vowel space that are not firmly committed to the L1" (Polka & Bohn, 2011, p. 8). Therefore, when Korean learners of English are exposed to /i/ and /ɪ/, it is predicted that they perceive /i/ as more salient than /ɪ/. Consequently, this might prevent Korean learners of English from detecting /ɪ/ and the /i/-/ɪ/ phonemic contrast, and in turn delay acquisition of the non-native phonemic contrast which is crucial in the target language.

2.2.4. The Linguistic Perception Model

The Linguistic Perception Model (LP; Boersma, Escudero, & Hayes, 2003) provides us with a more sophisticated approach to understanding how L2 learners perceive non-native phonemes and acquire them. This model starts from the framework of Functional Phonology (Boersma, 1998), which argues that linguistic knowledge underlies speech perception (Escudero, 2005). The most compelling component of this model is 'Boersma (1998)'s Optimality-Theoretic (OT; Prince & Smolensky, 1993) perception grammar'. According to the LP model, gradient acoustic signals are processed through the perception grammar; they in turn become categorical input to the recognition system, which enables access to lexical representation as illustrated in Figure 2:

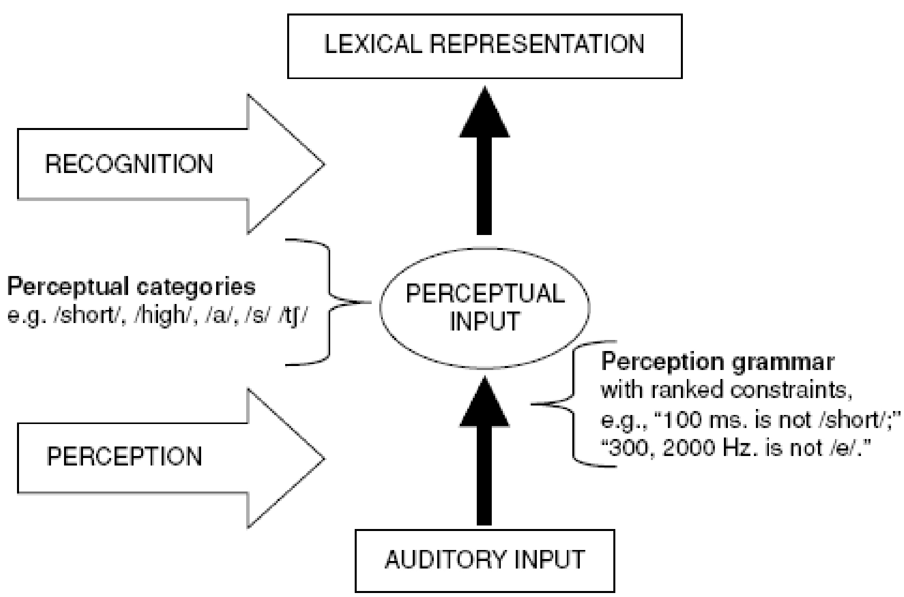

Figure 2. The Linguistic Perception Model (Escudero, 2005, p. 43; 2006, p. 125)

The perceptual grammar is operated by ranked constraints in that an optimal candidate is selected as a categorical input. According to the LP model, the constraints emerge from language experience such as human interaction, not from the innately given, such that different language speakers have different ranked constraints. In addition, compared to the previous models, the LP model addresses an explicit learning mechanism, the Gradual Learning Algorithm (GLA; Boersma & Hayes, 2001). It is comprised of four steps: (1) a datum, (2) generation, (3) comparison, and (4) adjustment. A datum refers to an adult surface form (i.e., input data). Generation is an infinite set of output candidates; among them, according to the current grammar, an optimal output becomes linguistic representation in response to a learning datum.
If there is no mismatch between a learning datum and a linguistic representation, the GLA is not activated. However, if a mismatch is detected during the process of comparison (e.g., learners notice mismatches through semantic contexts), the GLA pushes the current grammar to be adjusted by re-ordering each of constraints so that learners' grammar registers a 'no-mismatch' state between a learning datum and their output.

Based on the LP model, Escudero and Boersma (2004) suggested an L2 version of the LP model (L2LP). The key premises of the L2LP include three components: (a) L2 learners start L2 acquisition by copying the L2 perceptual grammar and lexical representation (Full Copying);

(b) as in the LP model, the GLA is the prominent learning mechanism (Full Access); and (c) L1 and L2 develop two separate systems such as the perceptual grammar wherein L2 learners achieve native or native-like perception in the L2 as well as L1 (Full Proficiency).

The F_1 of English /i/ and /ɪ/ corresponds approximately to 338 Hz and 437 Hz, respectively (Yang, 1996). By simulating Yang's (1996) data 1,000,000 times using the rnorm function in *R* (R Development Core Team, 2010), a F_1 boundary of 450 Hz between /i/ and /e/ was obtained in the Korean vowel inventory. For example, if the F_1 is below the boundary of 450 Hz, /i/ is the most optimal outcome, whereas it is over 450 Hz (preciously, 450 Hz $< F_1 <$ 650 Hz), /e/ overrides /i/ and becomes the most optimal outcome. With this simulation data and previous research studies regarding the L2LP (e.g., Escudero, 2005; Escudero & Boersma, 2004), perceptual grammars of native speakers of English and Korean are constructed as below (see Tables 3 and 4), which enable us to presuppose how differently both language group speakers categorize the same auditory input.

Table 3. The perception of acoustic input (340 Hz and 430 Hz) by a native speaker of English

Acoustic input [340 Hz]	340 Hz not /ɪ/	430 Hz not /i/	430 Hz not /ɪ/	340 Hz not /i/
☞ /i/				*
/ɪ/	*!			

Acoustic input [430 Hz]	340 Hz not /ɪ/	430 Hz not /i/	430 Hz not /ɪ/	340 Hz not /i/
/i/		*!		
☞ /ɪ/			*	

Table 4. The perception of acoustic input (340 Hz and 430 Hz) by a native speaker of Korean

Acoustic input [340 Hz]	340 Hz not /e/	430 Hz not /e/	430 Hz not /i/	340 Hz not /i/
☞ /i/				*
/e/	*!			

Acoustic input [430 Hz]	340 Hz not /e/	430 Hz not /e/	430 Hz not /i/	340 Hz not /i/
☞ /i/			*	
/e/		*!		

As Tables 3 and 4 suggest, both language listeners optimally perceive 340 Hz as /i/. However, in terms of 430 Hz, a native speaker of Korean is predicted to perceive it still as /i/ while a native speaker of English categorizes it as another phoneme, /ɪ/. Therefore, in spite of the same auditory input (that is, 430 Hz), both language groups result in different perceptual representation; Korean learners of English are less likely to categorize 430 Hz as /ɪ/. According to the LP and L2LP, it is because of the perceptual grammar.

2.3. Perceptual cue weighting

As previous studies have empirically investigated (Baker et al., 2002; Flege et al., 1997; Ingram & Park, 1997; Tsukada et al., 2005), Korean learners of English have difficulty categorizing and perceiving English /i/ and /ɪ/. So far, four speech perception models have been invoked to explain the source of difficulty. Intriguingly enough, in spite of the confusion regarding categorizing /i/ and /ɪ/, Korean learners of English are known to come up with their own strategy to categorize these vowels: using durational information. As Flege et al. (1997) found, Korean learners of English highly rely on durational information. For example, regardless of spectral information (mainly, the F_1), if it is long, they tend to consider it /i/; if it is short, they classify it as /ɪ/. It might be possible to argue that such phenomena are evidence of L1 transfer insofar as length can be used to infer a phonemic contrast in Korean. However, length contrast is in the process of disappearing in contemporary Korean and preserved only in older Korean speakers (i.e., over 50 years old) (Baker & Trofimovich, 2005; Ingram & Park, 1997; Magen & Blumstein, 1993). This begs the question as to how Korean learners of English are able to adopt durational information to contrast /i/ and /ɪ/.

Bohn (1995) attempted to answer this question by formulating a Desensitization Hypothesis. It posits that if L2 learners are desensitized to a spectral difference in contrastive phonemes, with respect to categorizing them, durational information is inadvertently entrenched. This is because "duration cues in vowel perception are easy to access whether or not listeners have had specific linguistic experience with them" (Bohn, 1995, p. 294). He claimed that the Desensitization Hypothesis is closely associated with a universal availability of durational information as a means of categorizing /i/ and /ɪ/.

In contrast, Escudero and Boersma (2004) attributed such tendency to evidence of Full Access (Schwartz & Sprouse, 1996). Full Access presupposes that if L2 learners are eager to acquire linguistic features that are not available in their L1 grammar, they gradually have access to an L1-like acquisition device. In terms of discerning a phonemic contrast, by citing 'distributional learning' from the research study (Maye, Werker, & Gerken, 2002), they argued that a length contrast is an example of category formation (/i/ vs. /i:/) that can be easily found as an L1 acquisition strategy, whereas splitting one assimilated sound is a case of category split (/i/ vs. /ɪ/) that can hardly be permitted in typical L1 acquisition. In sum, they concluded that L2 acquisition entails access to an L1-like acquisition device (Full Access) wherein L2 learners prefer to use length difference rather than to split one category into two different categories in order to make a contrast.

2.4. Effects of perceptual training

In light of the initial state of Korean learners of English discussed so far, the next question is: In spite of the difficulty, is it possible for them to improve their perceptual accuracy? In order to delve into this question, a number of empirical studies have been conducted. Table 5 summarizes key research studies.

Table 5. Key research studies of perceptual training

Author	Language group	Target	Training L2 learners?	Classroom-based training?	Feedback intervention?
Carney, Widin, & Viemeister (1977)	L1 English	/b/, /p/	No	No	Yes
Pisoni, Aslin, Perey, & Hennessy (1982)	L1 English	/b/, /p/, /pʰ/	No	No	Yes
Werker & Tees (1984)	L1 English	/ṭ/, /t/ ; /k̀/, /q̀/	No	No	Yes
Wayland & Guion (2004)	L1 English and Mandarin	Thai tone	No	No	Yes
Strange & Dittman (1984)	EFL Japanese	/r/, /l/	Yes	No	Yes
Jamieson & Morosan (1986, 1989)	EFL Canadian French	/θ/, /ð/	Yes	No	Yes
Lively, Logan, & Pisoni (1993)	EFL Japanese	/r/, /l/	Yes	No	Yes
Lively, Pisoni, Yamada, Tohkura, & Yamada (1994)	EFL Japanese	/r/, /l/	Yes	No	Yes
Bradlow, Pisoni, Akahane-Yamada, & Tohkura (1997)	EFL Japanese	/r/, /l/	Yes	No	Yes
Hardison (2003)	EFL Korean & Japanese	/r/, /l/	Yes	No	Yes
Wang (2002)	EFL Mandarin & Cantonese	/i/, /ɪ/; /u/, /ʊ/; /ɛ/, /æ/	Yes	No	Yes
Wang & Munro (2004)	EFL Mandarin & Cantonese	/i/, /ɪ/; /u/, /ʊ/; /ɛ/, /æ/	Yes	No	Yes
Lambacher, Martens, Kakehi, Marasinghe, & Molholt (2005)	EFL Japanese	/æ/, /ɑ/, /ʌ/, /ɔ/, /ɝ/	Yes	No	Yes
Thomson (2007)	EFL Mandarin	/i/, /ɪ/; /e/, /ɛ/, /æ/; /ɒ/, /ʌ/; /o/; /ʊ/, /u/	Yes	No	Yes
Lee (2008)	EFL Korean	/i/, /ɪ/; /ɛ/, /æ/	Yes	No	Yes
Wang, Spence, Jongman, & Sereno (1999)	L1 English	Mandarin tone	Yes	No	Yes

Along with Table 5, it is well established that learners benefit from training in developing perceptual accuracy. A great number of perceptual training studies (Bradlow et al., 1997; Carney et al., 1977; Hardison, 2003; Jamieson & Morosan, 1986, 1989; Lively et al., 1993; Lively et al., 1994; Pisoni et al., 1982; Strange & Dittman, 1984; Werker & Tees, 1984) have dealt with contrasts of consonants (e.g., exclusively for /r/-/l/ contrast); it turned out that L2 learners benefited from short-term laboratory training by re-categorizing confusable consonant contrasts. In addition, effects of perceptual training on non-native prosody have been also investigated in this field of study (Wang et al., 1999; Wayland & Guion, 2004). However, compared to consonants, a relatively small number of findings related to vowels exist in the field of study (Lambacher et al., 2005; Lee, 2008; Thomson, 2007; Wang, 2002; Wang & Munro, 2004). Among studies related to vowels, Wang and Munro (2004) investigated how Mandarin and Cantonese learners of English perceive /i/ and /ɪ/, and whether they could benefit from computer-based perceptual training. Their study showed that a group exposed to perceptual training outperformed a control group in discerning /i/ and /ɪ/, and they were also able to transfer this knowledge to new contexts. After performing a similar study, Lee (2008) also found similar findings with Korean learners of English.

However, a disproportionate number of studies have been conducted in the laboratory setting where L2 learners practice testing items such as forced-identification, A(B)X discrimination, or goodness judgment tasks used for measurement sessions. The only difference is that learners are provided with feedback and repetition when they are only wrong or both right and wrong. In terms of feedback, when learners choose a right answer, visual and (or) audio confirmation is given to learners. If learners select a wrong answer, visual and (or) audio signals are provided such that they help the learners to notice errors and select right answers by repetition (e.g., visual feedback; Leather, 1990).

Notwithstanding the effects of perceptual training embedded with feedback, it is unclear whether such laboratory training procedures are feasible and easily accessible for L2 learners. Given that people mostly start to learn EFL in the classroom setting with instructors, the previous research studies entail a lack of pedagogical relevance. Nevertheless, there is no specific classroom-based research that has investigated whether L2 learners benefit from classroom-based perceptual training including L2 instruction and oral feedback with instructor-student interaction.

What attracts our attention is that many laboratory-based studies have paid attention to the importance of feedback, albeit simply (e.g., 'right'-'wrong' feedback). For example, Logan et al. (1991) emphasized immediate feedback, while mentioning that feedback helped learners to consistently focus on a key feature of the target stimuli. Hardison (2003) also emphasized the role of feedback as the following:

> Training with a large and variable stimulus set using natural speech and minimal pairs contrasting /ɹ/ and /l/ in a variety of word positions produced by multiple talkers with feedback resulted in significant improvement in identification accuracy, generalization to novel stimuli and a new voice, transfer to production improvement, and retention of improved abilities. … feedback versus no feedback perceptual training was consistently superior, especially under conditions involving a high degree of within-category variability that characterizes the phonetic variants of AE /ɹ/ and /l/ (p. 498-499).

However, again, there are few research studies that have investigated the effects of feedback on speech perception even in laboratory-based perceptual training, as Wang and Munro (2004) stated: "This study also provides evidence that identification training with feedback can improve ESL speakers' performance on English vowel contrasts, a research focus that so far has received little attention" (p. 550).

2.5. CF and L2 speech perception

So, what leads L2 speech perception researchers to have a great interest in feedback? Since SLA researchers started to delve into the role of corrective feedback in L2 learning, an incredible number of research studies have been published concerning CF effectiveness (Li, 2010; Lyster & Saito, 2010; Mackey & Goo, 2007; Russell & Spada, 2006), different types of CF and corresponding effects (Ammar & Spada, 2006; Doughty & Varela, 1998; Ellis, Loewen, & Erlam, 2006; Lyster, 1998, 2004a; Lyster & Izquierdo, 2009; Lyster & Ranta, 1997; Sheen, 2007; Yang & Lyster, 2010), the relationship between CF and classroom context (Lyster & Mori, 2006), and CF on different linguistic targets (Lyster, Saito, & Sato, 2013).

Corrective feedback plays two main roles: (a) indicating errors and (b) carrying out error correction. In other words, corrective feedback could be defined as verbal movements indicating that errors occur and, at the same time, carrying out error correction by means of providing learners with target-like forms. In terms of 'who' enacts the verbal moves, the previous literature focused on instructors; however, peers have recently started to be highlighted as a potential agent (Sato & Ballinger, 2012; Sato & Lyster, 2012).

Despite of the corrective intention of feedback moves, unless learners notice that they made errors, the ultimate goal of corrective feedback (i.e., L2 development) turns out to be hard to accomplish.
In this regard, 'how' to indicate errors is one of the most important parts of corrective feedback procedure. 'How' to indicate errors is also closely related to 'by whom' correct forms are delivered. There are two main families of feedback in the literature; one is reformulations and the other

is prompts (Ranta & Lyster, 2007). Reformulations entail recasts and explicit correction in that correct forms per se are articulated by instructors. In other words, when L2 learners make errors in the target language, instructors provide them with accurate forms with the purpose of having L2 learners notice mismatches between their utterances and the correct forms. On the other hand, prompts aim to push them to self-correct with opportunities intentionally arranged by instructors.

For example, when L2 learners make an error, instructors give them information that their utterance is not acceptable in the target language without an explicit correct form, and then provide them with dialogical space to elicit the correct forms from the learners as illustrated below:

> Student: I goed to the school yesterday.
>
> Instructor: No, irregular verb!
>
> (Dialogical space)
>
> Student: I…go, went, gone — I … went to the school yesterday.

Sheen and Ellis (2011) present feedback as being either explicit or implicit. For example, although correct forms are articulated by instructors, depending on the explicitness of ‘how’ to indicate errors (e.g., conversational recasts vs. didactic recasts), learners notice mismatches more easily or not. In terms of prompts, the more explicitly corrective feedback conveys negative evidence (e.g., repetition vs. metalinguistic clue), the more likely learners are to respond with correct forms (at least notice their errors and try to correct them) such that they reflect on their linguistic knowledge or hypotheses.

Finally, when it comes to 'what' the feedback moves target, it could be morphosyntactic, lexical, phonological, or pragmatic errors (Lyster et al., 2013). However, a majority of research studies have focused on L2 learners' morphosyntactic errors; admittedly, L2 phonology-related studies received little attention after the golden age of audio-lingualism in the field of L2 education. A few studies recently started to investigate the effects of instruction and corrective feedback on L2 speech production, focusing on intelligible pronunciation (Saito, 2011; Saito & Lyster, 2012a, 2012b). Saito and Lyster (2012a) found that Japanese learners of English benefited from instruction and partial recasts on English /ɹ/; Saito and Lyster (2012b) also revealed the effects of instructional intervention on English vowels with Japanese learners of English. However, considering that speech production and perception are closely associated and that production accuracy reflects the degree of perceptual accuracy (Escudero, 2006; Flege, 1995; Leather, 1999; Llisterri, 1995), it is hard to discuss the development of L2 phonology without mentioning L2 speech perception. Nevertheless, again, there are no research studies delving into the effects of classroom-based perceptual training such as instruction and corrective feedback on L2 speech perception.

2.6. CF and learning mechanisms

As mentioned, a great number of CF-related research studies have empirically demonstrated that CF groups outperformed non-CF groups in acquiring L2 morphosyntactic and phonological targets (see meta-analyses by Li, 2010; Lyster & Saito, 2010; Mackey & Goo, 2007; Russell & Spada, 2006). However, while shedding light on the treatment effect, it is also very important to discuss how CF affects the L2 learning mechanisms.

The acquisitional role of CF could be explained within cognitive theory based on information-processing models (DeKeyser, 1998, 2001; Skehan, 1998). In cognitive theory, "representation refers to the nature of items stored in memory, while processing is concerned with access to these representations" (Lyster, 2004b, p. 323). According to Skehan's (1998) information-processing models, information processing, described as skill acquisition (Anderson, 1983, 1985; Johnson, 1996), refers to a gradual transition from declarative to procedural knowledge. Implementing Skehan's (1998) models into L2 acquisition, declarative knowledge stands for grammatical, phonological, and lexical information while procedural knowledge connotes actual abilities to access the declarative knowledge (Lyster, 2004b; Sato & Lyster, 2012).

At the initial stage, to facilitate the transformation of declarative knowledge into procedural knowledge, a great deal of attention and short-term memory are required (i.e., controlled processing), whereas automatic processing results from practice and long-term memory (Shiffrin & Schneider, 1977). In addition, in order to expedite proceduralization of declarative representations and to organize a better declarative representation, repeated practice and feedback are inevitably necessary (Anderson, Corbett, Koedinger & Pelletier, 1995; DeKeyser, 1998; McLaughlin, 1990). In this respect, Lyster (2004b) argued that, in the absence of feedback, interlanguage representations per se are highly likely to become automatized, which hinders L2 learners from attaining correct linguistic targets. In sum, based on cognitive theory, by restructuring interlanguage representations, corrective feedback helps declarative representations to be proceduralized with the linguistically accurate forms; it ultimately results in successful L2 development.

In addition, CF is compatible with three major SLA hypotheses: Krashen's Input Hypothesis (Krashen, 1982, 1985), Swain's Output Hypothesis (Swain, 1985, 1995, 2005), and Schmidt's Noticing Hypothesis (Schmidt, 1990, 1993, 1995, 2001). For example, recasts, defined as "the teacher's reformulation of all or part a student's utterance, minus the error" (Lyster & Ranta, 1997, p. 46), could provide L2 learners with comprehensible input (Krashen's Input Hypothesis). At the same time, prompts, which "push learners to self-repair" (Lyster & Mori, 2006, p. 271), could offer L2 learners opportunities to confirm whether what they know is acceptable in the target language (Swain's Output Hypothesis). In this respect, CF helps them to realize the gap between their interlanguage representation and the target language, and consequently, to update their interlanguage system to be more target-like (Schmidt's Noticing Hypothesis).

2.7. Motivation of the current study

Studies of L2 perceptual training have been conducted exclusively in the laboratory setting. However, if taking a close look at how the laboratory perceptual training was operationalized, surprisingly or quite obviously, they exhibit several teaching techniques adopted in the traditional L2 classroom. One of them is to provide L2 learners with feedback indicating errors and offering opportunities to correct errors. However, due to the limits of the laboratory-based training, most feedback types are rather simple, showing visual or audio cues to inform learners of wrong answers. As Wang and Munro (2004) mentioned concerning their laboratory-based studies, even from the laboratory-based perspective, the role of feedback in perceptual training is still waiting to be uncovered.

During the previous 20 years, many research studies have been done in the realm of corrective feedback. Particularly, in terms of phonological targets, there are a few research studies and all of them focus on the development of L2 speech production and the effects of feedback on learners' misarticulation (e.g., Saito, 2012; Saito & Lyster, 2012a; 2012b). However, it is hard to deny that speech production and perception are inextricably linked. Many researchers (Escudero, 2006;

Flege, 1995; Leather, 1999; Llisterri, 1995) emphasized the role of speech perception in the acquisition of L2 phonology while arguing that speech perception precedes production. In addition, the previous findings also suggest that L2 learners benefited solely from perceptual training without explicit pronunciation instruction on their L2 production (Bradlow et al., 1997; Jamieson & Rvachew, 1992; Rochet, 1995; Rvachew & Jamieson, 1995; Wang et al., 2003). It is thus speculated that CF may affect L2 learners' perception first, and then it triggers successful production of L2 sounds.

Strange and Dittmann (1984) stated that

> Native Japanese speakers learning English have difficulty perceptually differentiating the liquid consonants /r/ and /l/, even after extensive conversational instruction. ... Improved laboratory training tasks may be useful in establishing categorical perception of these contrasts. (p. 131)

However, we need to take into account that most of people start to learn EFL in the classroom with instructors. In this sense, with a general lack of studies of L2 vowel perception and, more importantly, the wide disparity between laboratory-based perceptual training and instructor-student interaction in the classroom where most of L2 learning practically occurs, it is timely to delve into an exploration of the extent to which L2 learners benefit from instruction and corrective feedback on L2 vowel perception.

2.8. Research questions

In light of the previous research studies and motivation of the current study, the research questions are as follows:

1. If instruction and corrective feedback are provided, to what extent do Korean learners of English improve perceptual accuracy in categorizing /i/ and /ɪ/?
   (a) Do the effects of Instruction-only and Instruction + CF differ? If so, how?
   (b) To what extent do Instruction-only and Instruction + CF help Korean learners of English to perceive the target vowels in words that had not appeared during the instructional sessions? Do they differ?
2. Are the effects of instruction and CF on perceptual accuracy retained after two weeks?
3. Can Instruction-only and Instruction + CF induce Korean learners of English to focus on spectral information regardless of duration differences? If so, do they differ?

In the next chapter, the research methodology adopted to address these questions will be outlined in detail.

CHAPTER 3

METHODOLOGY

In this chapter, the methodology designed to investigate the research questions is discussed. First, characteristics of participants and recruiting procedure are explained. Second, in order to measure the extent to which the L2 learners manage to categorize the target vowels, testing materials and the design of perception test are presented. Third, each step of this classroom-based experiment is illustrated: (a) a pretest, (b) instruction, (c) an immediate posttest, and (d) a delayed posttest. Specifically, after the pretest, the L2 learner participants were assigned to either Instruction-only or Instruction + CF group wherein they took five 1-hour pronunciation lessons with form-focused instruction. Focused tasks and corrective feedback sequences will be mainly elucidated while describing the instructional sessions. Finally, specific statistical designs of analyzing testing results are introduced, focusing on delving into the current research questions.

3.1. Participants

The current study involved five types of participants:

1. NS (Native-Speaker) Speakers (*n* = 2)
2. NS Listeners (*n* = 16)
3. ESL Instructors (*n* = 3)
4. L2 learners receiving instruction only (*n* = 16)
5. L2 learners receiving instruction with corrective feedback (*n* = 16)

The two NS participants comprising the NS Speakers group included one male and one female speaker recruited from General American (GA) speakers, which had been judged by their self-reporting (e.g., biographical information such as hometowns and places of education). Their recording samples were screened by another GA speaker and acoustically analyzed with Praat (Boersma & Weenink, 2013) to be compared with previous acoustic data (Yang, 1996).

The 16 NS perception participants comprising the NS Listeners group included four males and twelve females recruited from English monolingual speakers and English bilingual speakers, which had been verified by their self-report questionnaire. Finally, the three ESL instructors all met the criteria described for the NS Listeners group and had ESL teaching experience. They were recruited through advertisements posted at universities in Montreal, Canada. In addition, snowball sampling (i.e., requesting the selected participants to help identify and contact other individuals through their social and professional networks) was also encouraged.

The L2 learner participants were recruited from Korean learners of English. Assuming that the non-native speakers of English know target vowels, but still have difficulty perceiving them differentially, the researcher recruited only intermediate learners of English whose perceptual accuracy was measured during the pretest.

Participants were invited to participate in five 1-hour classes of pronunciation instruction at no cost. At the same time, they were asked to participate in three test sessions: a pretest, an immediate posttest, and a delayed posttest. The advertisements were posted at universities with ESL programs in Montreal, private language institutions, and community websites such as the Korean Community of Greater Montreal (http://kcc.montrealkorean.com/). Snowball sampling was highly promoted as well.

3.2. Testing materials

A series of studies (Lively et al., 1993; Lively et al., 1994; Logan et al., 1991; Wang et al., 1999) confirmed that perceptual accuracy is influenced by speaker variability. In light of these findings, each of the target words and distracters was recorded by one male and one female GA speaker, and the same sound files were consistently used for all of the test sessions. In addition, the L2 learner participants were equally exposed to three instructors such that they were under the same control regarding the influence of speaker variability and instructors per se.

Target words were chosen from minimal pairs of CVC monosyllabic words. Twelve sets of /i/-/ɪ/ pairs were collected from the Longman Communication 3000, which is a list of the 3000 most frequent words in both spoken and written English (Longman Dictionary of Contemporary English, 2009, p. 2,044-2,059). Moreover, so as to prevent participants from noticing the linguistic target of the experimental, 12 sets of distracters (e.g., half of them with only onset changed; the rest with only the nucleus changed) were also included as follows:

Table 6. Target stimuli (cf. A-(9) ~ (12) for the generalizability test)

No.	**A: Target words**	**B: Distracters**
(1)	beat-bit	book-look
(2)	cheap-chip	cut-shut
(3)	feel-fill	deal-seal
(4)	heat-hit	fat-hat
(5)	lead-lid	gun-sun
(6)	scene-sin	pin-thin
(7)	seat-sit	bag-big
(8)	sheep-ship	chat-cheat
(9)	leave-live	mass-miss
(10)	peak-pick	net-nut
(11)	reach-rich	rob-rub
(12)	seek-sick	tap-tip

The two native speakers of English in the NS Speakers group were instructed to produce each word (see Table 6) twice naturally. The words were recorded by a digital recorder through a unidirectional microphone installed in a sound attenuated room at the Department of Linguistics, McGill University. The productions were digitalized at 44,100 Hz, and one out of two productions was selected by each speaker's counterpart. Finally, the sound files were analyzed with Praat (Boersma & Weenink, 2013), and it was confirmed that they fell in the average of GA speakers' vowel formants (Yang, 1996).

Adding to sets of natural speech stimuli, in order to investigate the extent to which native and non-native speakers of English rely on durational and spectral information, each word — except for distracters — was synthesized with five different durations in equal increments, using a Klatt (1980) synthesizer. Each word was therefore prepared as 12 different tokens: Speaker (male and female speaker) × Duration (natural, 130 ms, 160 ms, 190 ms, 220 ms, and 250 ms).

Using the PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993), forced identification tasks were designed. Specifically, two words were shown on the computer screen. After participants heard a given word, they were asked to choose what they had heard; for example, the words 'sheep' and 'ship' appeared on the computer screen accompanied by the sound /ʃɪp/ (synthesized as 250 ms). If they were unsure about answers, they were told to make a guess. Again, distracter tasks were embedded alongside target tasks.

Following this procedure, the non-native participants were asked to complete a total of 240 identification tasks in the pretest and delayed posttest (Target tasks: 16 words × 12 tokens = 192; distracters: 24 words × 2 tokens = 48) and a total of 336 identification tasks in the immediate posttest (Target tasks: 24 words × 12 tokens = 288; distracters: 24 words × 2 tokens = 48). There was no pre-set interval time between tasks. Instead, participants needed to press the 'next' key before moving on to another task, so could take as much time as they needed.

3.3. Procedure

3.3.1. Synopsis

Table 7 summarizes the design of the study, which consists of four steps: pretest, instructional component, immediate posttest, and delayed posttest. First, before the instructional sessions, 16 native speakers of English and 32 Korean learners of English took the pretest at the research laboratory, Faculty of Education, McGill University. Taking account of their age, gender, length of residence (henceforth, LOA), and %-correct identification scores of the pretest, the 32 non-native speakers were assigned to either Instruction-only or Instruction + CF group. Within each group, one class was composed of 8 learners (i.e., 2 classes per group; thus a total of 4 classes). Instructional sessions consisted of five 1-hour pronunciation lessons. During every class, the corresponding instruction was provided to both groups, whereas CF was given only to the Instruction + CF group during focused tasks. Finally, after five 1-hour lessons, the Korean learners of English took an immediate posttest.

In order to see whether the effects of the instruction and CF were retained after the instruction, the participants also took part in a delayed posttest two weeks later. Each step is illustrated as follows:

Table 7. Design of this study

Steps	Procedures	Participants
Pretest	1) Forced identification tasks • For the L2 learners (n = 240), Words: A-(1) ~ (8) and B-(1) ~ (12) • For the NS Listeners (n = 336), Words: A-(1) ~ (8) and B-(1) ~ (12) + A-(9) ~ (12) for the generalizability test	1) NS Listeners (n = 16) 2) Instruction-only group (n = 16) 3) Instruction + CF group (n = 16)
Instruction (five 1-hour lessons during one week)	1) Pronunciation-focused instruction • Day 1: /i/ & /ɪ/ • Day 2: /ɛ/ & /æ/ • Day 3: /u/ & /ʊ/; /l/ & /r/ • Day 4: /ʃ/, /tʃ/, /ʒ/, & /dʒ/ • Day 5: /θ/ & /ð/; /f/ & /v/ 2) Instructional pairs: only A-(1) ~ (8) 3) Focused tasks: • Pick-a-card games • Bingo games • Fill-in-the-blank exercises 4) Corrective feedback: • Only for the Instruction + CF group during the focused tasks	1) ESL Instructors (n = 3) 2) Instruction-only group (n = 16) 3) Instruction + CF group (n = 16)
Immediate posttest	1) Forced identification tasks (n = 336) • Words: A-(1) ~ (8) and B-(1) ~ (12) + A-(9) ~ (12) for the generalizability test	1) Instruction-only group (n = 16) 2) Instruction + CF group (n = 16)
Delayed posttest (2 weeks later)	1) Forced identification tasks (n = 240) • Words: A-(1) ~ (8) and B-(1) ~ (12)	1) Instruction-only group (n = 16) 2) Instruction + CF group (n = 16)

3.3.2. Pretest

The Korean learners of English took a total of 240 forced identification tasks, and the native speakers of English in the NS Listeners group took a total of 336 forced identification tasks including the generalizability testing items. The pretest was administered in the research laboratory at the Faculty of Education, McGill University, and took 1 hour to complete. Detailed instructions were first explained to participants. In order to avoid confusion about tasks per se, the researcher demonstrated a few simulations, and the participants also conducted trials with distracters. As discussed before, there was no interval time between tasks; hence, if they wanted to take a rest, they were allowed to do so.

3.3.3. Instruction

3.3.3.1. Setting

Considering age, gender, LOA, and %-correct identification scores of the pretest (see Table 8), the 32 Korean learners of English were assigned to either the Instruction-only group or the Instruction + CF group. In order to guarantee active classroom interaction, one class within each group was set to have 8 learners.

Table 8. Participant information by group

Information	Instruction-only (*n* = 16)	Instruction + CF (*n* = 16)
Age	28.50 (*SD* = 5.85)	27.88 (*SD* = 5.26)
Male : Female (%)	25 : 75	25 : 75
LOA (months)	12.63 (*SD* = 13.69)	14.02 (*SD* = 14.02)
Total scores of the pretest (%)	58.24 (*SD* = 11.62)	58.86 (*SD* = 11.21)

Five 1-hour pronunciation lessons were taught at the McGill Faculty of Education by the ESL instructors recruited for the purpose of this study. Before the first class, the researcher had a meeting with the ESL instructors so that they fully understood the aims of this study, what to teach, how to manage lessons (e.g., focused tasks) and, more importantly, how to provide learners with corrective feedback appropriately.

Each lesson was designed for English consonants and vowels with which Korean learners of English have difficulty (see Table 7). More specifically, students were taught about the target vowels /i/ and /ɪ/ on the first day; the target vowels were taught through the instructor's metalinguistic explanations as well as focused tasks that were games and exercises requiring the categorization of the target vowels. Even after the first lesson, the learners were exposed to the target vowels throughout the rest of lessons in that the instructors continued to remind learners of the contrast of target vowels.

To incite learners to focus on the target vowels, all of the lessons were based on form-focused instruction (henceforth, FFI). With its goal of drawing learners' attention to target forms that would otherwise not be noticed, FFI is known to have acquisitional value as an effective teaching technique (e.g., Lyster, 2004a, 2004b; Spada, 1997). Specifically, with respect to phonological development, Saito and Lyster (2012a) found that Japanese English learners partially benefited from FFI in acquiring English /ɹ/. Therefore, to predispose learners to pay more attention to the target contrast, not only corrective feedback but also FFI was employed in this study following Saito and Lyster's (2012a) operationalization of FFI:

(a) Structured input (VanPatten, 2004): While showing pictures or words, instructors provide learners with accurate sounds of the target vowels.

(b) Typographically enhanced input (Han, Park, & Combs, 2008): On teaching aids, each target vowel is emphasized with it underlined and highlighted in bold.

(c) Focused tasks (Ellis, 2001, 2006): During the focused tasks embedded with the categorization of the target vowels, learners are asked to categorize and perceive the target vowels correctly.

Corrective feedback was provided to the Instruction + CF group during three focused tasks, which were conducted every class in order to induce learners to focus on /i/ and /ɪ/. The focused tasks are illustrated as below:

(a) Pick-a-card games: Learners are given word cards on which target sounds are written. For example, 'sheep' is written on the front and 'ship' on the back. The instructor asks each learner to pick up the card and show the right side that corresponds with what the instructor calls out. In the Instruction + CF group, corrective feedback is immediately given to the learner once s/he makes a perceptual error.

(b) Bingo games: Learners are asked to fill out a bingo board (5 × 5 squared boxes) with daily target words. The instructor calls a word, and then learners mark it in order. Once a student completes a horizontal, vertical, or diagonal line, he or she yells 'Bingo' and wins the game. The instructor also makes sure that the learners have marked the right word right after calling. As in the pick-a-card games, corrective feedback is individually provided in cases where learners selected the wrong word.

(c) Fill-in-the-blank exercises: This is a semi-dictation task in which learners are asked to fill out daily focused sounds including target vowels. As in the bingo games, the instructor calls a word and the learners write relevant sounds. Corrective feedback on error is provided only to the Instruction + CF group while the instructor checks if they wrote the sound(s) appropriately.

Regardless of daily topics, all of the target words (i.e., A-(1) ~ (8) in Table 6) were covered during the focused tasks. The Instruction-only group participated in the same tasks but instructors provided no corrective feedback nor any clues that might lead students to believe that they were right or wrong.

3.3.3.2. Feedback sequences

For the Instruction + CF group, if any perceptual error occurred during the focused tasks, the feedback was immediately given to learners individually. Please note that the focused tasks were designed to improve the L2 learners' perceptual accuracy, so corrective feedback was only given on perceptual errors, not on production errors. The feedback sequences used for this classroom-based experiment are schematized in Figure 3.

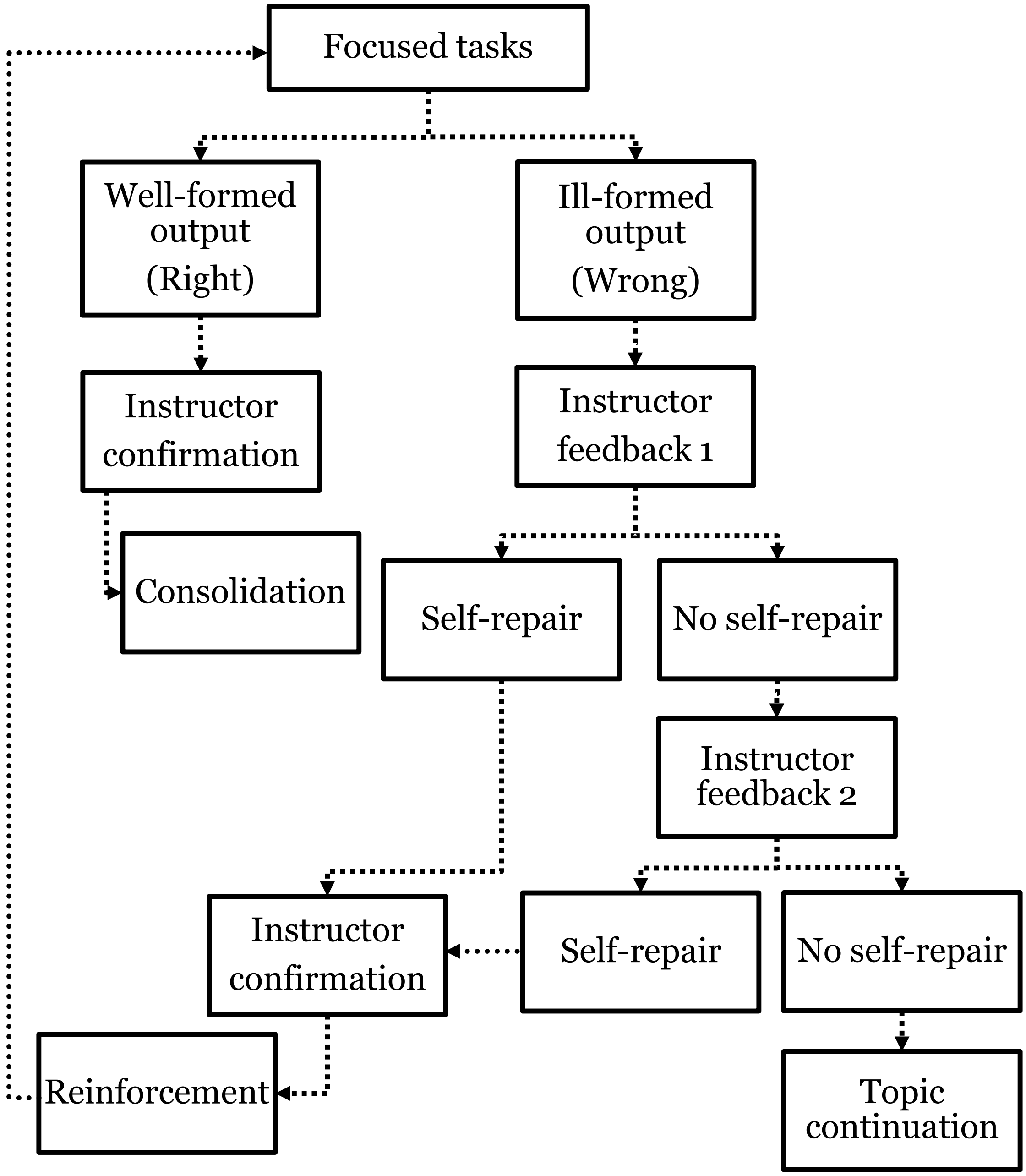

Figure 3. The feedback sequences

If learners responded to the tasks in the right way, the instructor provided them with a confirmation such as 'right, that is (a correct form)', 'perfect', 'good', or 'correct' with falling intonation.

On the other hand, if learners showed any perceptual error, the instructor immediately gave them CF1 at first. CF1 corresponded to a repetition of error. It was designed to draw their attention to the perceptual error in that the instructor repeated verbatim the learners' error — that is, the word selected or written by the learners incorrectly — with rising intonation without any further explanation. If self-repair took place, so as to guarantee that it was not arbitrary, the instructor asked them to pick up or write the relevant word again, which could be either the original word or its counterpart, as illustrated by this excerpt during the pick-a-card games.

Instructor (I): Could you show me 'fill'?

Student (S): (Showing the instructor a card 'feel')

I: Feel? ⇒ CF1

S: Umm. (Showing the other side of the card, 'fill') That one? ⇒ Self-repaired

I: Right, that is 'fill' (with falling intonation). ⇒ Instructor confirmation

S: (Looking at the card 'fill' again) OK.

I: Try again? Show me 'fill'. ⇒ Reinforcement

In spite of CF1, if the student did not self-repair, CF2 was provided: 'I said (the correct word), but you picked up/wrote (the wrong one). Right, then?' CF2 was compatible with explicit correction: 'don't say X, say Y' (Lyster & Ranta, 1997). It was to induce learners to notice their errors

more explicitly and to push them to self-correct. For instance, compared to CF1, it conveyed more elaborated explanation; the instructor produced both sounds as a means of conveying the message 'your choice is not what I said'. In addition, right after further explanation, the instructor provided them with an opportunity to self-correct by saying 'then?'
If the learners responded to this request in the right way, instructor confirmation and reinforcement were provided as in the CF1 sequence. The following illustrates the use of CF2 during the fill-in-the-blank exercises.

Instructor (I): Next, 'bit'.

Student (S): (Writing 'beat' on the worksheet)

I: 'beat'? ⟹ CF1

S: Yes.. Yes.. Yes.. ? ⟹ No self-repair

I: I said 'bit' but you wrote 'beat'. You see? ⟹ CF2

S: Say.. again please.

I: I said 'bit' but you wrote 'beat'. Right, then?

S: Aha, you said 'bit'. Okay. (Re-writing the word) This?
⟹ Self-repair

I: Correct. Could you write 'beat' next to it? ⟹ Instructor
confirmation and Reinforcement

Rarely did the L2 learner participants dismiss CF2. However, when CF2 did not work, topic continuation (e.g., the next word or another focused task) ensued in order to refresh the learners.

3.3.4. Immediate posttest

At the end of the instructional treatment, the 32 non-native participants took an immediate posttest, which included a total of 336 identification tasks. Adding to the pretesting materials, in order to see whether they would be able to transfer acquired knowledge to new words, four new minimal pairs were included as the generalizability test. The tasks were randomly ordered and conducted in the computer laboratory at the Faculty of Education, McGill University. It took from 1 to 1.5 hours to complete. Considering that there are two groups in this study (i.e., Instruction-only vs. Instruction + CF), it is of interest to know whether CF can help learners to be better at unfamiliar items than Instruction-only.

3.3.5. Delayed posttest

Two weeks after the immediate posttest, the 36 non-native speakers completed a delayed posttest at the research laboratory, which took 1 hour to complete. It corresponded to the tasks which they had as the pretest but in a different order. The delayed posttest was to investigate the extent to which L2 learners could retain the knowledge acquired from the instruction.

3.4. Analyses

The scores yielded by this study were based on percentages of correct responses and analyzed statistically. First, to delve into Research Question 1a, %-correct identification scores of the pretest and immediate posttest were compared using a three-way mixed ANOVA conducted with three factors: Group (Instruction-only vs. Instruction + CF) × Time (Pretest vs. Immediate posttest) × Sound type (/i/-natural vs. /i/-synthesized vs. /ɪ/-natural vs. /ɪ/-synthesized). In order to respond to Research Question 1b, %-correct identification scores of the generalizability test were compared using a two-way ANOVA with two factors: Group (Instruction-only vs. Instruction + CF) × Sound type (/i/-natural vs. /i/-synthesized vs. /ɪ/-natural vs. /ɪ/-synthesized).

Second, Research Question 2 was investigated as follows; a three-way mixed ANOVA was conducted with three factors: Group (Instruction-only vs. Instruction + CF) × Time (Pretest vs. Immediate posttest vs. Delayed posttest) × Sound type (/i/-natural vs. /i/-synthesized vs. /ɪ/-natural vs. /ɪ/-synthesized).

Finally, to discuss Research Question 3, %-correct identification scores of each vowel (i.e, /i/ and /ɪ/) for each duration (i.e., 130 ms, 160 ms, 190 ms, 220 ms, and 250 ms) were calculated and compared between the Instruction-only and Instruction + CF groups for the pretest and immediate posttest separately. Specifically, by investigating how the %-correct identification scores of each synthesized vowel were accounted for, it was possible to determine the extent to which the L2 learners relied on durational information and whether their strategy changed after the instructional sessions. Scores of native speakers were also used as a baseline for all comparisons.

3.5. Summary

A total of 21 NS participants (NS speakers, *n* = 2; NS Listeners, *n* = 16; ESL instructors, *n* = 3) and 32 Korean learners of English (Instruction-only, *n* = 16; Instruction + CF, *n* = 16) participated in this study. Twelve sets of /i/-/ɪ/ pairs and another twelve sets of distracters (onset changed pairs, *n* = 6; nucleus changed pairs, *n* = 6) were selected from a high-frequency ESL vocabulary list. All of the testing materials were chosen from CVC monosyllabic words and recorded by one male and one female GA speaker. As for the design of perception test, forced identification tasks were provided to the participants in three test sessions: (a) a pretest, (b) an immediate posttest, and (c) a delayed posttest. The immediate posttest consisted of the pretesting tasks but in a different order and the unfamiliar words that had not appeared during the pretest and instruction. Regarding the delayed posttest, it was comprised of the pretesting tasks but in a different order. Duration-synthesized as well as natural sounds were given to the participants to investigate the extent to which the Korean learners of English relied on durational information.

After the pretest, the Korean learners of English were assigned to either the Instruction-only group or the Instruction + CF group, both of which participated in five 1-hour pronunciation lessons targeting various English consonants and vowels. Specifically, in order to predispose them to pay more attention to the target vowels, FFI was employed. During the focused tasks component of the FFI, corrective feedback was provided to the Instruction + CF group when any perceptual error occurred. Results of the pretest, immediate posttest, and delayed posttest were analyzed with ANOVAs to address the research questions in this study.

CHAPTER 4

RESULTS

In this chapter, four analyses are discussed. First, in order to investigate the effects of instruction and corrective feedback, a comparison of the pretest and immediate posttest was conducted. Second, by measuring how the L2 learners responded to new words containing the target vowels, it was possible to see whether the instructional treatments had an effect beyond familiar items that had been taught during the instructional sessions. Third, with a comparison of the pretest, immediate posttest, and delayed posttest, it was possible to determine whether the effects of the instructional treatments were retained after two weeks. Finally, in terms of the duration-synthesized vowels, two facets were explored: (a) the extent to which the L2 learners relied on durational information when categorizing the target vowels, and (b) whether there was any change after the instructional sessions. While conducting the above four analyses, the guiding question was whether the effects of the treatments — the Instruction-only and Instruction + CF — differed within each analysis.

An alpha level of .05 was set for all statistical analyses in this study. Positive values of mean differences refer to an increase in %-correct identification scores; conversely, negative values refer to a decrease.

In addition, to quantify the effects of instructional treatments, Cohen's *d* (1988) was also calculated and classified as small (.20 ≤ *d* < .50), medium (.50 ≤ *d* < .80), or large (.80 ≤ *d*) effect sizes.

4.1. Pretest–immediate posttest comparison

To assess the effects of instruction and corrective feedback on the target vowels, a three-way mixed ANOVA was conducted with three factors: Group (Instruction-only vs. Instruction + CF) × Time (Pretest vs. Immediate posttest) × Sound type (/i/-natural vs. /i/-synthesized vs. /ɪ/-natural vs. /ɪ/-synthesized).

The means and standard deviations for the pretest and immediate posttest by group and by sound type appear in Table 9 and the means are displayed graphically in Figure 4 for the Instruction + CF group and in Figure 5 for the Instruction-only group.

Table 9. Means and standard deviations of pretest and immediate posttest scores by group and sound type

	Pretest		**Immediate posttest**	
	Instruction-only (*n* = 16)	*Instruction + CF* (*n* = 16)	*Instruction-only* (*n* = 16)	*Instruction + CF* (*n* = 16)
/i/-natural	68.359 (*SD* = 14.696)	72.656 (*SD* = 15.117)	85.156 (*SD* = 13.283)	84.062 (*SD* = 12.898)
/i/-synthesized	69.609 (*SD* = 16.569)	70.703 (*SD* = 13.632)	82.734 (*SD* = 14.048)	84.375 (*SD* = 11.474)
/ɪ/-natural	57.812 (*SD* = 15.729)	62.890 (*SD* = 16.373)	52.734 (*SD* = 23.382)	78.906 (*SD* = 12.263)
/ɪ/-synthesized	43.750 (*SD* = 19.223)	44.921 (*SD* = 11.157)	46.172 (*SD* = 21.117)	62.890 (*SD* = 20.401)

There was a significant three-way Group × Time × Sound type interaction, $F(3, 90) = 5.855$, $p = .001$. For the Instruction + CF group, after the instructional sessions, their perceptual accuracy increased for all sound types: (a) /i/-natural, $M_{diff} = 11.406$, $F(1, 240) = 4.014$, $p = .046$, $d = .81$; (b) /i/-synthesized, $M_{diff} = 13.672$, $F(1, 240) = 5.771$, $p = .017$, $d = 1.09$; (c) /ɪ/-natural, $M_{diff} = 16.016$, $F(1, 240) = 7.920$, $p = .005$, $d = 1.11$; and (d) /ɪ/-synthesized, $M_{diff} = 17.969$, $F(1, 240) = 9.969$, $p = .002$, $d = 1.09$. However, concerning the Instruction-only group, a simple main effect of Time was significant only for the /i/-natural and /i/-synthesized sounds: (a) /i/-natural, $M_{diff} = 16.797$, $F(1, 240) = 8.711$, $p = .003$, $d = 1.20$; and (b) /i/-synthesized, $M_{diff} = 13.125$, $F(1, 240) = 5.319$, $p = .022$, $d = .85$.

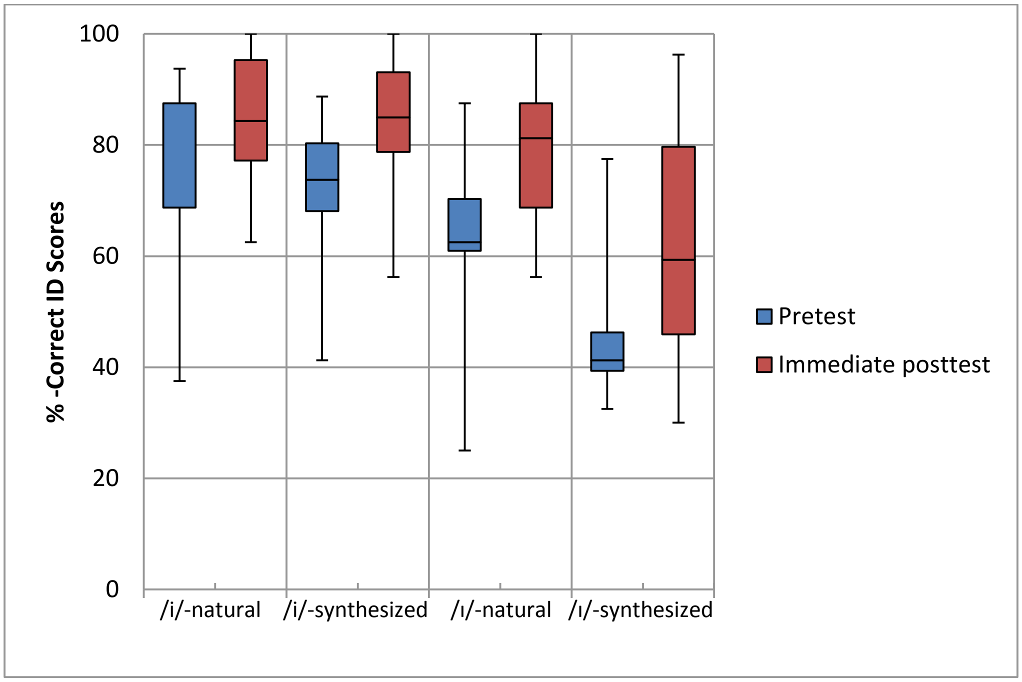

Figure 4. Means of pretest and immediate posttest scores of the Instruction + CF group for each sound type

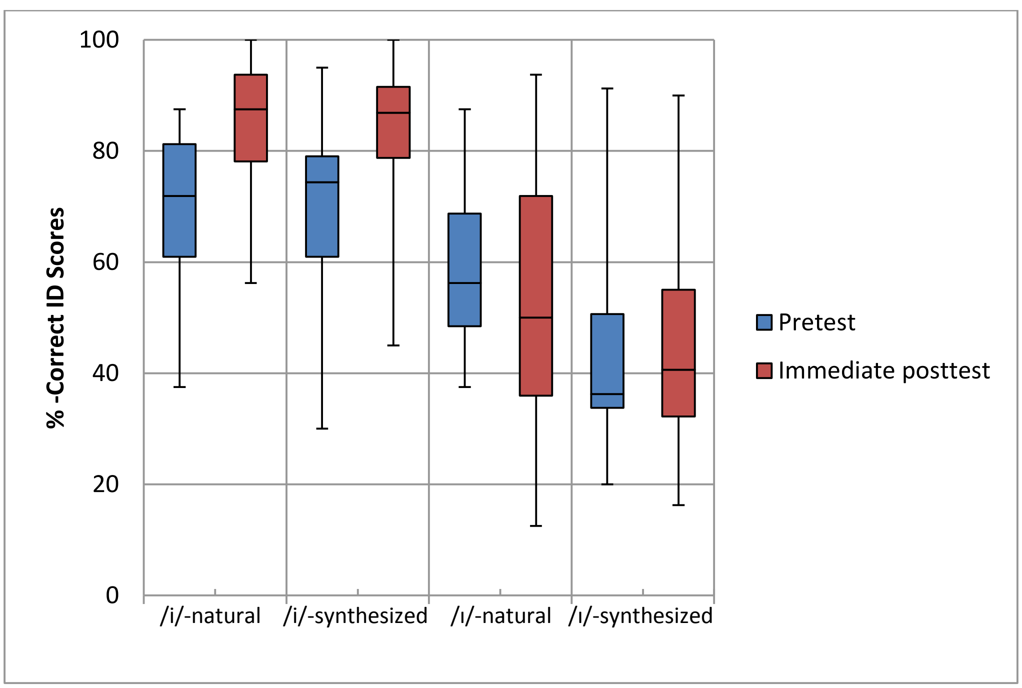

Figure 5. Means of pretest and immediate posttest scores of the Instruction-only group for each sound type

In the same vein, both groups showed similar perceptual accuracy across all sound types at the pretest: (a) /i/-natural, M_{diff} = 4.297, $F(1, 240) = .570$, $p = .451$; (b) /i/-synthesized, M_{diff} = 1.094, $F(1, 240) = .037$, $p = .848$; (c) /ɪ/-natural, M_{diff} = 5.078, $F(1, 240) = .796$, $p = .373$; and (d) /ɪ/-synthesized, M_{diff} = 1.171, $F(1, 240) = .042$, $p = .837$. However, a simple main effect of Group was significant for the /ɪ/-natural and /ɪ/-synthesized sounds at the immediate posttest wherein the Instruction + CF group outperformed the Instruction-only group in the /ɪ/-natural sounds, M_{diff} = 26.172, $F(1, 240) = 21.149$, $p < .001$, $d = 1.45$, and the /ɪ/-synthesized sounds, M_{diff} = 16.718, $F(1, 240) = 8.630$, $p = .004$, $d = .83$. Moreover, the main effect of Sound type on the perceptual accuracy was significant, $F(3, 90) = 52.841$, $p < .001$. Tukey HSD post hoc tests indicated that there was no statistical difference between the /i/-natural and /i/-synthesized sounds ($p = .996$), whereas all other comparisons were significant ($p < .001$) as follows: /i/-natural ($M = 77.558$), /i/-synthesized ($M = 76.855$) > /ɪ/-natural ($M = 63.086$) > /ɪ/-synthesized sounds ($M = 49.433$).

As expected, the native speakers of English showed almost perfect perceptual accuracy on all sound types: (a) /i/-natural, $M = 100$; (b) /i/-synthesized, $M = 99.375$; (c) /ɪ/-natural, $M = 100$; and (d) /ɪ/-synthesized, $M = 99.766$. There was no statistically significant difference between the sound types, $F(3, 60) = 2.382$, $p = .078$.

In sum, both Instruction-only and Instruction + CF groups demonstrated similar perceptual accuracy at the pretest. However, after the instructional sessions, the Instruction + CF group showed an increase in perceptual accuracy vis-à-vis all sound types while the Instruction-only group showed the improvement on the /i/-natural and /i/-synthesized sounds only. Specifically, note that there were large between-group effects for the /ɪ/-natural and /ɪ/-synthesized sounds at the immediate posttest (i.e., d = 1.45 and .83, respectively). Concerning the baseline data, the native speakers of English showed almost perfect perceptual accuracy regardless of sound type.

4.2. Generalizability test

In order to investigate the extent to which the L2 learners could perceive the target vowels in words that had not been covered during the instructional sessions, %-correct identification scores of the generalizability test were compared between the Instruction-only and Instruction + CF groups, using a two-way ANOVA with two factors: Group (Instruction-only vs. Instruction + CF) × Sound type (/i/-natural vs. /i/-synthesized vs. /ɪ/-natural vs. /ɪ/-synthesized).

The means and standard deviations for the generalizability test by group and by sound type are presented in Table 10 and the means are illustrated graphically in Figure 6 for both groups.

Table 10. Means and standard deviations of generalizability test scores by group and sound type

	Instruction-only (n = 16)	Instruction + CF (n = 16)
/i/-natural	68.125 (*SD* = 14.186)	75.781 (*SD* = 19.077)
/i/-synthesized	62.500 (*SD* = 17.512)	79.063 (*SD* = 20.308)
/ɪ/-natural	51.406 (*SD* = 17.981)	76.094 (*SD* = 19.230)
/ɪ/-synthesized	37.656 (*SD* = 9.594)	53.438 (*SD* = 27.536)

There was no significant interaction, $F(3, 120) = 1.095$, $p = .354$. However, the ANOVA revealed a main effect for Sound type, $F(3, 120) = 13.443$, $p < .001$ and more importantly, a main effect for Group, $F(1, 120) = 23.655$, $p < .001$. The Instruction + CF group was overall better than the Instruction-only group at the generalizability test, $M_{diff} = 16.172$.

As for the main effect of Sound type, Tukey HSD post hoc tests revealed that results for the /i/-natural, /i/-synthesized, and /ɪ/-natural sounds were not significantly different ($p > .05$) while the /ɪ/-synthesized sounds were statistically different from the rest of the sound types ($p < .001$) such that it showed the following pattern: /i/-natural ($M = 71.953$), /i/-synthesized ($M = 70.782$), /ɪ/-natural ($M = 63.750$) > /ɪ/-synthesized sounds ($M = 45.547$).

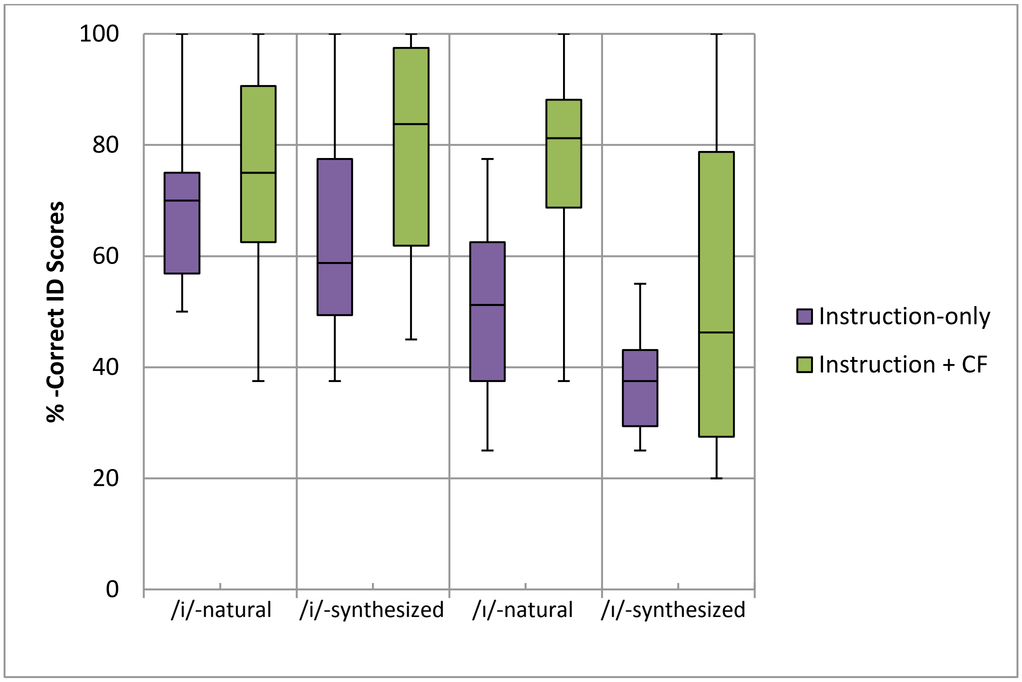

Figure 6. Means of generalizability test scores of the Instruction-only and Instruction + CF groups for each sound type

Effect size analyses also indicated small-to-large effects between the Instruction-only and Instruction + CF groups: d = .47 for the /i/-natural, d = .90 for the /i/-synthesized, d = 1.37 for the /ɪ/-natural, and d = .79 for the /ɪ/-synthesized sounds.

Again, the native speakers of English seldom made perceptual errors. They showed over 99.43% accuracy with no significant differences across all sound types, $F(3, 60) = 1.866$, $p = .145$.

In a nutshell, the Instruction + CF group overall outperformed the Instruction-only group at the generalizability test. Considering that both groups showed similar perceptual accuracy on the target vowels at the pretest, it was interesting to find that the Instruction + CF group was better than the Instruction-only group at the generalizability test. As in the pretest-immediate posttest comparison, the accuracy of /i/ was higher than that of /ɪ/. The native speakers of English again made few errors in the generalizability tasks.

4.3. Delayed posttest

So as to investigate whether the L2 learners retained the effects of instructional treatments after the instructional sessions, a delayed posttest was administered two weeks after the treatment and results were analyzed by a three-way mixed ANOVA with three factors: Group (Instruction-only vs. Instruction + CF) × Time (Pretest vs. Immediate posttest vs. Delayed posttest) × Sound type (/i/-natural vs. /i/-synthesized vs. /ɪ/-natural vs. /ɪ/-synthesized).

The means and standard deviations for the pretest, immediate posttest, and delayed posttest by group and by sound type are shown in Table 11 and the means are displayed graphically in Figure 7 for the Instruction + CF group and in Figure 8 for the Instruction-only group.

Table 11. Means and standard deviations of pretest, immediate posttest, and delayed posttest scores by group and sound type

	Instruction-only (*n* = 16)			**Instruction + CF (*n* = 16)**		
	Pretest	*Immediate posttest*	*Delayed posttest*	*Pretest*	*Immediate posttest*	*Delayed posttest*
/i/-natural	68.359 (*SD* = 14.696)	85.156 (*SD* = 13.283)	81.328 (*SD* = 11.287)	72.656 (*SD* = 15.117)	84.062 (*SD* = 12.898)	91.406 (*SD* = 10.674)
/i/-synthesized	69.609 (*SD* = 16.569)	82.734 (*SD* = 14.048)	80.625 (*SD* = 5.164)	70.703 (*SD* = 13.632)	84.375 (*SD* = 11.474)	82.969 (*SD* = 14.200)
/ɪ/-natural	57.812 (*SD* = 15.729)	52.734 (*SD* = 23.382)	57.812 (*SD* = 20.854)	62.890 (*SD* = 16.373)	78.906 (*SD* = 12.263)	76.922 (*SD* = 15.925)
/ɪ/-synthesized	43.750 (*SD* = 19.223)	46.172 (*SD* = 21.117)	50.937 (*SD* = 18.655)	44.921 (*SD* = 11.157)	62.890 (*SD* = 20.401)	69.296 (*SD* = 14.888)

As in the pretest-immediate posttest comparison, there was a significant three-way Group × Time × Sound type interaction, $F(6, 180) = 3.236$, $p = .005$. A simple main effect of Time was significant for the Instruction + CF group for all sound types, (a) /i/-natural, $F(2, 360) = 5.836$, $p = .003$; (b) /i/-synthesized, $F(2, 360) = 3.698$, $p = .026$; (c) /ɪ/-natural, $F(2, 360) = 4.983$, $p = .007$; and (d) /ɪ/-synthesized, $F(2, 360) = 10.440$, $p < .001$. Specifically, differences proved to be significant between the pretest and immediate posttest ($M_{diff} = 11.406$, 13.672, 16.016, and 17.969 for (a), (b), (c), and (d) respectively), and between the pretest and delayed posttest ($M_{diff} = 18.750$, 12.266, 14.032, and 24.375 each). However, they were not significant between the immediate posttest and delayed posttest ($M_{diff} = 7.344$, -1.406, -1.984, and 6.406 respectively). A simple main effect of Time was also significant for the Instruction-only group, but only for the /i/-natural and /i/-synthesized sounds: (a) /i/-natural, $F(2, 360) = 5.067$, $p = .007$; and (b) /i/-synthesized, $F(2, 360) = 3.248$, $p = .040$. As in the Instruction + CF group, the differences were significant between the pretest and immediate posttest ($M_{diff} = 16.797$ and 13.125, respectively), and between the pretest and delayed posttest ($M_{diff} = 12.969$ and 11.016, each), but not between the immediate posttest and delayed posttest ($M_{diff} = -3.828$ and -2.109, respectively).

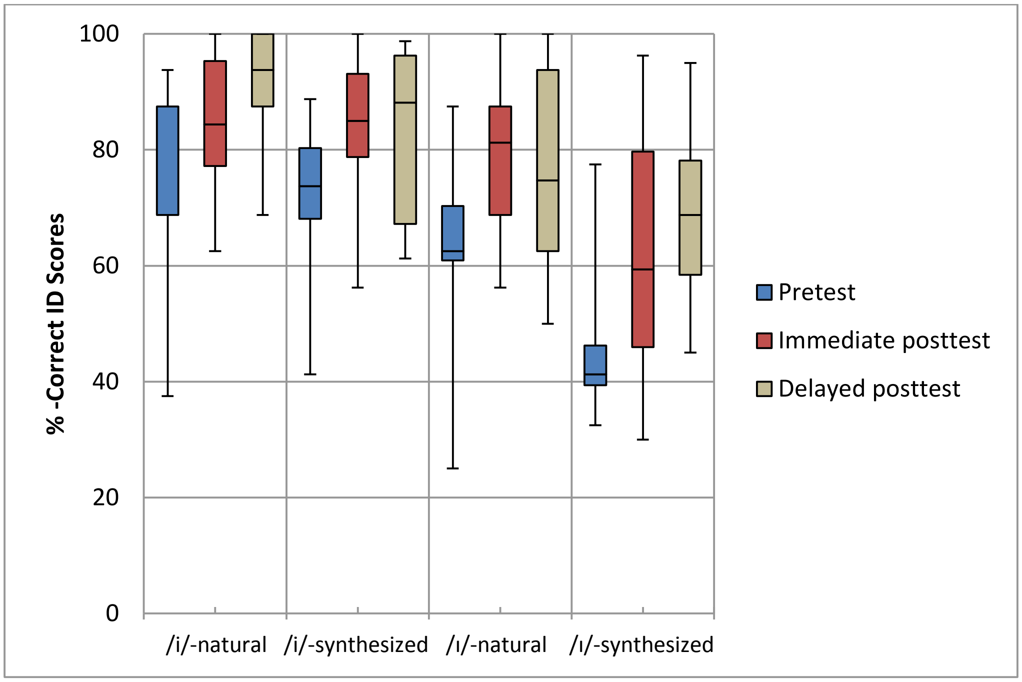

Figure 7. Means of pretest, immediate posttest, and delayed posttest scores of the Instruction + CF group for each sound type

In the comparison of the pretest and immediate posttest, as discussed before, both groups showed similar perceptual accuracy across all sound types at the pretest; the ANOVA revealed a significant main effect for Group for the /ɪ/-natural and /ɪ/-synthesized sounds at the immediate posttest, with the Instruction + CF group outperforming the Instruction-only group. A simple main effect of Group was also found at the delayed posttest — there were significant differences observed only at the /ɪ/-natural and /ɪ/-synthesized sounds. For both, the Instruction + CF group was better than the Instruction-only group and effect size analyses also indicated large effects between two groups:

(a) /ɪ/-natural, M_{diff} = 19.110, F(1, 360) = 11.937, p = .001, d = 1.06; and (b) /ɪ/-synthesized, M_{diff} = 18.359, F(1, 360) = 11.019, p = .001, d = 1.12.

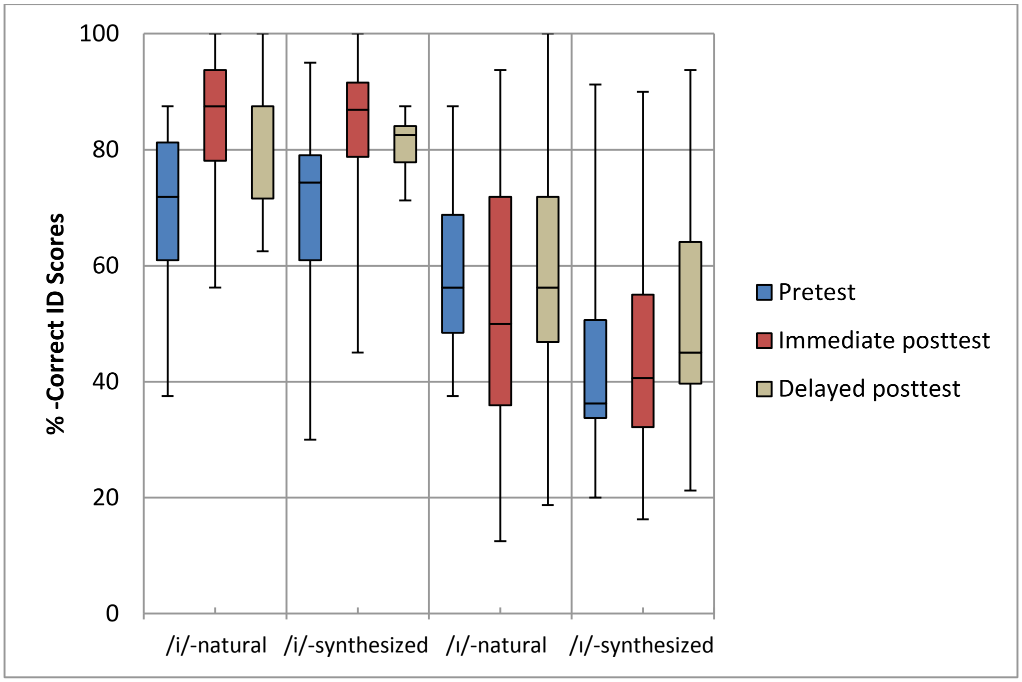

Figure 8. Means of pretest, immediate posttest, and delayed posttest scores of the Instruction-only group for each sound type

A significant effect for Sound type was also found in this analysis, F(3, 90) = 65.371, p < .001. Tukey HSD post hoc tests pointed out no difference between the /i/-natural and /i/-synthesized sounds, whereas all of the other comparisons were significant (p < .05) as: /i/-natural (M = 80.495), /i/-synthesized (M = 78.503) > /ɪ/-natural (M = 64.513) > /ɪ/-synthesized sounds (M = 52.994).

In sum, both Instruction-only and Instruction + CF groups retained their instructional treatment effects at the delayed posttest.

The ANOVA analyses confirmed that %-correct identification scores of the immediate posttest and delayed posttest were not statistically different from each other, which were statistically higher than those of the pretest.

4.4. The use of spectral vs. durational information

It is well known that Korean learners of English make extensive use of durational information when it comes to categorizing /i/ and /ɪ/ (Flege et al., 1997). Taking this fact into account, this study attempted to delve into whether instruction and corrective feedback could induce them to focus on spectral information. To tease apart this question, first of all, it was calculated how the %-correct identification scores of each synthesized vowel were accounted for. On the one hand, with respect to the /i/-synthesized sounds, the native speakers of English showed almost perfect accuracy (M = 99.375). Specifically, they showed 98.828, 99.219, 99.609, 99.609, and 99.609 for the /i/-130 ms, /i/-160 ms, /i/-190 ms, /i/-220 ms, and /i/-250 ms, respectively. A similar result was obtained with the /ɪ/-synthesized sounds as well (M = 99.766): 100, 99.609, 99.609, 100, and 99.609 for the /ɪ/-130 ms, /ɪ/-160 ms, /ɪ/-190 ms, /ɪ/-220 ms, and /ɪ/-250 ms, respectively. In other words, not only did native speakers of English show almost perfect accuracy on the synthesized vowels but they also showed little reliance on durational information while categorizing the target vowels.

On the other hand, as expected, the responses of the Korean learners of English varied according to durational information. For example, both Instruction-only and Instruction + CF groups showed statistically similar %-correct identification scores on the /i/-synthesized sounds (i.e., *M* = 69.609 and 70.703 each) at the pretest. Moreover, the longer the vowel was, the higher their accuracy was: (a) The Instruction-only group scored 52.734, 63.281, 72.656, 76.953, and 82.422 and (b) the Instruction + CF group scored 49.609, 63.672, 76.172, 78.906, and 85.156 for the /i/-130 ms, /i/-160 ms, /i/-190 ms, /i/-220 ms, and /i/-250 ms, respectively.

Conversely, regarding the /ɪ/-synthesized sounds (*M* = 43.750 and 44.921 for the Instruction-only and Instruction + CF groups, respectively), the shorter the vowel was, the higher their accuracy was: (a) The Instruction-only group scored 59.375, 50.000, 42.969, 40.625, and 25.781 and (b) the Instruction + CF group scored 68.750, 54.686, 43.745, 29.297, and 28.125 for the /ɪ/-130 ms, /ɪ/-160 ms, /ɪ/-190 ms, /ɪ/-220 ms, and /ɪ/-250 ms, respectively.

Based on the above results, Figures 9 and 10 were created to illustrate the extent to which the L2 learners relied on durational information at the pretest and immediate posttest compared to the native speakers of English. For example, as mentioned before, the NS accuracy of /i/-synthesized sounds was 98.828, 99.219, 99.609, 99.609, and 99.609 for the /i/-130 ms, /i/-160 ms, /i/-190 ms, /i/-220 ms, and

/i/-250 ms, respectively. Each value was marked on the axis of corresponding duration (e.g., 98.828 on the axis of 130 ms in Figure 9), and then these values were connected with solid black lines to represent the native baseline data. Following the same procedure, the pretest results of the L2 learners are illustrated with blue dots and their immediate posttest results with red dots. In addition, if the learners' accuracy at the immediate posttest became statistically different from the pretest, this is indicated by orange arrows. These figures thus serve to illustrate the extent to which the Korean learners of English decreased their use of durational information and became more similar to native speakers after the instructional sessions.

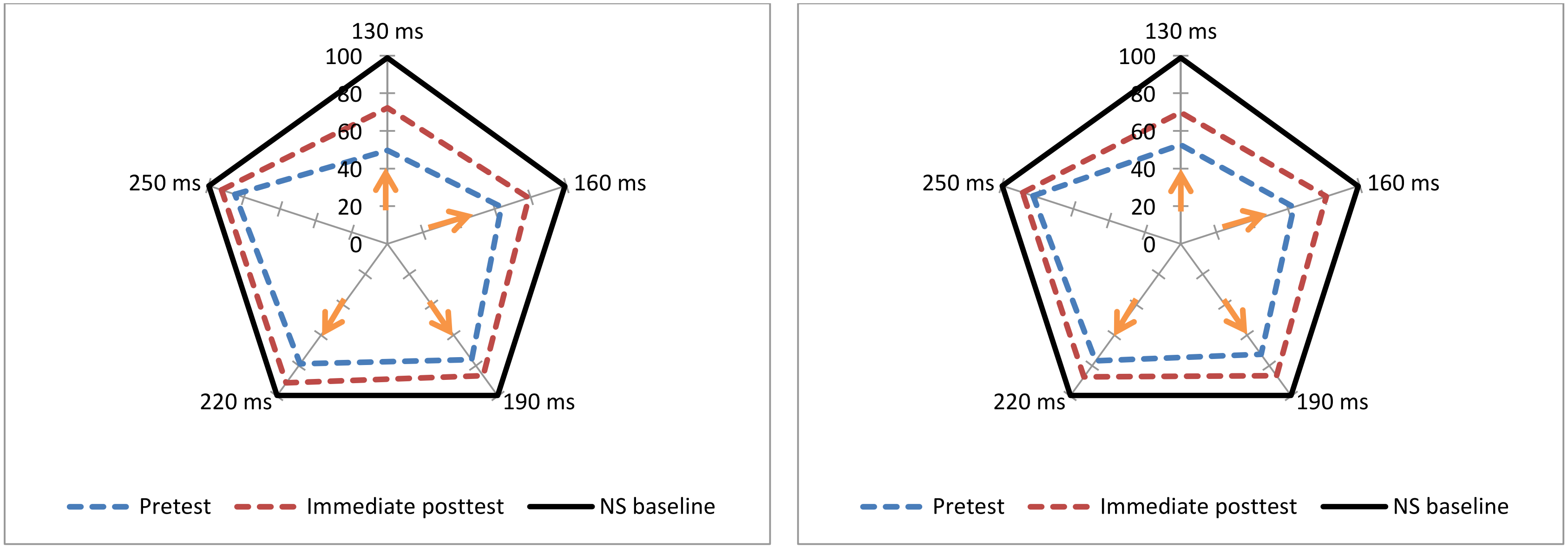

Figure 9. The usage of durational information and effects of instruction (and corrective feedback) on the /i/-synthesized sounds (left: the Instruction + CF group; right: the Instruction-only group)

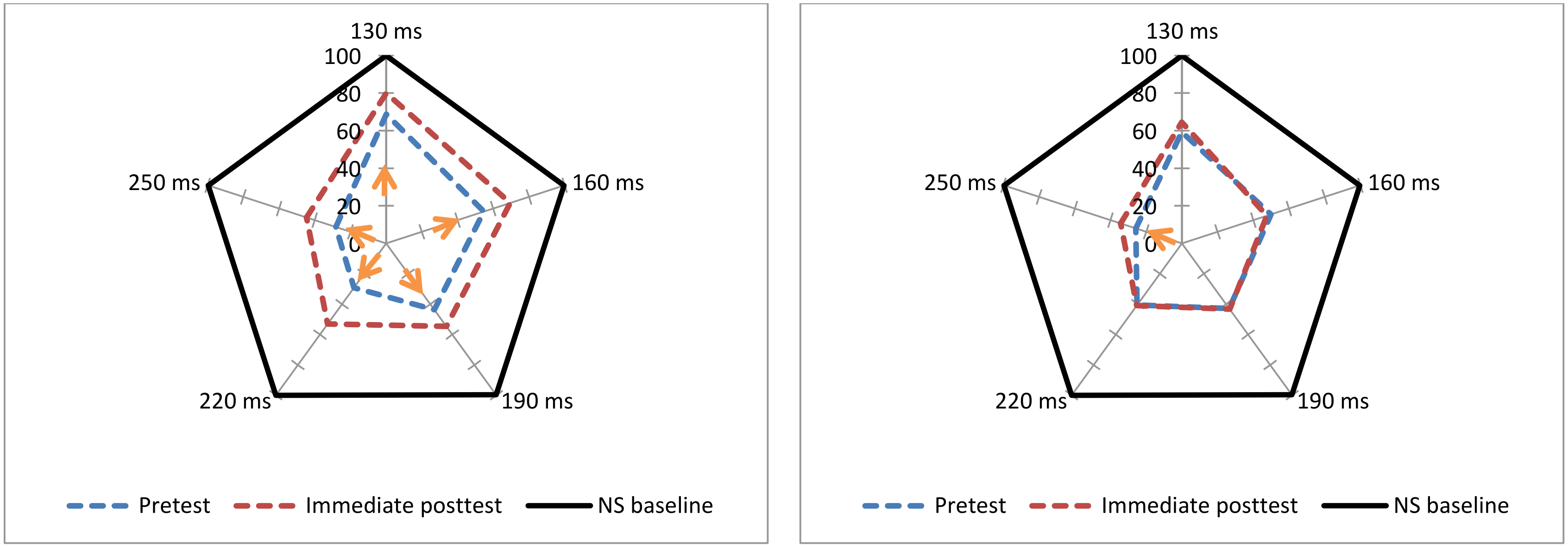

Figure 10. The usage of durational information and effects of instruction (and corrective feedback) on the /ɪ/-synthesized sounds (left: the Instruction + CF group; right: the Instruction-only group)

As Figures 9 and 10 indicate, the scores of the Korean learners of English were more affected by durational information in the case of /ɪ/-synthesized sounds. Despite the fact that both groups still relied on durational information after the instructional treatment, they started to pay more attention to spectral information. For example, in terms of the /i/-synthesized sounds, both the Instruction-only and the Instruction + CF groups demonstrate the same degree of positive change towards the NS baseline; however, with respect to the /ɪ/-synthesized sounds, the Instruction + CF group demonstrates more positive change than its counterpart.

To sum up, the following results were (re)confirmed through this analysis. First, the Korean learners of English relied on durational information when categorizing the /i/ and /ɪ/ (notably more with the /ɪ/), whereas the native speakers of English were not confused about the duration-synthesized sounds at all. Second, after the instruction, the L2 learners started to focus on spectral information; specifically, for the /i/-synthesized sounds, both groups showed similar improvement (see Figure 9), and for the /ɪ/-synthesized sounds, the Instruction + CF group was better than the Instruction-only group (see Figure 10).

4.5. Summary

In terms of categorizing and perceiving /i/ and /ɪ/, both Instruction-only and Instruction + CF groups benefited from its respective instructional treatment. However, the statistical analyses revealed differences between those groups. First, the Instruction + CF group showed an increase on all sound types at the immediate posttest, whereas the Instruction-only group showed improvement only on the /i/-natural and /i/-synthesized sounds at the immediate posttest. Second, the Instruction + CF group outperformed the Instruction-only group at the generalizability test, being better at categorizing the target vowels in words that had not been taught during the instructional sessions. Third, the above aspects were reconfirmed at the two-week delayed posttest. Interesting enough, there was one main effect that was consistently observed across all of the analyses in this study: the main effect of Sound type. The post hoc tests consistently indicated the following pattern: /i/-natural, /i/-synthesized > /ɪ/-natural > /ɪ/-synthesized sounds (except for the generalizability test, /i/-natural, /i/-synthesized, /ɪ/-natural > /ɪ/-synthesized sounds). Finally, the instruction and corrective feedback helped the Korean learners of English to pay more attention to spectral information, whereas they tend otherwise to use durational information when categorizing /i/ and /ɪ/ (Flege et al., 1997). Among the target sounds, the L2 learners seemed to be more affected by duration differences in the case of the /ɪ/-synthesized sounds. However, there was also a difference between two groups — compared to the /i/-synthesized sounds, the Instruction + CF group showed more positive improvement on the /ɪ/-synthesized sounds than the Instruction-only group.

CHAPTER 5

DISCUSSION

In this chapter, the perceptual default of Korean learners of English is at first predicted based on the results in this study. Taking their perceptual preference into account, I propose a framework of (mis)match-based perceptual decisions by Korean learners of English. Specifically, the most exciting finding in this study — that the Instruction + CF group showed improvement on all of the sound types while the Instruction-only group showed an increase only on the /i/-natural and /i/-synthesized sounds — is discussed within the proposed framework. Also addressed within the framework are the effects of instructional treatments including corrective feedback and focused tasks, which help L2 learners to build and test linguistic hypotheses and to generate new hypotheses if wrong. In addition, given the current results and framework, ultimate attainment of the target vowels by the Korean learners of English is also expected within the framework. Finally, while shedding light on the roles of corrective feedback and focused tasks during which the corrective feedback naturally took place, the relationship between focused tasks and corrective feedback is expounded.

5.1. Perceptual default

One of the key findings in this research is that the ANOVA indicated the main effect of Sound type across the pretest, immediate posttest, and delayed posttest with the following order: /i/-natural, /i/-synthesized > /ɪ/-natural > /ɪ/-synthesized sounds. In other words, the accuracy of /i/ was always higher than that of /ɪ/. In addition, given that there is a statistical difference between the /ɪ/-natural and /ɪ/-synthesized sounds but not between the /i/-natural and /i/-synthesized sounds, it may be the case that the duration manipulation hampered the Korean learners of English from categorizing the /ɪ/-synthesized sounds as much as the /ɪ/-natural sounds. These results allow us to confirm their perceptual preference with the target vowels. According to Table 12, when auditory input /i/ was given to the participants, their accuracy was approximately 71%; in other words, 29% of the auditory input /i/ was perceived as /ɪ/. Intriguingly enough, when the auditory input /ɪ/ was provided to the participants, the probability of it being perceived as /i/ was around 49%.

Table 12. Perceptual preference by Korean learners of English in the pretest

Auditory input / **Perceptual representation**	**/i/**	**/ɪ/**	**Perceptual preference**
/i/	71%	49%	60%
/ɪ/	29%	51%	40%

Simply put, regardless of the auditory input /i/ and /ɪ/, the Korean learners of English showed a tendency to perceive the auditory input as /i/. For instance, given the results of the pretest, 60% of the total auditory input was perceived as /i/ although the same ratio (50:50) of the auditory input /i/ and /ɪ/ was provided to the participants. In this regard, the participants exhibited a preference for /i/ as the perceptual representation.

In sum, /i/ is set to be the default in the perceptual state of the Korean learners of English. This speculation is also supported by the speech perception models. The PAM (Best, 1995) predicts that two non-native phonemes /i/ and /ɪ/ will be assimilated into the native phoneme /i/. Specifically, the PAM expects that Korean learners of English are likely to show a pattern of either single-category or category-goodness. In the present study, they tended to show a pattern of category-goodness, given their higher accuracy on the target phoneme /i/ than on /ɪ/. In other words, the fact that the L2 learners showed less accuracy on /ɪ/ could imply the target phoneme /ɪ/ was frequently perceived as /i/, considering that there were only two options /i/ and /ɪ/ available. Moreover, the NRV (Polka & Bohn, 2011) predicts that /i/ is the natural referent vowel because of its acoustic properties and perceptual salience. The model also predicts that the natural reference vowel becomes more prominent when L2 learners acquire a new vowel system.

In addition, the fact that the participants used durational information distinctively when they were exposed to /ɪ/ also suggests that /i/ is the default in their perceptual state. Because /i/ is prioritized in their perceptual state, they need a strong impetus to categorize the auditory input /ɪ/ as /ɪ/, which is of shorter duration. Because the length of /ɪ/ is acoustically shorter than that of /i/ (Hillenbrand et al., 1995) and because Korean learners of English are incognizant of the spectral information of /ɪ/ relative to /i/ which is prioritized in their perceptual state, they notice an acoustic cue of short duration while processing relevant input and struggling to categorize /ɪ/ correctly. They thus rely more on durational information when /ɪ/ is provided. In sum, based on the current results and theoretical support, /i/ seems to be the default in the perceptual state of Korean learners of English.

5.2. Perceptual decisions

Table 13 schematizes the framework of (mis)match-based perceptual decisions by Korean learners of English. As discussed above, it was found that the perceptual representation /i/ is frequently prioritized regardless of the auditory input /i/ and /ɪ/. In terms of the target vowels, /i/ is in turn presupposed to be the default in their perceptual state and the target phonemes /i/ and /ɪ/ consequently share the default /i/.

In order to perceive a non-native phoneme, L2 learners start to match the L2 phoneme with a default existing in their L1 inventory. For example, English /i/ has a corresponding default /i/ in the Korean vowel inventory. Similarly, English /e/ has a corresponding default, Korean /e/. Because English /i/ and /e/ do not share a default vowel, L2 learners match each L2 phoneme with its corresponding default; there is no perceptual conflict occurring when Korean learners of English are asked to distinguish English /i/ and English /e/ between English /i/ and /e/. They are thus supposed to show perfect discrimination. However, when more than one L2 auditory input shares one default (e.g., /i/ and /ɪ/), a perceptual conflict occurs.

Table 13. (Mis)match-based perceptual decisions by Korean learners of English

<table>
<tr><th>Default
(A)</th><th>Auditory input
(B)</th><th>Perception
(if perceiving as…)</th><th>Perceptual representation</th><th>Decision</th></tr>
<tr><td rowspan="4">/i/</td><td rowspan="2">/i/</td><td>(A) = (B)
∴ Match</td><td>/i/</td><td>Right</td></tr>
<tr><td>(A) ≠ (B)
∴ Illusion mismatch</td><td>/ɪ/</td><td>Wrong</td></tr>
<tr><td rowspan="2">/ɪ/</td><td>(A) = (B)
∴ Illusion match</td><td>/i/</td><td>Wrong</td></tr>
<tr><td>(A) ≠ (B)
∴ Mismatch</td><td>/ɪ/</td><td>Right</td></tr>
</table>

For example, as illustrated in Table 13, English /i/ and /ɪ/ share one default /i/. If English /i/ is provided as auditory input, their perceptual system determines whether the auditory input is matched with the default /i/. On the one hand, if the perceptual system declares 'match' status between the auditory input /i/ and the default /i/, the system registers /i/ as the perceptual representation. In this case, the perceptual system results in the right representation. On the other hand, although the auditory input /i/ and the default /i/ should be matched, if the perceptual system declares 'mismatch' status (so-called 'illusion mismatch' status in this study) between the auditory input and the default, the system registers the other potential representation sharing the default. For instance, because /i/ and /ɪ/ share one default, /ɪ/ is in turn selected as the representation. In this case, the system results in the wrong representation.

However, in terms of /ɪ/ as auditory input, the scenario becomes completely different from the above. Once the system confirms 'match' status between the auditory input /ɪ/ and the default /i/ (that is, 'illusion match' status), the system incorrectly registers /i/ as the representation. The decision turns out to be wrong. On the other hand, if the system declares 'mismatch' status between the auditory input /ɪ/ and the default /i/, it manages to register the other representation sharing the default, that is, /ɪ/. It becomes the right decision.

As discussed before, it was found that /i/ is frequently prioritized as the perceptual representation regardless of the auditory input such that it is predicted that their perceptual system is less likely to declare overall 'mismatch' status between the auditory input and the default. It is also compatible with the fact that the Korean learners of English showed less accuracy on /ɪ/ compared to /i/. That is because the system should declare 'mismatch' status in order to classify the auditory input /ɪ/ as the perceptual representation /ɪ/; however, the system is less likely to do so.

5.3. Linguistic hypotheses and CF

5.3.1. Linguistic hypotheses and CF with /ɪ/

After the five 1-hour instructional sessions, the most exiting finding is that the Instruction + CF group showed an increase in accurately perceiving both /i/ and /ɪ/, whereas the Instruction-only group showed improvement only on /i/. It allows us to conclude that corrective feedback played a key role in helping the L2 learners to categorize the auditory input /ɪ/ correctly. If so, how?

Based on the framework described earlier and the results of the current study, Table 14 illustrates the results of the pretest, immediate posttest, and generalizability test within the framework of (mis)match-based perceptual decisions.

Table 14. A summary of results in the framework of (mis)match-based perceptual decisions (a = Instruction-only group; b = Instruction + CF group)

Default (A)	**Auditory input (B)**	**Perception (if perceiving as…)**	**Perceptual representation**	**Decision**	**Pretest (a : b)**	**Feedback treatment**	**Immediate posttest (a : b)**	**Generalizability test (a : b)**
/i/	/i/	(A) = (B) ∴ Match	/i/	Right	69% : 72%		84% : 84%	65% : 78%
		(A) ≠ (B) ∴ Illusion mismatch	/ɪ/	Wrong	31% : 28%	√	16% : 16%	35% : 22%
	/ɪ/	(A) = (B) ∴ Illusion match	/i/	Wrong	49% : 47%	√	51% : 29%	56% : 36%
		(A) ≠ (B) ∴ Mismatch	/ɪ/	Right	51% : 53%		49% : 71%	44% : 64%

As mentioned before, 'match' status is preferred in these learners' perceptual system. For example, during the pretest, if /i/ was provided as auditory input, the Korean learners of English were highly likely to select /i/ as its perceptual representation, which is the right decision. Regarding /i/, the perceptual accuracy thus became higher. However, the match-biased perceptual state impedes them from showing higher accuracy on /ɪ/; the learners were more likely to choose /i/ in response to the auditory input /ɪ/. In other words, in order to attain higher accuracy on /ɪ/ within this framework, their perceptual system needs to declare 'mismatch' status with the default appropriately. In a nutshell, based on this framework, the fact that the Instruction + CF group showed improvement on /ɪ/ after the instructional sessions suggests that the Instruction + CF group declared more 'mismatch' status between the default and the auditory input /ɪ/, whereas the Instruction-only group was less likely to do so.

In order to declare 'mismatch' status, L2 learners at first have to acknowledge that what they hear might be something else beyond what they've currently perceived. For this, the L2 learners need to notice the existence of two different phonemes /i/ and /ɪ/ in the target language. In order to build accurate linguistic information regarding /i/ and /ɪ/, metalinguistic information and input were conveyed with instructional support as illustrated below:

Instructor: (By showing the word card ‘ship’) This is ‘ship’.
You know a boat? (By showing the other card ‘sheep’). This is ‘sheep’, an animal. Listen again (by showing relevant cards), ‘ship’ and ‘sheep’ (...) ‘sheep’ and ‘ship’. You see the difference? There are /i/ and /ɪ/, which are different. I heard that there is only one /i/ in Korean. Don’t be confused in English. The sounds are totally different; /i/ and /ɪ/. ‘sheep’ with /i/ and ‘ship’ with /ɪ/. Okay? (...) Then, could you show me ‘ship’, everyone?

Once they are asked to participate in focused tasks containing a number of categorization tasks between /i/ and /ɪ/, they start to recall the metalinguistic information and previous input in order to deal with the challenges imposed by the tasks. Based on the metalinguistic knowledge and previous input, they build linguistic hypotheses to declare ‘match’ or ‘mismatch’ status. However, due to the match-biased perceptual state, they make a great number of perceptual errors when /ɪ/ is provided as auditory input during the focused tasks. In the case of the Instruction + CF group, corrective feedback is immediately provided on the learners’ errors such that the match-biased perceptual state is challenged by negative evidence. In addition, the feedback types used in this study entail verbal moves that push learners to find the right representation. Such demands induce the match-biased perceptual state to be transformed, so that the system declares ‘mismatch’ status and registers /ɪ/ appropriately. While pushing the current perceptual state to be restructured, corrective feedback also provides the learners with opportunities to test whether their linguistic hypotheses are sufficient to declare ‘mismatch’ when /ɪ/ is given. For example, when the instructor asks learners to pick up the card ‘ship’, the learners are confronted with a decision point (i.e., ‘ship’ or ‘sheep’). Due to the match-biased perceptual state, the learners are likely to perceive the auditory input ‘ship’ as ‘sheep’. However, with respect to the Instruction + CF group, corrective feedback predisposes them to notice that their linguistic hypotheses, which are mainly built on the match-biased perceptual state, are not acceptable in the target language and pushes them to generate new hypotheses. Once they generate the new hypotheses and their hypotheses override the match-biased perceptual state, the perceptual system declares ‘mismatch’ status such that the L2 learners select the right representation ‘ship’ in response to the auditory input ‘ship’. Finally, with instructors’ confirmation such as ‘right, that is (a correct form)’, ‘perfect’, ‘good’, or ‘correct’, the learners are able to consolidate their linguistic hypotheses. Through the procedure of reinforcement embedded in the feedback sequence of this study, the learners are again asked to categorize ‘ship’ between ‘ship’ and ‘sheep’, and are able to double check the new hypotheses.

The same amount of metalinguistic information and input was provided to the Instruction-only group. However, in the absence of the feedback interaction described above, the effects of the metalinguistic information and input were not seemingly manifested, especially in the sense of predisposing them to declare 'mismatch' status in response to the auditory input /ɪ/.

*5.3.2. Linguistic hypotheses and CF with /*i*/*

Due to the match-biased perceptual state, the participants showed higher accuracy on /i/ than /ɪ/. Both Instruction-only and Instruction + CF groups indicated similar accuracy on /i/ at the pretest (around 69% and 72%, respectively). After the instructional sessions, both groups showed the same degree of improvement, which is around 84% for both groups (see also Table 14). If so, do the roles of corrective feedback disappear in contrast to the effects of it on /ɪ/?

Quantitatively speaking, both groups showed similar improvement on /i/. However, it is also important to discuss this result qualitatively: Are the accuracy of the Instruction-only group and that of the Instruction + CF group at the immediate posttest really the same? The results of the generalizability test provide us with a clue to delve into this question. Although both groups showed similar improvement on /i/ at the immediate posttest, intriguingly enough, the Instruction + CF group significantly outperformed the Instruction-only group on the overall /i/

sounds at the generalizability test. More specifically, with respect to the Instruction-only group, the accuracy of the generalizability test was similar to that of the pretest; however, concerning the Instruction + CF group, the accuracy of the generalizability test was significantly higher than that of the pretest.

In response to /i/ as auditory input, in spite of the match-biased perceptual state pervading their perceptual system, the system sometimes declared ‘illusion mismatch’. The Korean learners of English thus chose /ɪ/ as the perceptual representation. Again, in order to declare ‘mismatch’ status, the L2 listeners need in advance to notice the existence of the different non-native phonemes sharing the same default. For example, with respect to naïve Korean listeners of English, they are unlikely to have a background regarding the existence of /i/ and /ɪ/ in English. Therefore, regardless of /i/ and /ɪ/, they are predicted to show a high match-biased perceptual state; /i/ might in turn become the perceptual representation in most cases. In other words, considering that the Korean learners of English participating in this study were recruited from the intermediate level, it is highly possible that, through previous learning experiences, they were already familiar with the existence of /i/ and /ɪ/ without necessarily being able to accurately distinguish them. Because of the emergence of /ɪ/ in their perceptual system, the system incorrectly declared ‘illusion mismatch’ status in response to the auditory input /i/.

With respect to the Instruction + CF group, once they exhibited 'illusion mismatch' status during the focused tasks, corrective feedback was immediately provided. As in the case of /ɪ/, corrective feedback used in this study entailed a number of opportunities to test linguistic hypotheses; in response to non-target choices, they were pushed to generate new hypotheses. By means of the instructor's confirmation and reinforcement, they correctly regenerated linguistic hypotheses such that their perceptual system appropriately declared 'match' status in response to the auditory input /i/. Because the framework of (mis)match-based perceptual decisions in this study is based on 'match' and 'mismatch' status at the segmental level, the Instruction + CF group showed improved accuracy on words that had not appeared in the previous tests and instruction.

On the other hand, regarding the Instruction-only group, they might have acquired lexical items per se, not the target segments. In spite of the absence of confirmation and corrective feedback, a great deal of metalinguistic information and input were equally provided to the Instruction-only group. That is, during the five 1-hour instruction, they were highly exposed to the /i/-embedded testing words such as 'sheep' and 'feel'. Because /i/ is the default in their perceptual system, compared to the testing items with /ɪ/, the lexical items embedded with /i/ could have been easily stored as their mental representation.

Nevertheless, in the absence of corrective feedback, they could not have any opportunity to confirm whether their linguistic hypotheses at the segmental level are correct or not. This explains why they showed less accuracy on the unfamiliar items.

5.3.3. Ultimate attainment of the target vowels

During the focused tasks and feedback interaction, the Instruction + CF group was given repeated opportunities to test their linguistic hypotheses; when wrong, they were pushed to generate new ones and test them again. Consequently, in response to /i/ as the auditory input, the perceptual system began to stabilize in declaring 'match' status and so would register /i/; in addition, in declaring 'mismatch' status over the match-biased perceptual state in response to the auditory input /ɪ/, so the perceptual representation /ɪ/ appeared to the auditory input /ɪ/. Given the repeated opportunities to test hypotheses, those of the Instruction + CF group became increasingly correct and thus the decision making — 'match' to the auditory input /i/ and 'mismatch' to the auditory input /ɪ/ — became more automatized. It is in this way that Korean learners of English were able to exhibit fast and excellent categorization regarding /i/ and /ɪ/.

5.4. Focused tasks and types of CF

In order to guarantee natural space for the feedback interaction between instructors and students, three types of focused tasks were administrated during the lessons: (a) pick-a-card games, (b) bingo games, and (c) fill-in-the-blank exercises. In terms of corrective feedback, two types of corrective feedback were provided on the learners' perceptual errors. Specifically, if any perceptual error occurred, CF1 (a repetition of error verbatim) was delivered with rising intonation. Unless self-repair was elicited after the CF1, CF2 (explicit correction: Not 'X', but 'Y') was provided next.

After counting feedback turns in the 10-hour Instruction + CF classroom video tapes (2 classes of the Instruction + CF group × five 1-hour lessons), around 450 feedback turns were found: 56% of the feedback turns involved only CF1; in other words, immediate self-repairs were elicited after CF1. The rest of the feedback turns (i.e., 44%) were composed of double moves of CF1 and CF2; that is, because self-repair was not available after CF1, instructors provided them with CF2 so that self-repairs were elicited. After the feedback interaction (CF1-only or the double moves), self-repairs were perfectly guaranteed (100%) at all times.

Repetition is known as one type of prompt that signals an error and pushes learners to self-repair. According to Ellis (2006) and Loewen and Nabei (2007), repetition might be more implicit to the learners in that they are less likely to self-repair. However, interestingly enough, this study found that whether repetition resulted in self-repair or not depended on the focused tasks. The feedback turns were overall observed across three types of focused task. During the fill-in-the-blank exercises, 85% of the feedback turns consisted of the double move of CF1 and CF2, whereas double moves only accounted for 30% and 22% during the pick-a-card and bingo games, respectively. In sum, during the pick-a-card and bingo games, repetition was explicit enough to induce the learners to notice their errors and to result in the right representation; however, during the fill-in-the-blank exercises, the learners required more elaborated negative evidence beyond repetition in order to notice their errors and result in the right representation.

Concerning the pick-a-card and bingo games, the target pairs (e.g., 'sheep'-'ship') per se were written on the cards or bingo boards. During the games, what they needed to consider was only the sounds of the target vowels. For example, once they got negative evidence from CF1, they were easily able to notice that the error was related to the target vowel and to try to find the other vowel of the minimal pair.

On the other hand, with respect to the fill-in-the-blank exercises, learners needed to write relevant vowel(s) in response to what instructors said. These exercises required two kinds of linguistic knowledge:
(a) phonological (/i/ or /ɪ/) and (b) orthographic. For example,
if an instructor asks them to write 'lead' in the blank 'L___d', the learners at first need to decide whether it is /lid/ or /lɪd/. Once they decide that

it is /lid/, they should think of corresponding spelling such as ‘leed’ or ‘lead’. Although the instructor gave CF1 with rising intonation, the learners might have been confused about whether their phonological representation (i.e., /i/ or /ɪ/) or the orthographic representation (i.e., ‘ee’ or ‘ea’) was wrong, or both. Due to such ambiguity, learners hesitated to self-repair, and this was followed by the second feedback move along with further explanation.

5.5. Summary

With the main effect of Sound type, it was found that Korean learners of English showed a tendency to perceive the target vowels as /i/ regardless of spectral differences. It is thus hypothesized that /i/ is prioritized as the default in their perceptual system.

When it comes to perceiving a non-native phoneme, L2 learners start to match the non-native phoneme with the default. According to the framework of (mis)match-based perceptual decisions proposed in this chapter, /i/ and /ɪ/ share the default /i/ such that Korean learners of English are predicted to have difficulty categorizing these phonemes. In addition, due to the match-biased perceptual state, they are hypothesized to show more difficulty perceiving the auditory input /ɪ/ as /ɪ/.

After the instructional treatments, the most interesting finding is that the Instruction + CF group showed an increase on /ɪ/, whereas the

Instruction-only group did not. In this regard, with respect to /ɪ/, it turned out that the feedback interaction provided the Instruction + CF group with opportunities to notice their errors, generate new hypotheses, and test whether the hypotheses were sufficiently correct to override the match-biased perceptual state, so the system would correctly register /ɪ/ in response to the auditory input /ɪ/. However, regarding the Instruction-only group, metalinguistic information and input did not seem to be enough to generate linguistic hypotheses beyond the match-biased perceptual state.

Nevertheless, both groups showed similar improvement on /i/ at the immediate posttest. In this chapter, it is argued that their improvement is qualitatively different: For the Instruction + CF group, while testing and modifying the linguistic hypotheses with corrective feedback, they acquired segmental-level knowledge such that they showed higher accuracy on the unfamiliar items containing /i/. On the other hand, concerning the Instruction-only group, they might have acquired the /i/-embedded lexical items per se with the aid of a great deal of input. Considering that /i/ is perceptually prioritized in their perceptual state, the lexical items might be easily stored in their representational system. Nevertheless, with a general lack of segmental-level knowledge, they became relatively poor at categorizing the unfamiliar items with /i/.

Within the framework described in this study, ultimate attainment of the target vowels is also predicted. With the continuous feedback interaction, learners are given a great number of opportunities to test their linguistic hypotheses and modify them so that the system eventually becomes automatized in declaring 'match' status to the auditory input /i/ and 'mismatch' status to the auditory input /ɪ/. By doing so, they show fast and excellent categorization between the target vowels.

In order to provide the learners with a natural environment for the feedback interaction, three types of focused tasks were assigned every lesson. Among the feedback moves, CF1 (repetition) accounted for 56%. In case of the absence of self-repair after the CF1, CF2 (explicit correction) followed: Double moves of CF1 and CF2 accounted for 44% of all feedback. Intriguingly enough, depending on the type of focused task, different patterns were observed. For the pick-a-card and bingo games, CF1 managed to elicit the self-repair; however, the double moves and further explanation were highly exploited during the fill-in-the-blank exercises. It was because repetition itself was inadequate for enabling learners to self-repair, considering that these exercises required two kinds of linguistic knowledge: phonological and orthographic.

CHAPTER 6

CONCLUSION

In this chapter, a brief conclusion is presented. While summarizing the current study, the implications of the present study are addressed in relation to the fields of L2 education and speech perception research. In addition, limitations of the study are presented, and suggestions for future research are proposed.

6.1. Implications of the present study

Let us go back to Kuhl (2000): "No speaker of any language perceives acoustic reality; in each case, perception is altered in the service of language" (p. 11,852). As mentioned before, this statement provided us with another conundrum: What is meant by 'the service of language'?

In this study, the service of language could be defined as the L2 learners' perceptual system that is intrinsically linked to L1 background. Specifically, it is argued that Korean learners of English have difficulty categorizing /i/ and /ɪ/ due to the match-biased perceptual state. However, it is also very important to stress that the service of language has alterability — the service of language can evolve with appropriate language exposure. Specifically, with respect to L2 speech perception, considering that most of EFL learning takes place in the classroom setting, appropriate language exposure occurs mainly with instruction, teaching aids, and interaction between instructors and students. Among them, this study sheds light on the roles of corrective feedback on L2 speech perception training in the classroom setting. During focused tasks and feedback interaction, L2 learners have opportunities to build their own linguistic hypotheses and test whether the hypotheses are sufficient to lead to the right perceptual decisions. If not, corrective feedback also provides them with room to generate new hypotheses, so the perceptual system can ultimately register the right representation.

Previous research findings regarding laboratory-based perceptual training have emphasized the roles of feedback. However, what has so far received little attention is why and how corrective feedback helps L2 learners to perceive non-native phonemes correctly during the training sessions. In this regard, the current study suggests a pivotal role for corrective feedback and its importance in speech perceptual training. That is, corrective feedback provides opportunities for L2 learners to (re)confirm their L2 perceptual representations and, through repeated instances, can lead to a restructuring of the perceptual system.

In addition, many SLA studies have found effects for corrective feedback on L2 grammatical, lexical, and pragmatic targets, and recently on L2 speech production. Given that there are no previous studies investigating the effects of feedback on L2 speech perception, this study is expected to expand horizons in regard to the roles attributed to corrective feedback.

6.2. Limitations and future studies

6.2.1. Limitations

There was no control group in the current study that received no instruction and no CF. While addressing the result that the Instruction-only group showed improvement on /i/ at the immediate posttest, this study attributed the improvement to the effects of instructional treatments (except for CF) such as metalinguistic information and input. However, considering the number of testing words embedded with /i/, the improvement could have resulted from test-retest effects.

The testing words were selected from minimal pairs of CVC monosyllabic words that fell into a high-frequency ESL vocabulary list. However, there is also a possibility that the degree of frequency would vary within the testing words, which could affect the L2 learners' perceptual accuracy. In addition, the fact that the vowel quality slightly differs depending on the surrounding consonants might have an influence on their perceptual accuracy as well. Therefore, the reason why the Instruction-only group did not show an increase on /i/ at the generalizability test might be because of testing item effects.

6.2.2. Future studies

Compared to laboratory-based perceptual training, one of the less promising results in this study is that both Instruction-only and Instruction + CF groups did not drastically reduce their use of durational information in the case of synthesized sounds.

Although the Instruction + CF group showed an increase on the /i/-synthesized and /ɪ/-synthesized sounds, and the Instruction-only group showed improvement on the /i/-synthesized sounds, they seemed to be frequently affected by durational information even after the instruction, especially in the /ɪ/-synthesized sounds. When comparing the results with previous laboratory-based studies of perceptual training, it becomes more distinct. For example, Wang and Munro (2004) found that, after the training, trainees showed a high level of accuracy on the target vowels regardless of duration differences while mentioning that "the trainees did indeed learn to focus their attention away from duration and toward vowel quality" (Wang & Munro, 2004, p. 548). However, such a disparity might stem from the fundamental difference between laboratory-based and classroom-based perceptual training — participants are trained with the synthesized sounds per se in the laboratory (especially, computer-assisted) setting. However, in the classroom setting, considering that instructors are the primary providers of input and corrective feedback, it is hard to expect participants to be amply exposed to the synthesized sounds (e.g., /hɪt/ with 130 ms, 160 ms, 190 ms, 220 ms, and 250 ms) from human speech. They have therefore fewer opportunities to encounter these synthesized sounds and to receive corrective feedback or confirmation of their perceptual representations in response to the synthesized sounds.

Nevertheless, given that both groups started to pay more attention to spectral information after the instruction, it could be a matter of the length of instruction; the participants were trained for two months in the Wang and Munro's (2004) study, while the participants took five 1-hour lessons in this study. In future research, it would be worth delving into the question as to whether L2 learners would benefit even more from instruction and corrective feedback with a longer instructional period.

It is well known that speech production and perception are highly linked to each other. Previous empirical studies found that L2 learners benefited solely from perceptual training without explicit pronunciation instruction on their L2 production (Bradlow et al., 1997; Jamieson & Rvachew, 1992; Rochet, 1995; Rvachew & Jamieson, 1995; Wang et al., 2003). In contrast to this, Saito (2013) recently revealed that L2 learners' perceptual accuracy was improved after production-focused training in a classroom setting. It would thus be interesting to investigate (a) the extent to which L2 learners can improve their production accuracy of the target vowels /i/ and /ɪ/ solely after classroom-based perceptual training and (b) the extent to which they can improve their perceptual accuracy of the target vowels only after classroom-based production training. These future studies are expected to reveal the concrete relationship between L2 speech production and perception.

In spite of the importance of speech perception, it has been predominately with respect to speech production that L2 pedagogical issues have been discussed. Yet Borden, Gerber, and Milsark (1983) suggested that perception abilities might be a prerequisite for successful L2 speech production. In pursuing the question of how to teach L2 pronunciation most effectively, the fruitful solution to this question may lie in the realm of L2 speech perception.

REFERENCES

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

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R. (1985). *Cognitive psychology and its implications*. New York: Freeman.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of Learning Sciences, 4*(2), 167-207.

Baker, W., & Trofimovich, P. (2005). Interaction of native-and second-language vowel system (s) in early and late bilinguals. *Language and Speech, 48*(1), 1-27.

Baker, W., Trofimovich, P., Mack, M., & Flege, J. E. (2002). The effect of perceived phonetic similarity on non-native sound learning by children and adults. In S. Fisch, B. Scarabela, & A.-H. Do (Eds.), *Proceedings of the 26th Boston University Conference on Language Development* (Vol. 26, pp. 36-47). Somerville, MA: Cascadilla Press.

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

Boersma, P. (1998). *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. (Doctoral dissertation), University of Amsterdam, The Hague, Netherlands.

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. *Proceedings of the 15th international congress of phonetic sciences*, 1013-1016.

Boersma, P., & Hayes, B. (2001). Empirical tests of the gradual learning algorithm. *Linguistic inquiry, 32*(1), 45-86.

Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program] (Version 5.3.41). Retrieved from http://www.praat.org/

Bohn, O.-S. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech perception and linguistic experience: Issue in cross-language research* (pp. 279-304). Timonium, MD: York Press.

Bohn, O.-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics, 11*(3), 303-328.

Bohn, O.-S., & Polka, L. (2001). Target spectral, dynamic spectral, and duration cues in infant perception of German vowels. *The Journal of the Acoustical Society of America, 110*, 504-515.

Borden, G., Gerber, A., & Milsark, G. (1983). Production and perception of the /r/-/l/ contrast in Korean adults learning English. *Language Learning, 33*(4), 499-526.

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

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology, 6*(2), 201-251.

Carney, A. E., Widin, G. P., & Viemeister, N. F. (1977). Noncategorical perception of stop consonants differing in VOT. *The Journal of the Acoustical Society of America, 62*(4), 961-970.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, and Computers, 25*(2), 257-271.

Cook, V. J. (2002). Background to the L2 user. In V. J. Cook (Ed.), *Portraits of the L2 user* (pp. 1-28). Clevedon: Multilingual Matters.

DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42-63). Cambridge: Cambridge University Press.

DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125-151). Cambridge: Cambridge University Press.

Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114–138). New York: Cambridge University Press.

Ellis, R. (2001). Introduction: Investigating form-focused instruction. *Language Learning, 51*(s1), 1-46.

Ellis, R. (2006). Researching the effects of form-focussed instruction on L2 acquisition. *AILA Review, 19*(1), 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. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization.* (Unpublished doctoral dissertation), University of Utrecht, Utrecht, Netherlands.

Escudero, P. (2006). Second language phonology: The role of perception. In M. Pennington (Ed.), *Phonology in context* (pp. 109-134). New York: Palgrave Macmillan.

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

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

Flege, J. E. (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. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*(4), 437-470.

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics, 14*(1), 3-28.

Han, Z., Park, E., & Combs, C. (2008). Textual enhancement of input: Issues and possibilities. *Applied Linguistics, 29*(4), 597-618.

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

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical society of America, 97*(5), 3099-3111.

Ingram, J. C., & Park, S. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of phonetics, 25*(3), 343-370.

Jamieson, D. G., & Morosan, D. E. (1986). Training non-native speech contrasts in adults: Acquisition of the English/ð/-/θ/contrast by francophones. *Perception & Psychophysics, 40*(4), 205-215.

Jamieson, D. G., & Morosan, D. E. (1989). Training new, nonnative speech contrasts: A comparison of the prototype and perceptual fading techniques. *Canadian Journal of Psychology/Revue canadienne de psychologie, 43*(1), 88-96.

Jamieson, D. G., & Rvachew, S. (1992). Remediating speech production errors with sound identification training. *Journal of Speech-Language Pathology and Audiology, 16*, 201-210.

Johnson, K. (1996). *Language teaching and skill learning*. Oxford: Blackwell.

Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *The Journal of the Acoustical Society of America, 67*, 971-995.

Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Press.

Krashen, S. (1985). *The input hypothesis*. London: Longman.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences USA, 97*(22), 11850-11857.

Lambacher, S. G., Martens, W. L., Kakehi, K., Marasinghe, C. A., & Molholt, G. (2005). The effects of identification training on the identification and production of American English vowels by native speakers of Japanese. *Applied Psycholinguistics, 26*(2), 227-247.

Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In J. Leather & A. James (Eds.), *New Sounds 90: Proceedings of the Amsterdam Symposium on the Acquisition of Second Language Speech* (pp. 72-97). Amsterdam: University of Amsterdam.

Leather, J. (1999). Second-language speech research: An Introduction. In J. Leather (Ed.), *Phonological Issues in Language Learning* (pp. 1-58). Oxford: Blackwell.

Lee, H. (1993). Korean. *Journal of the International Phonetic Association, 23*(1), 28-31.

Lee, K. (2008). *The effect of perceptual training on the perception and production of English vowel contrasts by Korean speakers.* (Unpublished master's thesis), Hanyang University, Seoul, Korea.

Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning, 60*(2), 309-365.

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

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English/r/and/l/. III. Long-term retention of new phonetic categories. *The Journal of the Acoustical Society of America, 96*(4), 2076-2087.

Llisterri, J. (1995). Relationships between speech production and speech perception in a second language. In K. Elenius & P. Branderud (Eds.), *Proceedings of the 13th International Congress of Phonetic Sciences* (pp. 92-99). Stockholm: KTH/Stockholm University.

Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 361-377).

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English/r/and/l: A first report. *The Journal of the Acoustical Society of America, 89*(2), 874-886.

The Longman Dictionary of Contemporary English (2009). (M. Mayor Ed. 5th ed.). New York: Pearson Longman.

Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition, 20*(1), 51-81.

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

Lyster, R. (2004b). Research on form-focused instruction in immersion classrooms: Implications for theory and practice. *Journal of French Language Studies, 14*(3), 321-341.

Lyster, R., & Izquierdo, J. (2009). Prompts versus recasts in dyadic interaction. *Language Learning, 59*(2), 453-498.

Lyster, R., & Mori, H. (2006). Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition, 28*(2), 269-300.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition, 19*(1), 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.

Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407–452). Oxford: Oxford University Press.

Magen, H. S., & Blumstein, S. E. (1993). Effects of speaking rate on the vowel length distinction in Korean. *Journal of Phonetics, 21*(4), 387-409.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*(3), B101-B111.

McLaughlin, B. (1990). Restructuring. *Applied Linguistics, 11*, 113-128.

Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder Education.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America, 24*(2), 175-184.

Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America, 32*(6), 693-703.

Pisoni, D. B., Aslin, R. N., Perey, A. J., & Hennessy, B. L. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance, 8*(2), 297-314.

Polka, L., & Bohn, O.-S. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *The Journal of the Acoustical Society of America, 100*, 577-592.

Polka, L., & Bohn, O.-S. (2003). Asymmetries in vowel perception. *Speech Communication, 41*(1), 221-231.

Polka, L., & Bohn, O.-S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics, 39*(4), 467-478.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20*(2), 421-435.

Prince, A., & Smolensky, P. (1993). *Optimality theory: Constraint interaction in generative grammar*. Technical Report CU-CS-696-93, Department of Computer Science, University of Colorado at Boulder, and Technical Report TR-2, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ.

R Development Core Team. (2010). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

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

Rochet, B. L. (1995). Perception and production of second-language speech sounds by adults. In W. Strange (Ed.), *Speech perception and linguistic experience: Issue in cross-language research* (pp. 379-410). Timonium, MD: York Press.

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133-163). Amsterdam: John Benjamins.

Rvachew, S., & Jamieson, D. G. (1995). Learning new speech contrasts: Evidence from adults learning a second language and children with speech disorders. In W. Strange (Ed.), *Speech perception and linguistic experience: Issue in cross-language research* (pp. 411-432). Timonium, MD: York Press.

Saito, K. (2011). Examining the role of explicit phonetic instruction in native-like and comprehensible pronunciation development: An instructed SLA approach to L2 phonology. *Language Awareness, 20*(1), 45-59.

Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly, 46*(4), 842-854.

Saito, K. (2013). Communicative focus on second language phonetic form: Teaching Japanese learners to perceive and produce English /ɹ/ without explicit instruction. *Applied Psycholinguistics*, 1-33.

Saito, K., & Lyster, R. (2012a). 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.

Saito, K., & Lyster, R. (2012b). Investigating the pedagogical potential of recasts for L2 vowel acquisition. *TESOL Quarterly, 46*(2), 387-398.

Sato, M., & Ballinger, S. (2012). Raising language awareness in peer interaction: A cross-context, cross-methodology examination. *Language Awareness, 21*(1-2), 157-179.

Sato, M., & Lyster, R. (2012). Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization. *Studies in Second Language Acquisition, 34*(4), 591-626.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied linguistics, 11*(2), 129-158.

Schmidt, R. (1993). Consciousness, learning, and interlanguage pragmatics. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 21-42). Oxford: Oxford University Press.

Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1-63). Honolulu: University of Hawaiʻi Press.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). New York: Cambridge University Press.

Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research, 12*(1), 40-72.

Sheen, Y. (2007). The effect of corrective feedback, language aptitude, and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 301-322). Oxford: Oxford University Press.

Sheen, Y., & Ellis, R. (2011). Corrective feedback in language teaching. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. 2, pp. 593-610). New York: Routledge.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*(2), 127-190.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

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

Strange, W., & Dittmann, S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception & Psychophysics, 36*(2), 131-145.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125-144). Oxford: Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471-483). Mahwah, NJ: Erlbaum.

Thomson, R. I. (2007). *Modeling L1/L2 interactions in the perception and production of English vowels by Mandarin L1 speakers: A training study*. (Unpublished doctoral dissertation), University of Alberta, Alberta, Canada.

Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Flege, J. E. (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics, 33*(3), 263-290.

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

Wang, X. (2002). *Training Mandarin and Cantonese speakers to identify English vowel contrasts: Long-term retention and effects on production*. (Unpublished doctoral dissertation), Simon Fraser University, British Columbia, Canada.

Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System, 32*(4), 539-552.

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America, 113*, 1033-1043.

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *The Journal of the Acoustical Society of America, 106*, 3649-3658.

Wayland, R. P., & Guion, S. G. (2004). Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language Learning, 54*(4), 681-712.

Werker, J. F., & Tees, R. C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *The Journal of the Acoustical Society of America, 75*(6), 1866-1878.

Yang, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics, 24*(2), 245-261.

Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learners' acquisition of regular and irregular past-tense forms. *Studies in Second Language Acquisition, 32*(2), 235-263.