Effects of web-based auditory training on the perception of Korean sounds by Mandarin learners of Korean

by

Na-Young Ryu

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy

Linguistics
University of Toronto

© Copyright by Na-Young Ryu 2019
Effects of web-based auditory training on the perception of Korean sounds by Mandarin learners of Korean

Na-Young Ryu
Doctor of Philosophy
Linguistics
University of Toronto
2019

Abstract

This dissertation is concerned with the perceptual acquisition of Korean sounds by adult second language (L2) learners and suggests web-based auditory training as an effective learning tool to enhance L2 learners’ perceptual accuracy of Korean sounds.

Experiment 1 explores the role of experience in the perception of Korean three-way laryngeal contrasts of stops and affricates by 44 native Mandarin listeners differing in Korean language experience. The results revealed that most Mandarin listeners did not reach the same level of perceptual accuracy as native Korean listeners; however, advanced learners perceived contrasts more accurately than beginner and intermediate learners. This indicates that the greater the adult learners’ experience in the L2, the better their perceptual accuracy of L2 contrasts.

The goal of Experiment 2 is to examine whether L1 inventory size affects L2 perception by comparing 28 English and 28 Mandarin listeners’ identification accuracy for Korean vowels and codas. The results showed that English listeners, whose L1 has a rich vowel and coda inventory, outperformed Mandarin listeners, who have a small L1 vowel and coda inventory, suggesting that L1 inventory size has a significant effect on L2 perception.
Experiment 3 investigates the effects of web-based auditory training and compares the efficacy of attention on the perception of Korean vowels and codas. For this study, 45 Mandarin listeners were randomly assigned to one of two training groups to receive either vowel- or coda-training, both exposed to identical stimuli, or to a control group to receive no training. Results demonstrated that both training groups outperformed the control group for the identification of target sounds, suggesting a beneficial effect of web-based training of L2 sounds. In addition, both the vowel- and coda-trained group yielded significant improvement in identifying Korean vowels; however, only the coda-trained group achieved a significant increase in coda identification accuracy. These findings indicate that attention type differentially affects perception depending on the target structure and highlights the benefits of explicit over implicit training for at least some L2 sound structures. Furthermore, learners were able to generalize knowledge to a novel phonetic context, providing evidence that robust learning has occurred during training.
Acknowledgements

The journey to achieve my PhD in the Department of Linguistics at the University of Toronto has been a truly life-changing experience, and not something I would have been able to accomplish without the constant support and encouragement I have received.

First and foremost, I would like to express my sincerest gratitude to my supervisor, Dr. Yoonjung Kang for her theoretical and professional advice throughout my Ph.D research studies. She has greatly contributed to improving the quality of my work - from conference abstracts and presentations, grant applications, manuscript submissions as well as this dissertation. With each meeting, my intellectual excitement has grown as she has shared her exceptional scientific knowledge, and motivated me towards a more constructive and beneficial direction, both in terms of my research and personal life. She is the star that has illuminated my path, allowing me to find my way whilst inspiring me and strengthening me so that I may surpass what I believed were my limits, and do more than I ever thought myself capable of. Moreover, her professional guidance has infused me with all the energy and light necessary to share what I have learned from her in a new world.

I feel extremely lucky to have Dr. Philip Monahan and Dr. Jessamyn Schertz as my committee members. Their challenging questions and critiques of my thesis have pushed me to refine my thinking, and make my arguments stronger and my writing clearer. I am very grateful for all the detailed feedback they have given me on my work, as well as their technical and emotional support during my Ph.D program.

I am also deeply indebted to Dr. Kyoungrok Ko and Dr. Yoojung Choi who have provided me with the opportunity to teach Korean as an instructor in the Department of East Asian Studies, and who have shared their Korean teaching expertise with me. I also cannot go without mentioning my wonderful students who have taken my Korean language and linguistics courses. It is with absolute confidence that I say, my students are my teacher. It is through them that I am able to continuously evolve my teaching, and it is because of them that I have chosen to pursue a career as an educator at the academic level. I want to be the same star to them, as my supervisor was to me; both a guiding, and illuminating light.
A warm thanks to my cohort - Erin Hall, Emilia Melara, Patrick Murphy, Ruth Maddeaux - for going through this long process with me. Especially Erin Hall, who has been enthusiastic in helping me both in an academic and non-academic way. She has tirelessly edited reams of my writing and emotionally supported my work. I really appreciate her friendship.

I am also lucky enough to be surrounded by amazing people who are always there for me. Connie Ting, Rob Macgregor, Moon-Sung Seo, Kyla Munden, each of whom has always wished me well and whose time I have thoroughly enjoyed. I am truly thankful for their always prompt help, whenever I needed it. I am also grateful to my parents, two little sisters, Da-young and Ka-young for their everlasting support, encouragement, and love which helped me continue to follow my dreams.

Last but not least, I would like to thank, from the bottom of my heart, my husband Hyoung-seok Kwon for supporting me, and accompany me on this journey in Canada. He has always given me immeasurable love and care. If it is possible to meet you in my next life; I would like to be your wife again to return what I received from you in this life. I love you.
# Table of Contents

Acknowledgements........................................................................................................................ iv  
Table of Contents........................................................................................................................... vi  
List of Tables ................................................................................................................................. ix  
List of Figures ................................................................................................................................ xi  
List of Appendices ....................................................................................................................... xiii  
Chapter 1 Introduction .....................................................................................................................1  
  1.1 Purpose of the study.............................................................................................................1  
  1.2 Research questions and predictions .....................................................................................2  
  1.3 Outline of the study..............................................................................................................3  
Chapter 2 Literature review .............................................................................................................5  
  2.1 L2 speech models.................................................................................................................5  
  2.2 The role of foreign language experience in L2 perception..................................................7  
  2.3 The role of L1 inventory size in L2 perception ...................................................................9  
  2.4 Perceptual training studies .................................................................................................10  
  2.5 Implicit and explicit instruction in L2 phonology .............................................................15  
  2.6 Standard Korean, Mandarin and English phoneme inventories.........................................18  
  2.6.1 Korean, Mandarin and English vowel inventory ...................................................18  
Chapter 3 Experiment 1: Effects of L2 experience on the perception of Korean three-way  
contrasts of stops and affricates by native Mandarin listeners..................................................22  
  3.1 Introduction........................................................................................................................22  
  3.2 Research questions and predictions ...................................................................................24  
  3.3 Method..................................................................................................................................24  
  3.3.1 Overview.........................................................................................................................24  
  3.3.2 Participants.....................................................................................................................25  
  3.3.3 Stimuli............................................................................................................................26
Chapter 4 Experiment 2: Effects of L1 inventory size on the perception of Korean vowels and codas by native Mandarin and English listeners

4.1 Introduction

4.2 Research questions and predictions

4.3 Method

4.3.1 Overview

4.3.2 Participants

4.3.3 Stimuli

4.3.4 Procedure

4.3.5 Statistical analysis

4.4 Results

4.4.1 Effects of L1 inventory size on L2 vowel perception

4.4.2 Effects of L1 inventory size on L2 coda perception

4.5 Summary and discussion

Chapter 5 Experiment 3: Effects of web-based auditory training on the perception of Korean vowels and codas by native Mandarin listeners

5.1 Introduction

5.2 Research questions and predictions

5.3 Method

5.3.1 Overall design

5.3.2 Participants
5.3.3 Stimuli ........................................................................................................................68
5.3.4 Procedure ....................................................................................................................69
5.3.5 Statistical analysis ......................................................................................................73
5.4 Results ................................................................................................................................74
5.4.1 Effects of web-based auditory training on L2 perception ....................................74
5.4.2 Effects of implicit and explicit training on L2 perception ....................................75
5.4.3 Effects of implicit and explicit training on individual L2 sounds .........................79
5.4.4 Individual learners’ development of L2 perception ...........................................83
5.4.5 Identification performance during online training ................................................84
5.4.6 Response patterns for L2 sounds ...........................................................................85
5.4.7 Effects of generalization on L2 perception ............................................................87
5.5 Summary and discussion .............................................................................................92

Chapter 6 Conclusion .........................................................................................................97
6.1 Summary of the present study ..................................................................................97
6.2 Theoretical implications .............................................................................................98
6.3 Pedagogical implications ............................................................................................102
6.4 Contribution of this study .......................................................................................103
6.5 Limitations and future directions .............................................................................104

References .........................................................................................................................106

Appendices ........................................................................................................................118
List of Tables

Table 1. Inventory of Standard Korean vowels (adapted from Shin 2015, p.6) ....................... 18

Table 2. Inventory of Standard Mandarin vowels (adapted from Lin 2007, p.82) ................. 18

Table 3. Inventory of Standard Canadian English vowels (adapted from Labov, Ash and Boberg 2008, p. 217-220) ............................................................................................................................... 19

Table 4. Inventory of Standard Korean consonants ................................................................ 19

Table 5. Inventory of Standard Mandarin consonants ............................................................ 20

Table 6. Inventory of Standard Canadian English consonants ............................................. 20

Table 7. Syllable structure and phonotactic restrictions of Korean, Mandarin and English ....... 21

Table 8. Characteristics of the four groups ............................................................................. 25

Table 9. Examples of target stimuli used in the AX task ....................................................... 26

Table 10. Results of the mixed-effects model for the Korean three-way contrast in word-initial stops and affricates in the AX discrimination task ................................................................................ 32

Table 11. Post-hoc analysis of interaction between CONDITION and L2 EXPERIENCE ........ 34

Table 12. Post-hoc analysis of interaction between L2 EXPERIENCE and CONDITION ........ 35

Table 13. Results of the mixed-effects model of the Korean three-way contrast in word-initial stops and affricates in the identification task ................................................................................ 39

Table 14. Post-hoc analysis of interaction between LARYNGEAL CATEGORY and L2 EXPERIENCE .............................................................................................................................................. 41

Table 15. Information about participants in the study ............................................................ 47

Table 16. Information on stimuli in the identification task ...................................................... 48

Table 17. Results of the mixed-effects model of perception accuracy of Korean vowels ....... 51
Table 18. Results of the mixed-effects model of perception accuracy of Korean codas

Table 19. Post-hoc analysis of interaction between CODA TYPE and L1

Table 20. Design of the current experiment

Table 21. Information about participants in the study

Table 22. Information on stimuli at pre-test, training, post-test, and generalization test

Table 23. Mean % scores for Korean vowels and codas at pre-test and post-test and mean % of improvement by group

Table 24. Results of the mixed-effects model of perception accuracy of Korean vowels

Table 25. Post-hoc analysis of interaction between GROUP and TEST for Korean vowels

Table 26. Results of the mixed-effects model of perception accuracy of Korean codas

Table 27. Post-hoc analysis of interaction between TEST and GROUP for Korean codas

Table 28. Mean % scores for Korean vowels and codas at pre-test, post-test and generalization test

Table 29. Results of the mixed-effects model of generalization effects of Korean vowels

Table 30. Post-hoc analysis of interaction between GROUP and TEST for generalization effects of Korean vowels

Table 31. Results of the mixed-effects model of generalization effects of Korean codas

Table 32. Post-hoc analysis of interaction between TEST and GROUP for generalization effects of Korean codas
List of Figures

Figure 1. Example of instructions for Mandarin participants in the AX discrimination task ........ 28

Figure 2. Example of instructions for Mandarin participants in the identification task .......... 29

Figure 3. Discrimination accuracy of Korean pairwise contrasts by L2 experience .......... 30

Figure 4. Effects of L2 experience on the discrimination of different sound pairs of Korean stops and affricates by Mandarin listeners ................................................................. 31

Figure 5. Effects of condition on the discrimination of Korean stops and affricates by Mandarin listeners ............................................................................................................................................ 32

Figure 6. Identification accuracy of the Korean three-way categories by L2 experience .... 37

Figure 7. Distribution of response patterns of the Korean three-way contrast by Mandarin listeners ............................................................................................................................................ 38

Figure 8. Distribution of response patterns of the Korean three-way contrast by L2 experience 39

Figure 9. Example of stop coda identification trial ................................................................. 49

Figure 10. Boxplots of identification accuracy of Korean vowels by L1 group .................... 50

Figure 11. Identification accuracy of individual Korean vowels by L1 group ....................... 52

Figure 12. Distribution of response patterns for Korean vowels by English and Mandarin listeners ............................................................................................................................................ 53

Figure 13. Boxplots of identification accuracy of Korean codas by L1 group ....................... 54

Figure 14. Identification accuracy of Korean nasal codas by L1 group ............................... 55

Figure 15. Identification accuracy of Korean stop codas by L1 group ................................. 56

Figure 16. Distribution of response patterns of Korean nasal codas by English and Mandarin listeners ............................................................................................................................................ 57
Figure 17. Distribution of response patterns of Korean stop codas by English and Mandarin listeners
......................................................................................................................................................... 57

Figure 18. Examples of web-based training. Left side: Korean coda identification task, right side:
Korean vowel identification task ........................................................................................................ 71

Figure 19. Examples of immediate feedback on the training website ........................................ 72

Figure 20. Identification accuracy for Korean vowels (left side) and codas (right side) across
three groups at pre-test and post-test ............................................................................................. 74

Figure 21. All groups’ identification accuracy of individual Korean vowels and codas at pre-test
and post-test ........................................................................................................................................ 80

Figure 22. Identification accuracy of individual Korean vowels at pre-test and post-test by group
.......................................................................................................................................................... 81

Figure 23. Perception improvement for individual Korean vowels between pre-test and post-test
.......................................................................................................................................................... 81

Figure 24. Identification accuracy of individual Korean codas at pre-test and post-test by group
.......................................................................................................................................................... 82

Figure 25. Perception improvement of individual Korean codas between pre-test and post-test. 82

Figure 26. Individual learners’ perception development for Korean vowels (left side) and codas
(right side) at pre-test and post-test ............................................................................................... 83

Figure 27. Identification accuracy across eight training sessions (S1-S8) by group ................. 85

Figure 28. Distribution of response patterns of the Korean vowels at pre-test and post-test ...... 86

Figure 29. Distribution of response patterns of the Korean codas at pre-test and post-test ........ 87

Figure 30. Identification accuracy for Korean vowels (left side) and codas (right side) across two
groups at pre-test, post-test and generalization test ........................................................................ 88
List of Appendices

Appendix A. Korean stops and affricates in word-initial position (Real words) ...................... 118

Appendix B. Korean stops and affricates in word-initial position (Nonce words) .................... 119

Appendix C. Korean monosyllabic words for the vowel perception task ................................. 119

Appendix D. Korean monosyllabic words including stop codas for the coda perception tasks 120

Appendix E. Korean monosyllabic words including nasal codas for the coda perception tasks 120

Appendix F. Korean monosyllabic words including all vowels and codas used at pre-test, training and post-test (Onset: /h/) ...................................................................................................................... 121

Appendix G. Korean monosyllabic words including all vowels and codas used at generalization test (Onset: /k/) .................................................................................................................................................. 121

Appendix H. Instruction form for online perceptual training ................................................... 122

Appendix I. [Experiment 2] A post-hoc analysis of interaction between Korean individual vowels and L1 for the perception accuracy of Korean vowels ................................................. 123

Appendix J. Online language background questionnaire for Experiment 3 ............................. 123
Chapter 1
Introduction

1.1 Purpose of the study

In second language (L2) acquisition of phonology, adult learners often experience difficulty perceiving non-native sounds. To account for the perceptual difficulty in L2 category learning, many researchers highlighted the importance of language-internal factors such as transfer (Best, 1994; Flege, 1995, 2003) and markedness (Broselow & Xu, 2004; Eckman, 1997), and language-external factors including L2 experience (Best & Strange, 1992; Cebrian, 2006; Flege & MacKay, 2004), average age of L2 acquisition (Hyltenstam & Abrahamsson, 2003), length of L2 immersion (Flege, Frieda, & Nozawa, 1997) and extent of daily L2 vs. L1 usage (Jia, Aaronson, & Wu, 2002).

To enhance L2 learners’ perception abilities for non-native sounds, a number of studies over the past 25 years have shown relatively consistent results: L2 learners’ perception clearly benefits from perceptual training (Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Lopez-Soto & Kewley-Port, 2009; Rato, 2013; Sakai & Moorman, 2017; among many others). However, there is a scarcity of studies which directly compare the effects of implicit and explicit attention on the perception of L2 sounds through perceptual training (Carlet, 2017; Carlet & Cebrian, 2015; Nozawa, 2015) and 97% of training studies analyzed English as a first or target language (Lee, Jang, & Plonsky, 2014).

This dissertation therefore investigates how adult L2 learners perceive Korean vowels and consonants and compares the efficacy of implicit and explicit attention of the perception of Korean sounds through web-based auditory training. The purpose of this dissertation is threefold. The first goal of this dissertation is to examine how native Mandarin listeners with different amounts of Korean language experience in an academic setting in Canada perceive the Korean three-way laryngeal contrasts in stops and affricates tested by Experiment 1. The second goal is to explore whether L1 vowel and coda inventory sizes affect the perception of L2 sounds by comparing English and Mandarin listeners’ identification patterns for Korean vowels and codas tested by Experiment 2. These experiments provide baseline data showing which kinds of Korean sounds are difficult for L2 learners, which in turn will be used to develop the training paradigm used in Experiment 3. Finally, to help native Mandarin learners improve their
perception ability for Korean vowels and codas, I developed a web-based auditory training program. In Experiment 3, I demonstrate that the web-based auditory training using the high variability phonetic training (HVPT) method can successfully orient Mandarin learners’ attention to Korean phonetic information and thereby improve their target sound perception. In addition, I compare the effectiveness of implicit and explicit attention on the identification of Korean sounds through training.

1.2 Research questions and predictions

This dissertation is guided by the following research questions and associated predictions.

[Effects of experience on L2 perception]

Question 1: Are more experienced Mandarin learners of Korean more likely to have better identification and discrimination accuracy in Korean three-way contrasts of stops and affricates which do not occur in their L1 compared to less experienced Mandarin learners of Korean?

Prediction 1: Based on the positive effect of L2 experience on the perception of L2 sounds (Best, 1995; Best & Strange, 1992; Best & Tyler, 2007; Bohn, 1995; Bohn & Flege, 1990; Broselow, Chen, & Wang, 1998; Flege, 1992, 1995; Flege, Bohn, & Jang, 1997; Flege & MacKay, 2004; Ingram & Park, 1997; Major & Faudree, 1996), I hypothesize that the greater the language experience of Mandarin-speaking learners of Korean, the better their perception accuracy of Korean sounds.

[Effects of L1 phonemic inventory size on L2 perception]

Question 2: Does L1 inventory size affect the perception of L2 sounds?

Prediction 2: Based on findings that support the predictive role of L1 inventory size in L2 perception (Bradlow, 1995; Bundgaard-Nielsen, Best, & Tyler, 2011; Escudero, Sisinni, & Grimaldi, 2014; Flege & Wang, 1989; Fox, Flege, & Munro, 1995; Iverson & Evans, 2007, 2009; Scholes, 1968), I hypothesize that native Canadian English listeners, whose vowel and coda inventories are larger than those of Korean, will be more successful in perceiving Korean vowels and codas than native Mandarin listeners, whose vowel and coda inventories are smaller than those of Korean.
[Effects of web-based auditory training on L2 perception]

Question 3: Do Mandarin learners of Korean improve their perceptual accuracy in identifying Korean vowels and codas as a result of web-based auditory training?

Prediction 3: I hypothesize that native Mandarin speakers who receive web-based training will show improved performance in the perception of Korean vowels and codas, compared to a control group that does not receive web-based training.

Question 4: Does explicit attention in perceptual training lead to greater improvement in the perception of Korean vowels and codas than implicit attention, even when learners are exposed to identical L2 input during training?

Prediction 4: I hypothesize with respect to vowel perception performance that the vowel-trained group with explicit attention will show more improvement than the coda-trained group that is implicitly exposed to Korean vowels during training. Similarly, the coda-trained group with explicit attention will display more improvement for Korean coda perception than the vowel-trained group which is passively exposed to Korean codas during training. In other words, improvement on trained sounds with explicit attention is expected to be greater than improvement on untrained sounds with passive exposure.

Question 5: Can learning effects gained from web-based auditory training be successfully transferred to novel stimuli?

Prediction 5: I hypothesize that perceptual improvement obtained through web-based training will transfer to novel words.

1.3 Outline of the study

This dissertation consists of six chapters, which are briefly summarized as follows.

Chapter 1 introduces the purpose of this study, research questions and predictions, and the structure of the dissertation. In Chapter 2, I review previous literature on current theories of L2 speech learning, the effects of L2 experience and L1 inventory size on the perception of L2 sounds, followed by the effects of perceptual training studies and implicit and explicit attention on L2 speech acquisition. In addition, I present standard Korean, Mandarin and English phoneme
inventories. Chapter 3 describes the methods and results of Experiment 1, which investigates the perception of Korean three-way laryngeal contrasts in stops and affricates by English- and Mandarin-speaking learners of Korean using identification and discrimination tasks. Chapter 4 reports on how Mandarin learners of Korean perceive Korean vowels and codas using an identification task in Experiment 2. Chapter 5 shows the methodological design of the training program and the results of Experiment 3, which examines the effects of web-based auditory training on the perception of Korean vowels and codas by Mandarin-speaking learners of Korean. In addition, it shows how and to what extent implicit and explicit attention in training affects the perception of Korean vowels and codas. In Chapter 6, I summarize the findings from Experiments 1, 2 and 3 and discuss theoretical and pedagogical implications, address some limitations of the current study and suggest future research.
Chapter 2
Literature review

2.1 L2 speech models

In this section, I review the two dominant models in L2 speech acquisition: Speech Learning Model (SLM, Flege, 1995) and Perceptual Assimilation Model (PAM, Best, 1993, 1994, 1995).

The SLM focuses primarily on the L2 speech production of experienced learners, although it also has a strong perceptual component. The SLM proposes that adults retain the ability to accurately perceive the properties of L2 sounds and to form new phonetic categories throughout their lifetime, but the success of L2 learning depends on the perceived cross-language similarity between L2 and L1 sounds, as well as the age at which learning starts and the amount of L1 use. According to the SLM, L2 sound perception is dominated by equivalence classification. That is, learners’ perceptual difficulties can be predicted by the amount of acoustic similarity between L1 and L2 phones. The SLM classifies L2 target phones as “identical”, “new” or “similar” to their L1 phonetic categories in terms of perceived phonetic differences. If L2 phones are “identical” to L1 phones, perceptual difficulties are not expected as a result of positive transfer. In addition, “new” L2 phones, which do not exist in the L1, are perceived with little interference from the L1 and learners eventually establish a new L2 category, given enough exposure to the L2, and facilitated by early age of onset of learning (AOL) and little L1 use. Here, the concept “new” refers to the ability L2 learners have to create a “new” phonetic category once they perceive it as a different sound from a corresponding sound in their L1. On the other hand, the concept “similar” refers to sounds that are only slightly different from their L2 sounds. “Similar” L2 sounds, which are perceived as phonetically close to an L1 category, are expected to be more difficult to acquire, because learners are less likely to create separate L2 categories for these acoustically similar sounds.

While the SLM makes its predictions based on a comparison of individual sounds between the L1 and L2 to explain non-native perception, the PAM (Best, 1995; Best & Strange, 1992), and its extension to L2 acquisition, PAM-L2 (Best & Tyler, 2007), compare L2 sound contrasts to L1 categories.
With regard to cross-language perception, the PAM proposes that accuracy in the discrimination of non-native sound contrasts depends on the way the members of each contrast are perceptually equated to L1 sound categories. For example, when two non-native sounds are assimilated to two different L1 categories (two-category assimilation), discrimination is predicted to be very good. On the other hand, when two non-native sounds are mapped onto a single L1 category (single-category assimilation), discrimination is predicted to be poor. In cases where both members of the L2 contrast are mapped to the same L1 sound, but one member is a “better” exemplar of the native category than the other, discrimination is expected to vary depending on the degree of the category-goodness difference. Furthermore, L2 sounds can be uncategorizable, that is, still heard as speech sounds but not mapped onto any single L1 category or heard as non-speech.

In Best’s PAM, the greatest perceptual difficulty is expected when two distinct L2 sounds are mapped to a single L1 category. On the other hand, L2 sounds which are not assimilated to any L1 sounds or are mapped to two distinct L1 sounds will be accurately perceived, regardless of L2 experience.

The PAM was developed to explain naive listeners’ non-native speech perception, while the SLM was designed to address L2 learners’ production of L2 speech. Neither model was developed to address both situations. Thus, the aim of the PAM-L2 (Best & Tyler, 2007) was to examine whether and how the SLM can be used as a starting point to extend PAM’s non-native speech perception framework to L2 learners. Reinterpreting SLM tenets in PAM terms, PAM-L2 makes predictions about whether L2 learners will form a separate category for an L2 sound based both on its perceived distance from an L1 phonetic category and on a comparison of the L1 and L2 phonological contrasts. Thus, for instance, in the case of category-goodness assimilation, the member of an L2 contrast that is perceived as a good exemplar of a native category is likely to be merged with the L1 sound, whereas the poorer exemplar may be learned as a separate L2 category.

As seen in this section, both the SLM and PAM are influential in L2 speech acquisition. The goal of this current dissertation is to build on the shared insight of these models and to investigate whether increased experience with an L2 can improve the ability to distinguish L2 sounds even in adulthood. This is suggested by the two models which predict that L2 perceptual difficulty may be overcome with sufficient L2 experience.
2.2 The role of foreign language experience in L2 perception

Many researchers have provided promising evidence that increased L2 experience can lead to improved performance on L2 segmental and suprasegmental acquisition (Best, 1995; Best & Strange, 1992; Best & Tyler, 2007; Bohn 1995; Bohn & Flege, 1990; Broselow et al., 1998; Flege, 1992, 1995; Flege et al., 1997; Flege & MacKay, 2004; Ingram & Park, 1997; Major & Faudree, 1996 for L2 segmental property; Towell, 2002; Trofimovich & Backer, 2006 for L2 suprasegmental property).

A number of previous studies investigating the role of experience on L2 consonants have shown that perceptual accuracy of difficult L2 consonants increases with target language experience. For example, Japanese-speaking learners’ discrimination of English /r/ − /l/ has been shown to improve as they gain experience in English and improvement is observed as a consequence of training and explicit feedback (Best & Strange, 1992; Bradlow et al., 1997; Lively, Logan, & Pisoni, 1993; MacKain, Best, & Strange, 1981; Miyawaki et al., 1975). These results showed that improved perception of L2 consonants is given a sufficient amount of native-speaker input in a target language environment. In addition, Flege and Liu (2001) examined the identification of English word-final consonants /b, d, g, p, t, k/ by adult listeners of Chinese with varying experience in English based on their length of residence (LOR) in the United States (the short LOR group: 0.5 to 3.8 years vs. the long LOR group: 3.9 to 15.5 years) and their primary occupation (students vs. non-students). The results showed that students with relatively long LORs obtained significantly higher scores in the identification task than students with relatively short LORs. However, there was no significant difference between the non-students differing in LOR. These findings suggest that adult learners can improve L2 sound perception over time, but only if they receive a substantial amount of L2 input.

With respect to the role of experience on L2 vowel perception, Bohn and Flege (1990) investigated the perception of two English vowel contrasts /i/-/ɪ/ and /ɛ/-/ɛ/ by two native German groups differing in English-language experience. The inexperience and experienced native German groups had been in an English-speaking environment for average of 0.6 and 7.5 years, respectively. Their performance was evaluated through responses to members of synthetic continua (beat-bit, bet-bat). Results demonstrated no significant difference between groups in the identification of /i/-/ɪ/. Since English vowels /i/-/ɪ/ are perceptually similar to German vowels /i,
ɪ/, these results suggest that a vowel contrast involving similar vowels across languages would not be affected by the amount of L2 experience. On the other hand, only the experienced German group successfully identified the /æ/-/ɛ/ contrast, for which the English vowel /æ/ is a new vowel for German learners of English. This suggests that experienced German listeners, relied on spectral rather than temporal cues, similar to native English listeners. Taken together, these findings indicate that foreign language experience affects the perception of two L2 vowel contrasts differently and these differences can be explained by the cross-language correspondences. In a similar study involving more L1 groups, Flege et al. (1997) examined the effect of L2 experience on the perception of English vowels in synthetic beat-bit (/i/-/ɪ/) and bat-bet (/æ/-/ɛ/) continua by German, Spanish, Mandarin and Korean learners of English. They observed that more experienced native German, Spanish, Mandarin and Korean learners of English who had resided in the United States between 5.4 and 9.0 years identified English vowel contrasts more accurately than less experienced learners who had resided in the United States between 0.4 and 0.9 years. This finding shows a positive correlation between length of residence and L2 perception accuracy.

By contrast, other studies have not found a facilitative effect of experience on the perceptual acquisition of L2 sounds (Cebrian, 2006; Flege, Munro, & Fax, 1994; Imai, Flege, & Wayland, 2002). Flege et al. (1994) had English monolinguals and Spanish learners of English rated the dissimilarity of a number of English and Spanish vowels. They observed that experienced Spanish learners who, on average, had lived in the United States for 7 years differ in their dissimilarity ratings from inexperienced learners with an average of 1.8 years in the United States. This finding showed no positive effect of experience on L2 vowel perception. In addition, Cebrian (2006) compared English vowel categorization performance by Catalan speakers living in Barcelona, Spain and Toronto, Canada. Surprisingly, the experienced Catalan speakers who had spent on average 25 years in Canada did not outperform the inexperienced Catalan speakers, who were undergraduate students of English at a university in Barcelona, in English vowel categorization. This result suggests that we should consider other external factors, such as learners’ motivation and the amount of formal instruction in order to account for L2 speech performance. Moreover, Fullana and Mora (2009) investigated the perception and production of voicing contrasts in English word-final obstruents by Catalan/Spanish advanced learners of English varying in amount of exposure to the foreign language (school exposure vs. extra
exposure) in a formal learning setting and age of onset of foreign language learning (before vs. after 8 years). Results of this study indicated that neither exposure nor starting age played a significant role in the perception and production of the English voicing contrasts.

As seen above, there are mixed results on the effects of L2 experience on L2 segmental perception, as assessed by length of residence in a target language setting. In this dissertation, I examine whether learners’ L2 experience plays a fundamental role in L2 perceptual learning by assessing Korean vowel identification and discrimination accuracy by three groups of native Mandarin listeners differing in Korean language experience in an academic setting.

2.3 The role of L1 inventory size in L2 perception

A number of studies have examined the role of L1 inventory size in the perception of L2 sounds yielding somewhat conflicting results. Some studies have found that L1 vowel inventory size plays a critical role in L2 vowel perception (Bradlow, 1995; Bundgaard-Nielsen et al., 2011; Escudero et al., 2014; Fox et al., 1995; Iverson & Evans, 2007, 2009; Scholes, 1968). Perceptual difficulties arise when the L1 vowel inventory is smaller than the L2 vowel inventory, as multiple vowels in the L2 are likely to correspond to a single L1 category perceptually. In contrast, a larger L1 than L2 vowel inventory often facilitates L2 perception because there are enough contrastive L1 categories for all L2 sounds to map distinctly. For example, Iverson and Evans (2007) reported that native speakers of German and Norwegian, both having larger vowel systems than English, identified English vowels more accurately than French and Spanish native speakers with smaller L1 vowel inventories. This result suggests that having a larger and more complex vowel system facilitates L2 vowel learning. However, a number of recent studies have shown that L1 inventory size is not always a good predictor of L2 perceptual difficulty (Alispahic, Mulak, & Escudero, 2017; Elvin, Escudero, & Vasiliev, 2014). For instance, Elvin et al. (2014) demonstrated that Iberian Spanish speakers, who have a smaller vowel inventory in comparison to the speakers of Australian English, outperformed Australian English speakers in the discrimination of six Brazilian Portuguese vowel contrasts. This result confirms that cross-linguistic acoustic properties, rather than L1 vowel inventory sizes, successfully predict L2

---

1 French has 11 oral monophthongs and four nasal vowels. In this study, authors used only the oral vowels in the analysis.
discrimination difficulty. In other words, this indicates that the L1 vowel inventory size may not be sufficient for accurately predicting L2 discrimination difficulty.

Similar to the effect of L1 vowel inventory size on the perception of L2 vowels, there are mixed findings about the influence of L1 coda inventory size and phonotactics on the performance of L2 codas. Flege and Wang (1989) demonstrated that Cantonese speakers, whose L1 has /p, t, k/ in word-final position, perceived the English final voicing contrast more accurately than Mandarin speakers, whose L1 does not allow any coda obstruents, suggesting that L1 coda inventory size affects L2 coda perception. On the other hand, de Jong & Hao (2015) found that speakers of Mandarin, whose language prohibits obstruent codas, achieved higher accuracy results for English coda obstruents than speakers of Korean, whose language permits voiceless stop codas. This finding indicates that having codas in the L1 does not necessarily facilitate L2 coda perception.

As seen above, there is contradictory evidence on whether L1 inventory size can successfully predict non-native perception difficulty. This dissertation further examines this issue by evaluating two language groups’ identification of L2 vowels and codas: Canadian English listeners with large vowel and coda inventory sizes and permissive coda phonotactics vs. Mandarin listeners with small vowel and coda inventory sizes.

2.4 Perceptual training studies

Over the past 25 years, a considerable number of empirical studies have investigated how and to what extent perceptual training is effective in improving learners’ ability to acquire L2 segmental and suprasegmental contrasts (Akahane-Yamada, Tohkura, Bradlow, & Pisoni, 1996; Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999; Hardison, 2003; Hazan, Sennema, Iba, & Faulkener, 2005\(^2\); Iverson & Evans, 2007; Pisoni, Lively, & Logan, 1994; Strange & Dittmann, 1984; among others for L2 consonants; Aliaga-Garcia, 2010; Aliaga-García & Mora, 2009; Cenoz & Garcia-Lecumberri, 1999; Iverson & Evans, 2007; Lacabex, García-Lecumberri, &

---

\(^2\) This study investigates whether Japanese learners of English can be trained to make better use of phonetic information from visual cues in their perception and pronunciation of English consonants. Results showed that audiovisual training was more effective than audio training in improving the perception of the English labial/labiodental contrast. In addition, Japanese-speaking learners’ pronunciation of English /l/-/r/ improved significantly after they received audiovisual natural training, suggesting that the facial gestures of the talker leads to a greater improvement in pronunciation.
Cooke, 2008; Nobre-Oliveira, 2007; Rato, 2013; Thomson, 2011; among others for L2 vowels; Francis, Ciocca, Ma, & Fenn, 2008; Wang, Jongman, & Sereno, 2003; Wang, Spence, Jongman, & Sereno, 1999; Wayland & Guion, 2004; Wayland & Li, 2008; among others for L2 suprasegments). A substantial amount of this research has focused on the effects of perceptual training on L2 perception. Perceptual training studies have shown relatively consistent outcomes on the performance of L2 perception: learners’ ability to perceive L2 sounds improves with training, and the improvement is extended to new phonetic contexts and is retained long after training (Akahane-Yamada et al., 1996; Bradlow et al., 1997; Bradlow et al., 1999; Jamieson & Morosan, 1986; Lively et al., 1993; Lively, Yamada, Tohkura, & Pisoni, 1994; Logan, Lively, & Pisoni, 1991; Pisoni, Aslin, Perey, & Hennessy, 1982; Rato, 2013; Rochet & Chen, 1992; Strange & Dittman, 1984; Trapp & Bohn, 2000; and many others).

Many researchers have suggested implementing perceptual training using diverse phonetic training methods to contribute to the understanding of perception and production of L2 sounds. This section introduces various details of phonetic training methods, such as speaker variability (e.g., high variability vs. low variability), number of trained segments (e.g., full sets vs. subsets), training tasks (e.g., identification and/or discrimination tasks) and the location of training (online at home vs. laboratory setting). See Sakai & Moorman (2017) for a recent comprehensive review of L2 perception training studies.

Most of the success noted in perceptual training research has been attributed to the high variability perceptual training (HVPT) paradigm (Aliaga-García & Mora, 2009; Bradlow et al., 1997; Herd, Jongman, & Sereno, 2013; Iverson & Evans, 2009; Iverson, Hazan, & Bannister, 2005; Lambacher, Martens, Kakehi, Marasinghe, & Molholt, 2005; Lengeris, 2008; Lively et al., 1993; Logan et al., 1991; Nishi & Kewley-Port, 2007, 2008; Rato, 2013; Thomson, 2011, 2012; Trapp & Bohn, 2002; Wong, 2013). The HVPT method proposed by Logan et al. (1991) makes use of multiple talkers and multiple phonetic contexts to provide natural phonetic variability within a phonetic category so that L2 learners can learn to detect crucial phonetic cues associated with target sounds, as well as cues relevant to their category identification in different conditions. Since the introduction of HVPT, studies investigating phonetic training have found that HVPT leads to significantly greater improvement compared to the low variability phonetic training (LVPT) paradigm.
To compare the effects of the HVPT and LVPT training methods, Lively et al. (1993) trained Japanese listeners to identify English /r/ and /l/ with the two types of training: in Experiment 1, listeners were trained with multiple talkers and three different environments. In Experiment 2, listeners were trained with a single-talker and five different phonetic environments. The results from Experiment 1 showed that subjects yielded a significant increase in accuracy between the pre-test and post-test and were able to generalize both new words produced by a familiar talker and new words produced by an unfamiliar talker. On the other hand, results from Experiment 2 showed that subjects who trained with only a single talker showed improvement in pre-test and post-test performance, but showed little evidence of generalization to new tokens or new talkers. Iverson and Evan (2009) similarly showed that German and Spanish learners improved their perception of English monophthongs and diphthongs at post-test after HVPT training. Moreover, the improvement was sustained for 4-5 months after training for both language groups, suggesting that HVPT has long-term enhancement effects on L2 learners’ vowel identification ability. In the same vein, Thomson (2012) trained Mandarin speakers to perceive ten Canadian English vowels using the HVPT technique. The results revealed that learners’ ability to identify the Canadian English vowels significantly improved after training. In addition, a delayed post-test indicated that the improvement was maintained for one month after training. Pierce (2014) trained Chinese and Korean speakers to improve their perception of English vowels through different paradigms (HVPT vs. LVPT). His study found that learners who were trained with HVPT were able to generalize their learning to new stimuli, while learners who were trained with LVPT were not.

In addition to training L2 segmental contrasts, other studies that investigated perceptual training with the HVPT method showed significant improvement in the perception and production of L2 suprasegmental contrasts. For example, Wang et al. (1999) examined how American learners of Mandarin improve in perceiving four Mandarin tones through HVPT. They demonstrated that improvement in identification was transferred to new stimuli and novel talkers, and that perceptual learning was retained for over six months. The extended study by Wang et al. (2003) examined whether the Mandarin tone contrasts learned during training with the HVPT method can be transferred to production. The study found significant improvement in tone production after training. In sum, results of previous studies on the HVPT method suggest that talker variability plays a critical role in perceptual learning and that this type of training is effective for
learning to generalize L2 learning to new words, and for promoting retention of learning over time.

Another important factor in perceptual training is the number of segments in the training set used to improve learners’ perception of L2 sounds. Nishi and Kewley-Port (2007) trained Japanese speakers on American English (AE) monophthongs using two training sets: the full set of nine vowels vs. three difficult vowels. The results showed that Japanese speakers who were trained on the set of all nine AE monophthongs improved in all nine vowel categories, while those who were trained on the three most difficult vowels improved only for those three vowels and found no improvement for the six untrained vowels. In addition, the significant change in performance was observed only for the full-set group in the generalization task using real words. These results suggest that a training method with a full set of vowels in the target language can help learners more accurately identify L2 sounds. In their extended study (Nishi & Kewley-Port, 2008), the same effects were found for Korean listeners’ perception of the same AE vowels, confirming the importance of including a large set in training.

Previous studies have also explored the influence of the type of training task. Most studies make use of identification and/or discrimination tasks to train learners’ perception and/or production of L2 sounds and have shown mixed results. Several studies suggest that the efficacy of identification training is superior to discrimination training as an L2 training tool (Ellis, 2005; Jamieson & Morosan, 1986; Lively et al., 1993; Logan & Pruitt, 1995; Carlet and Cebrain; 2015; among others). For example, Lively et al. (1993, 1994) and Logan et al. (1991) showed that learners who received identification training generalized perceptual learning to untrained stimuli including /r/ and /l/ in several phonetic environments. Strange and Dittmann (1984) demonstrated that learners who received same-different discrimination training along a synthesized ‘rock-lock’ continuum training did not generalize learning to novel stimuli or natural recordings. Several other researchers, however, found that both identification and discrimination training effectively improve the accuracy of L2 contrasts (Shinohara & Iverson, 2018; Wayland & Li, 2008). Shinohara and Iverson (2018) compared the effects of identification training and discrimination training for Japanese speakers learning English /r/-/l/. A total of 41 adult Japanese speakers took part in both identification and discrimination training and their performance of pre-, mid- post-test of identification, auditory discrimination, category discrimination and /r/-/l/ production were assessed. Results showed that both identification and discrimination training
increased their accuracy in perception and production of the contrast when both training methods incorporated high variability. That is, both identification and discrimination training appeared to have similar effects. However, other studies have shown that discrimination tasks are not optimal for training learners (Jamieson & Morosan, 1986, 1989; Logan & Pruitt, 1995). For example, Jamieson and Morosan (1986) trained Canadian francophone adults to acquire the English /θ/-/ð/ contrast using identification and discrimination tasks. They concluded that discrimination training may not be successful in increasing identification performance because the task asks listeners to pay attention to differences between stimuli, rather than grouping stimuli in terms of their similarity to a prototype. These discrepancies among previous results leave room for further investigation of the effects of training task type.

Another variable considered in L2 perception training is the location of training, namely whether training takes place at home versus in the laboratory. Sakai and Moorman (2017) reported that training at home results in greater gains than training in the laboratory. Laboratory-based training is usually carried out according to a fixed schedule, which does not permit exploration of the benefits of a student-centered and long-term approach in terms of flexibility and accessibility. For example, home-based training allows students to take part in training at any time that is convenient to them and when they are most motivated, rather than being restricted to training at set scheduled lab-training dates/times. Furthermore, once laboratory-based training is complete, students often do not have the opportunity to continue training, whereas training at home can be an ongoing self-motivated endeavor. Additionally, it has been known that traditional laboratory-based research limits the efficiency, feasibility and reliability of research projects, because it is difficult to obtain large samples and is very time-consuming and expensive. To resolve these problems, web-based experiments have emerged in recent years. Many researchers have suggested Mechanical Turk, implemented through a website, as a feasible alternative to laboratory-based experiments in linguistics studies because it yields valid data, while minimizing the cost in time, effort, and expense (Horton, Rand, & Zeckhauser, 2011; Munro et al., 2010; Schnoebelen & Kuperman, 2010; Sprouse, 2011; Wang, Huang, Yao, & Chan, 2015). Moreover, in terms of the effectiveness of web-based versus classroom instructions, Sitzmann, Kraiger, Stewart, and Wisher (2006) reported in their meta-analytic study that web-based instruction is more effective than classroom instruction for teaching declarative knowledge due to the immediate feedback learners receive during training. These findings suggest that web-based
training can be an effective and efficient alternative to laboratory-based training for linguistics and applied linguistics studies.

2.5 Implicit and explicit instruction in L2 phonology

The effectiveness of different types of instruction have been extensively discussed within L2 syntax and morphology in the field of second language acquisition (de Graff, 1997; Dekeyser, 1995; Doughty, 1991; House, Pierrand & Van Daele, 2005; Hulstijn & de Graaff, 1994; Norris & Ortega, 2000; Robinson, 1996; Sanz & Morgan, 2004; Spada & Tomita, 2010; and many others). In recent years, a great deal of attention has been placed on the effects of instruction type on L2 pronunciation (Alves & Magro, 2011; Derwing & Munro, 2005; Derwing, Munro, & Wiebe, 1998; Derwing & Rossiter, 2003; Kennedy & Trofimovich, 2010; Lee et al., 2014; Neri, Mich, Gerosa, & Giuliani, 2008; Saito, 2007, 2011, 2012; Saito & Lyster, 2012; Venkatagiri & Leivs, 2007).

Recently there has been a growing consideration of the importance of instruction on L2 pronunciation development in terms of accuracy, comprehensibility, accentedness, intelligibility and fluency. In a review of the effectiveness of L2 pronunciation instruction, Thomson and Derwing (2014) surveyed 75 studies on L2 pronunciation with a focus on the methods and their results. They found that explicit instruction of phonological forms showed improvement in comprehensibility and intelligibility3 because it orients learners’ attention to phonetic information which promotes L2 learning; however, they noticed that many pronunciation studies did not include a control group, since as Thomson (2011) mentioned, learners may want to receive instruction. Thus, it cannot be concluded that considerable improvement in production was obtained as a direct result of pronunciation instruction.

In another study of instruction type, Alves and Magro (2011) examined the effects of explicit instruction on the production of the English aspirated voiceless bilabial plosive /p/ in word-initial position by Brazilian Portuguese speakers, following the observation that most Brazilian Portuguese learners of English do not aspirate it in word-initial position (Schwartzhaupt, Alves,

---

3 According to Munro and Derwing (1995) and Derwing & Munro (1997), intelligibility refers to the extent to which listeners can correctly identify the words they hear, whereas comprehensibility refers to the listeners’ perception of the ease or difficulty with which they can make out a speaker’s meaning.
& Fontes, 2012). To learn the aspiration of the plosive /p/, the experimental group received contextualized instruction including five steps: (1) description and analysis, (2) listening discrimination, (3) controlled practice and feedback (4) guided practice and feedback instruction, (5) communicative practice and feedback. The results demonstrated that explicit instruction had a positive effect on the acquisition of aspiration, suggesting that it helped learners notice differences between aspirated /p/ in English and unaspirated /p/ in Brazilian Portuguese and provided a more effective learning opportunity in a communicative context. Similarly, Saito and Lyster (2012) measured the efficacy of form-focused instruction (FFI)\(^4\) with and without corrective feedback (CF) for improving Japanese speakers’ pronunciations of English /r/. Their findings revealed an advantage for a treatment consisting of FFI plus CF. On the other hand, the group with FFI only and the control group did not show significant improvement in the pronunciation of English /r/. These results indicate that providing feedback on learners’ mispronunciations during treatment plays a vital role in changing their L2 production performance. Additionally, Abe (2011) examined the effect of FFI on the learning of weak forms in English by Japanese EFL learners in a regular classroom setting and found that the form-focused group which received explicit instruction showed a large amount of improvement and maintained their gains in the delayed post-test, whereas the non-form-focused instruction group demonstrated small improvement only at the immediate post-test.

These previous studies regarding the role of instruction in L2 pronunciation have some limitations, in that they have exclusively used explicit instruction in a laboratory setting or in a classroom setting. Thus, other learning platforms such as computer-based pronunciation tools and web-based perception training programs should be further developed for L2 learners who may have individual learning paces.

There are few studies that directly compare the effects of implicit and explicit instruction on the perception of L2 sounds. It should be noted that the term *attention*, rather than *instruction*, has been invoked in discussions of the perceptual learning of non-native sounds (Francis & Nusbaum, 2002; Iverson & Kuhl, 1995; Pisoni et al., 1994).

---

\(^4\) Form-focused instruction is defined as “any pedagogical effort which is used to draw the learners’ attention to language” (Spada, 1997)
Attention has been demonstrated to play a role in L2 perception, because learners are able to directly attend to their input and have a better chance of noticing sounds with instruction. This leads to more rapid establishment of L2 sounds, which in turn can be reinforced through sufficient L2 input. That is to say, learning L2 sounds through explicit instruction may be considered in terms of learning to focus one’s attention on target sounds in the area of L2 speech acquisition. To date, only a handful of perceptual training studies have examined the effects of explicit instruction which draws learners’ attention to target sounds (Carlet, 2017; Carlet & Cebrian, 2015; Nozawa, 2015; Pederson & Guion-Anderson, 2010). These studies provided mixed findings about whether learners’ ability to perceive non-native vowels and consonants is affected by their attention to target sounds. For example, Pederson and Guion-Anderson (2010) trained two monolingual English speaker groups, a vowel-oriented group and a coda-oriented group, on Hindi vowels and initial stops with the same stimuli. They demonstrated that identification training only promoted improvement for trained/attended sounds, not untrained/unattended sounds, suggesting that consciously directed attention during training facilitates learning of non-native segments. It is important to note that participants in this study were monolingual speakers, not L2 learners. Therefore, the nature of learning may be different from L2 acquisition.

Expanding the study of attention effects to L2 learners, Carlet (2017) and Carlet and Cebrian (2015) looked at the effects of training and attention on English vowels and initial and final stops by Spanish/Catalan bilingual learners of English. The same stimuli were used for both a vowel-trained group and a consonant-trained group within the two types of perceptual tasks, identification and discrimination. Results showed significant effects of explicit learning for both identification and discrimination tasks. However, a significant effect of implicit learning was found only in discrimination tasks, not in identification tasks. This is might be due to the fact that the identification task strictly directed learners’ attention to their target sounds, while the discrimination task might enhance within-category variability awareness.
2.6 Standard Korean, Mandarin and English phoneme inventories

2.6.1 Korean, Mandarin and English vowel inventory

Korean, Mandarin and English differ in the sizes of their vowel inventories. The traditional Korean vowel system consists of ten monophthongs, including two front rounded vowels /y, ø/ (Sohn, 2001); however, in contemporary Korean, these are realized as /wi, we/, respectively, and analyzed as diphthongs. In addition, the merger of /e-ε/ is prevalent in the Seoul dialect as well as in many others (Eychenne & Jang, 2015; Lee & Ramsey, 2000). Previous studies have therefore mostly agreed that standard Korean has seven monophthongs (Shin, 2015; Sohn, 2001). Table 1 presents these seven monophthongs and the other three vowels (in brackets) in the traditional ten-vowel system.

Table 1. Inventory of Standard Korean vowels (adapted from Shin 2015, p.6)

<table>
<thead>
<tr>
<th></th>
<th>Front unrounded</th>
<th>Front rounded</th>
<th>Back unrounded</th>
<th>Back rounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>i</td>
<td>(y) ~ (wi)</td>
<td>i</td>
<td>u</td>
</tr>
<tr>
<td>Mid</td>
<td>e</td>
<td>(ø) ~ (we)</td>
<td>Λ</td>
<td>o</td>
</tr>
<tr>
<td>Low</td>
<td>(ε)</td>
<td></td>
<td>α</td>
<td></td>
</tr>
</tbody>
</table>

Standard Mandarin has a five-vowel system, including three high vowels /i, y, u/, a mid vowel /ə/ and a low vowel /a/. Some of these vowels also have allophones that appear as variations in certain contexts. The vowel /a/ has three allophones [a, æ, ε] and the vowel /ə/ has three allophones [e, o, ɤ]. In addition, there are three glides [j, ɥ, w] which occur in syllable onset position as allophones of the three high vowels /i, y, u/, respectively (Duanmu, 2007; Lin 1989, 2007). The Mandarin vowel inventory is shown in Table 2.

Table 2. Inventory of Standard Mandarin vowels (adapted from Lin 2007, p.82)

<table>
<thead>
<tr>
<th></th>
<th>Front unrounded</th>
<th>Front rounded</th>
<th>Central</th>
<th>Back rounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>i</td>
<td>y</td>
<td></td>
<td>u</td>
</tr>
<tr>
<td>Mid</td>
<td></td>
<td></td>
<td>ə</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>a</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Compared to Korean and Mandarin, Standard Canadian English has a relatively large vowel inventory. According to Boberg (2010), there are 10 monophthongs /i, ɪ, æ, e, ɛ, ɑ, ʌ, ʊ, o, u/ in Canadian English, including four high vowels /i, ɪ, u, ʊ/, four mid vowels /e, ɛ, ʌ, o/ and two low vowels /æ, ɑ/ as shown in Table 3.

Table 3. Inventory of Standard Canadian English vowels (adapted from Labov, Ash and Boberg 2008, p. 217-220).

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>i</td>
<td></td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>ɪ</td>
<td></td>
<td>u</td>
</tr>
<tr>
<td>Mid</td>
<td>e</td>
<td>ɛ</td>
<td>ʌ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>o</td>
</tr>
<tr>
<td>Low</td>
<td>æ</td>
<td></td>
<td>ɑ</td>
</tr>
</tbody>
</table>

The consonant inventories of Korean, Mandarin and English differ in terms of both phonation types and places of articulation. In terms of obstruents, Korean has nine stops, three fricatives and three affricates, while Mandarin has six stops, five fricatives and three affricates. English has seven stops, nine fricatives and two affricates. The key differences are in the stops and affricates: Korean has a typologically marked three-way laryngeal contrast (aspirated, lenis, fortis). Standard Mandarin, on the other hand, only has a two-way laryngeal distinction between unaspirated and aspirated stops, consisting of six voiceless stops and six affricates. Additionally, English has a two-way contrast in stops and affricates: voiced and voiceless (Ladefoged & Maddieson, 1996; Lisker & Abramson, 1964). The inventory of Korean, Mandarin and English consonants are presented in Table 4, 5 and 6, respectively.

Table 4. Inventory of Standard Korean consonants

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Alveolar</th>
<th>Palato-alveolar</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop</td>
<td>p</td>
<td>pʰ</td>
<td>p’</td>
<td>k</td>
<td>kʰ</td>
</tr>
<tr>
<td>Affricate</td>
<td>tʃ</td>
<td>tʃʰ</td>
<td>tʃ’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>s</td>
<td>s’</td>
<td>s’</td>
<td>h</td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td></td>
<td>ŋ</td>
<td></td>
</tr>
<tr>
<td>Liquid</td>
<td>ɾ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Korean has a robust process of neutralization in codas, whereby word-final obstruents are phonetically realized as lenis consonants in syllable final position (Chung, 1980; Kim & Jongman, 1996; Kim-Renaud, 1974). For example, the three-way laryngeal distinction in stops (lenis, aspirated, fortis) is neutralized in favor of the voiceless lenis series in coda position (e.g., /p, pʰ, pʼ/ → [p]). Consequently, only the three lenis obstruents /p, t, k/, as well as the four sonorants /m, n, ŋ, l/ are allowed in syllable-final position (Yoo, 1996). It should also be noted that coda consonants are not released.

Table 5. Inventory of Standard Mandarin consonants

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Alveolar</th>
<th>Dental</th>
<th>Retroflex</th>
<th>Alveo-palatal</th>
<th>Velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop</td>
<td>p</td>
<td>pʰ</td>
<td>t</td>
<td>tʰ</td>
<td></td>
<td>k</td>
</tr>
<tr>
<td>Affricate</td>
<td></td>
<td></td>
<td>ts</td>
<td>tsʰ</td>
<td>tʂ tɕʰ</td>
<td>tɕʰ</td>
</tr>
<tr>
<td>Fricative</td>
<td>f</td>
<td></td>
<td>s</td>
<td>ñ</td>
<td></td>
<td>c</td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td>ŋ</td>
</tr>
<tr>
<td>Approximant</td>
<td></td>
<td></td>
<td>l</td>
<td></td>
<td></td>
<td>l</td>
</tr>
</tbody>
</table>

Table 6. Inventory of Standard Canadian English consonants

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Labiodental</th>
<th>Interdental</th>
<th>Alveolar</th>
<th>Alveo-Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop</td>
<td>p</td>
<td>b</td>
<td>t</td>
<td>d</td>
<td>k</td>
<td>g</td>
<td>?</td>
</tr>
<tr>
<td>Affricate</td>
<td></td>
<td></td>
<td>tʃ</td>
<td>dʒ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>f</td>
<td>v</td>
<td>θ</td>
<td>ð</td>
<td>s</td>
<td>z</td>
<td>h</td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td></td>
<td>n</td>
<td></td>
<td></td>
<td>ŋ</td>
<td></td>
</tr>
<tr>
<td>Liquid</td>
<td></td>
<td></td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glide</td>
<td></td>
<td></td>
<td></td>
<td>j</td>
<td>w</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

With respect to syllable structure among the three languages, Korean allows V, CV and CVC syllables, and has a broader range of licit codas /p, t, k, n, m, ŋ, l/. On the other hand, Mandarin
has a relatively restricted syllable structure: C(G)V(G) and C(G)V(N), where (G) and (N) are optional. The only coda consonants permitted in Mandarin are the two nasals /n/ and /ŋ/.

Compared to Korean and Mandarin, English has a rich coda inventory and the maximum syllable is CCCVCCCC. The syllable structure and coda restriction of the three languages are listed in Table 7.

Table 7. Syllable structure and phonotactic restrictions of Korean, Mandarin and English

<table>
<thead>
<tr>
<th></th>
<th>Korean</th>
<th>Mandarin</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syllable structure</td>
<td>CVC</td>
<td>CVN</td>
<td>CCCVCCCCC</td>
</tr>
<tr>
<td>Licit codas</td>
<td>voiceless [p, t, k]</td>
<td>-</td>
<td>voiceless [p, t, k, f, s, tʃ, ɾ, h, tʃ]</td>
</tr>
<tr>
<td></td>
<td>nasal [n, m, ŋ]</td>
<td>nasal [n, ŋ]</td>
<td>voiced [b, d, g, v, z, ʒ, ð, dʒ]</td>
</tr>
<tr>
<td></td>
<td>liquid [l]</td>
<td>liquid [ɾ, l]</td>
<td>nasal [n, m, ŋ]</td>
</tr>
</tbody>
</table>

Given the different syllable structures and inventory sizes across the three languages, as shown in this section, I investigate whether L2 learners of Korean who have different L1 language backgrounds perform differently in their perception of Korean vowels and consonants.
Chapter 3
Experiment 1: Effects of L2 experience on the perception of Korean three-way contrasts of stops and affricates by native Mandarin listeners

3.1 Introduction

Previous studies have found that accurate perception of L2 sounds that do not exist in the L1 is one of the most difficult aspects of L2 acquisition. A prominent example is the notorious difficulty L1 speakers of Japanese have differentiating the /r/-/l/ contrast in English as an L2 (Goto, 1971; MacKain et al., 1981; Yamada & Tohkura, 1992). As another case, even relatively advanced L1 Korean learners of L2 English struggle with the perception and production of the English /i/ and /ɨ/ contrast, since there is no tense/lax distinction in the Korean vowel system (Yang, 1992). However, current models of L2 acquisition, such as the SLM and PAM, argue that L2 perceptual difficulty may be overcome with sufficient L2 experience.

As cited in section 2.2, investigations that have examined the effects of L2 experience on the perception of non-native sounds have yielded inconclusive results. Some research has shown a positive effect of L2 experience, based on learners’ length of residence (LOR) in a target language, on the perception of L2 sounds and suggested that LOR is a good predictor of L2 acquisition (Best, 1995; Best & Strange, 1992; Best & Tyler, 2007; Bohn, 1995; Bohn & Flege, 1990; Broselow et al., 1998; Flege, 1995; Flege et al., 1997; Flege & MacKay, 2004; Ingram & Park, 1997; Major & Faudree, 1996 for L2 segmental property; Towell, 2002; Trofimovich & Backer, 2006 for L2 suprasegmental property). However, other studies did not find evidence of whether L2 perception is significantly influenced by L2 experience (Cebrian, 2006; Flege et al., 1994; Fullana & Mora, 2009; Imai et al., 2002). Considering contradictory evidence of the effect of L2 experience, the aim of this study is to determine the role of experience in the perceptual acquisition of L2 contrasts.

It is important to note that most previous studies are primarily focused on experience as measured by LOR in a target language country; however, comparatively little research has been carried out on L2 in a foreign language context. As Flege and Liu (2001) mentioned, simply living in a target language country is not sufficient to enable adult learners to perceive their L2 sounds more accurately. Therefore, differences in formal education must be considered when
examining the role of experience on L2 acquisition. In this paper, following Simon and D’Hulster (2012) I use the term *experience* defined in terms of the amount of formal instruction and naturalistic exposure to the L2 in a FLA context. More specifically, I examine identification and discrimination accuracy of Korean three-way laryngeal contrasts of stops and affricates by native Mandarin listeners who differ in their experience with the Korean language as measured by the amount of instruction at the University of Toronto, Canada.

In addition, this study is concerned with how and to what extent L1 transfer contributes to L2 learners’ perception by examining how Mandarin listeners, whose L1 has only a binary laryngeal contrast, acquire the Korean three-way laryngeal contrast in stops and affricates in word-initial position. Moreover, there are differences in the phonetic cues utilized to encode the laryngeal contrasts in the two languages. The majority of previous studies have reported that Korean listeners make use of both the voice onset time (VOT) of stops and the fundamental frequency (f0) of following vowels to distinguish the contrast in phrase-initial position in Seoul Korean (Chao & Chen, 2008; Cho, Jun, & Ladefoged, 2002; Kang, 2014; Kang & Guion, 2008; Kim, 2004; Kong, Kang, & Seo, 2014; Lisker & Abramson, 1964; Schertz, Cho, Lotto, & Warner, 2015; Silva, 2006). More specifically, fortis consonants have the shortest VOT, while aspirated consonants have the longest. Thus, fortis stops are principally distinguished from lenis and aspirated stops by VOT. In term of f0, lenis is the lowest, while aspirated is the highest. That is, f0 is a primary cue for Seoul speakers of Korean to distinguish lenis stops from aspirated and fortis stops; however, the realization of the stop contrast varies depending on age, gender and dialect. For example, Kang (2014) found effects of age and gender on cue weighting strategies for the aspirated-lenis stop contrast in a speech corpus in her investigation of sound change in the three-way stop distinction. In particular, older males rely primarily on VOT and younger males use both VOT and f0 in signaling the contrast. Meanwhile, younger females rely solely on the f0 difference, suggesting f0 enhancement concurrent with the loss of the VOT distinction. Overall, this work has shown that both consonantal VOT and vocalic f0 cues influence the distinction of the three-way contrast in Seoul Korean.

Mandarin Chinese listeners, on the other hand, are expected to put perceptual weight on VOT in differentiating the two-way contrast of Mandarin stops, since aspiration is a distinctive feature in Chinese (Chao & Chen, 2008; Cho & Ladefoged, 1999; Lisker & Abramson, 1964; Rochet & Yanmei, 1991).
Based on these findings, it is reasonable to expect that Mandarin listeners are more likely to discriminate stop contrasts that primarily utilize VOT as their cue. That is, distinguishing aspirated from fortis consonants in their L2 should be easier than identifying and discriminating lenis from fortis or lenis from aspirated, which relies on f0. By examining perception patterns of Korean three-way contrasts of stops and affricates by Mandarin-speaking learners of varying levels of experience, I provide empirical evidence to explain how L2 experience relates to learners’ ability to accurately perceive certain L2 sounds.

3.2 Research questions and predictions

The primary aim of Experiment 1 is to determine the effects of L2 experience on the perception of L2 consonant contrasts by examining whether experienced Mandarin listeners are more successful in perceiving Korean three-way laryngeal contrasts than inexperienced Mandarin listeners.

Question: Are more experienced Mandarin learners of Korean more likely to have better identification and discrimination accuracy in Korean three-way contrasts of stops and affricates which do not exist in their L1 than less experienced Mandarin learners of Korean?

Prediction: I hypothesize that the greater the Mandarin-speaking learners’ Korean language experience, the better their perception accuracy of Korean three-way contrasts of stops and affricates.

3.3 Method

3.3.1 Overview

Both identification and discrimination tasks were designed to assess Mandarin listeners’ ability to perceive Korean three-way contrasts in stops and affricates. Specifically, the identification task was used to investigate how Mandarin listeners who have different amounts of experience with the Korean language identify the Korean contrasts. In addition, the AX discrimination task (Carney, Widin, and Viemeister 1977; Pisoni and Lazarus, 1974) ⁵ was used to examine the effect

---

⁵ Research has shown that acoustic processing is facilitated in the AX discrimination procedure which has low memory demand relative to the ABX task which has high memory demand.
of experience on adult Mandarin listeners’ ability to discriminate phonetic similarities and differences between pairs of Korean sounds.

3.3.2 Participants

This study included 44 native speakers of Mandarin (37 females, 7 males, mean age: 20 years, range: 18-25 years), all of whom are undergraduate students learning Korean as a foreign language at the University of Toronto, Canada. All were born and raised in China except for two, who were born and raised in Taiwan. Participants were recruited and divided into three groups according to the amount of formal language instruction and exposure they had to the Korean language: The first group (henceforth referred to as the beginner group) consisted of 20 participants who completed a first-year Korean course. The second group (intermediate group) consisted of 14 participants who completed a second-year Korean course. Finally, the third group (advanced group) had 10 participants who were enrolled in either a third- or fourth-year Korean course at the time of the experiment.

The learners in the beginner group had taken Modern Korean I for two semesters and had no experience learning Korean before attending Modern Korean I. The intermediate learners had taken Modern Korean II, had learned Korean for more than four semesters. Finally, the advanced learners had learned Korean for more than six semesters, were taking either Modern Korean III or Modern Korean IV at the University of Toronto. To take Modern Korean II, III and IV, students must have either achieved the minimum grade in the prerequisite course or attended the mandatory April interview and received placement in one of the courses. Overall, both the intermediate and advanced groups of learners had more experience with Korean than the beginner groups in terms of formal Korean language instruction at the University of Toronto and amount of exposure to Korean resources such as K-pop, K-drama, and movies. Generally, Mandarin learners’ Korean-language experience and Korean media exposure (number of hours per week) are correlated, as shown in Table 8. This information was acquired based on a language background questionnaire completed at the time of testing.

Table 8. Characteristics of the four groups

<table>
<thead>
<tr>
<th>L2 experience group</th>
<th>No. of participants</th>
<th>Average age</th>
<th>Gender</th>
<th>Korean media exposure (hours/ week)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>Count</td>
<td>Age (SD)</td>
<td>Gender</td>
<td>Average Age</td>
</tr>
<tr>
<td>--------------</td>
<td>-------</td>
<td>------------</td>
<td>------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Beginner</td>
<td>20</td>
<td>19.8 (2.03)</td>
<td>16 females, 4 males</td>
<td>19.3</td>
</tr>
<tr>
<td>Intermediate</td>
<td>14</td>
<td>20.8 (1.56)</td>
<td>13 females, 1 male</td>
<td>29.3</td>
</tr>
<tr>
<td>Advanced</td>
<td>10</td>
<td>21.6 (1.51)</td>
<td>8 females, 2 males</td>
<td>32.2</td>
</tr>
<tr>
<td>Native Korean</td>
<td>13</td>
<td>28.3 (6.54)</td>
<td>7 females, 6 males</td>
<td>NA</td>
</tr>
</tbody>
</table>

Note. Standard deviations are in parentheses.

In addition, 13 native speakers of Standard Korean (7 females, 8 males, mean age: 28 years) participated as a comparison group. All are speakers of Standard Korean who were born and raised in Seoul or Gyeonggi province in South Korea and were living in Toronto, Canada at the time of the experiment. All participants provided written informed consent and were compensated $10 for their participation. They reported normal hearing and no speech disorders.

### 3.3.3 Stimuli

#### 3.3.3.1 Discrimination task

In the discrimination task, the target stimuli consisted of 48 real words: 16 minimal triplets containing word-initial stops and affricates. The stimuli were balanced by place and manner of articulation. Table 9 shows examples of the target words used in the AX task. The full set of stimuli is provided in Appendix A.

<table>
<thead>
<tr>
<th>Aspirated</th>
<th>Lenis</th>
<th>Fortis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirated</td>
<td>/pʰul/ ‘glass’</td>
<td>/pul/ ‘fire’</td>
</tr>
<tr>
<td>Alveolar</td>
<td>/tʰal/ ‘mask’</td>
<td>/tal/ ‘moon’</td>
</tr>
<tr>
<td>Velar</td>
<td>/kʰe/ ‘to dig’</td>
<td>/ke/ ‘dog’</td>
</tr>
<tr>
<td>Palato-alveolar</td>
<td>/cʰata/ ‘to kick’</td>
<td>/cata/ ‘to sleep’</td>
</tr>
</tbody>
</table>

The auditory stimuli were natural recordings by one female (the author, 39 years old) and one male (45 years old), both of whom are native speakers of Standard Seoul Korean. Each item was recorded five times using Praat (Boersma & Weenink, 2016) with a DPA 4011 cardioid
microphone, and sampled at 44.1 kHz. The stimuli were normalized to have a mean intensity of 70 dB SPL. From these recordings, the most natural and clear tokens were selected for the stimuli by the author. There were 96 total tokens (48 words each produced by 2 speakers).

### 3.3.3.2 Identification task

The stimuli for the identification task included a combination of the 16 minimal triplets of real words (16 * 3 = 48 words) used in the discrimination task and 16 minimal triplets of disyllabic nonce words (e.g., 파아 /pha.a/, 바아 /pa.a/, 뱀아 /p’a.a/, 16 * 3 = 48 words). The full set of stimuli is provided in Appendix B. Auditory stimuli for the task were recorded by the same talkers and using the same recording parameters as in the discrimination task. Clear and natural tokens were selected by the author to be used in the identification task.

### 3.3.4 Procedure

#### 3.3.4.1 Discrimination task

In the discrimination task, all combinations of lenis, aspirated and fortis pairs were tested. The order of tokens in the AB word-pair was counterbalanced. For instance, in the case of the triplet [pʰul-pul-p’ul], there were three ‘different’ AB word-pairs [pʰul-pul], [pul-p’ul], [pʰul-p’ul], and three additional AB word-pairs in the reversed order [pul-pʰul], [p’ul-pul] [p’ul-pʰul]. All combinations of ‘same’ word-pairs were also presented: [pʰul-pʰul], [pul-pul], [p’ul-p’ul]. Thus, a total of 144 word-pairs were used in the AX task (96 different word-pairs + 48 same word-pairs). On each trial, tokens were produced by different talkers and talker order was counterbalanced. For example, in an AB word-pair, if A is [pʰul] recorded by the female talker, then B is [pul] recorded by the male talker. The inter-stimulus interval was 500 ms\(^6\) and the inter-trial interval was 1000 ms.

The experiment was deployed using the OpenSesame experiment builder software (Mathôt, Schreij & Theeuwes, 2012) and conducted individually in a sound attenuated booth located in the University of Toronto Phonetics Laboratory. Participants were instructed to listen to pairs of stimuli over Sennheiser HD 280 Pro headphones and asked to determine whether the two stimuli

---

\(^6\) Researchers have found that in cross-language speech-perception experiments when the inter-stimulus interval is very short (e.g., approximately 500ms), the acoustic/ phonetic information is available and discrimination can be good for within-category non-native contrasts (Crowder; 1982; Pisoni, 1973; Werker and Logan, 1985).
they heard were the ‘same’ or ‘different’. If they thought the two sounds were the ‘same’, they were instructed to press q on the keyboard and p if ‘different’. All instructions were displayed in Mandarin for the Mandarin speakers and in Korean for the Korean speakers. See Figure 1 for the example of instructions for Mandarin participants. Participants were instructed to decide as quickly, yet accurately as possible.

Figure 1. Example of instructions for Mandarin participants in the AX discrimination task

Before the actual experiment, each participant completed a practice session to ensure familiarity with the task. The practice session had four trials which were not used in the actual experiment. All trials were pseudo-randomly presented for each participant. Each subject took approximately 20 minutes to complete the task.

3.3.4.2 Identification task

In the identification task, participants were instructed to listen to a Korean stimulus and determine whether the auditory stimulus matched one of three visual targets presented on a computer screen. For example, on a given trial, a participant might hear ‘불’ [pul]. The task was then to choose from three visually presented Korean stimuli, i.e., ‘풀’ [phul], ‘불’ [pul], ‘뿔’ [p’ul], by clicking on the corresponding number on the keyboard. This experiment lasted approximately 15 minutes. All instructions were visually provided in Mandarin for Mandarin speakers and in Korean for Korean speakers (See Figure 2 for an example of instructions for Mandarin participants).
Before the experiment, each participant completed a practice session to ensure familiarity with the task. The practice session had four trials with a separate set of real words. The identification task consisted of 96 trials: 48 real words (24 - female trials & 24 - male trials) and 48 nonce words (24 - female trials & 24 - male trials). The inter-trial interval was 1000 ms and all 96 trials were randomized for each participant. There was no time-out and participants had to respond to proceed to the next trial.

3.3.5 Statistical analysis

To perform statistical analyses of the discrimination accuracy of Korean stops and affricates by Mandarin learners of Korean, a linear mixed effects model (Baayen, Davidson, & Bates, 2008) was used in the R Statistical Environment (R Core Team, 2017). In particular, the packages used were lme4 (Bates et al., 2017), lmerTest (Kuznetsova, Christensen, & Brockhoff, 2013) and phia (De Rosario-Martinez, 2015) for post-hoc comparisons. Satterthwaite approximations were used to estimate the degrees of freedom to derive p-values. FDR corrections for multiple comparisons were used in post-hoc analyses.

In the model, d-prime (d’), a signal-detection measurement of how discriminable two stimuli are for listeners (Macmillan & Creelman, 2004), was the dependent variable with fixed effects of L2 EXPERIENCE (beginner, intermediate, advanced, native Korean), CONDITION (aspirated-fortis, fortis-lenis, aspirated-lenis) and their interaction. Dummy-coding was used in this analysis. A random by-participant intercept was included for the random effects structure of the model.

In addition, to analyze Mandarin listeners’ identification accuracy of the three Korean laryngeal categories, a mixed-effects model with a logistic link function was used. This was carried out...
using the *glmer* function in the *lme4* library of the R Statistical Environment. The model included L2 EXPERIENCE, LARYNGEAL CATEGORY and their interaction as fixed effects, with response (correct response: 1 vs. incorrect response: 0) as the dependent variable. Both L2 EXPERIENCE and LARYNGEAL CATEGORY were dummy coded. This model also included random by-subject slope for LARYNGEAL CATEGORY, and random intercepts for items and subjects.

3.4 Results

3.4.1 Discrimination task

Figure 3 presents the results of the AX discrimination task by comparing response patterns of Korean pairwise contrasts in terms of Mandarin learners’ Korean language experience.

![Figure 3. Discrimination accuracy of Korean pairwise contrasts by L2 experience](image)

*Note.* Error bars represent the standard error.

As seen in Figure 3, native Korean listeners were extremely accurate (99.3%), whereas none of the Mandarin groups of learners achieved native-speaker accuracy levels. This figure also shows, however, that adult learners’ L2 experience influences perception of the L2 contrast. Overall, Mandarin listeners in the advanced group had the highest accuracy (90.8%) followed by those in the intermediate group (88%) and then the beginner group (85.4%) in the AX task.
Figure 4 illustrates adult Mandarin learners’ sensitivity, measured in $d'$, of the Korean pairwise contrasts based on their Korean-language experience.

![Figure 4. Effects of L2 experience on the discrimination of different sound pairs of Korean stops and affricates by Mandarin listeners](image)

**Note.** Error bars represent the standard error.

In Figure 4, higher $d'$ values indicate learners’ better discrimination ability for different sound pairs of Korean stops and affricates. As seen in the figure, the increase in perceptual sensitivity from beginner learners to native Korean listeners suggests that adult learners’ L2 experience is closely related to their discrimination accuracy. This indicates that the greater the Mandarin learners’ Korean language experience, the better their discrimination of the Korean laryngeal contrast.
Figure 5. Effects of condition on the discrimination of Korean stops and affricates by Mandarin listeners

Figure 5 demonstrates that Mandarin listeners were most likely to perceive [aspirated-fortis] as different sound pairs regardless of their Korean language experience. As expected, Mandarin listeners showed high discrimination sensitivity for the contrast [aspirated-fortis], suggesting that they rely mainly on VOT differences for discriminating the contrast. Compared to this contrast, Mandarin listeners had difficulty discriminating [lenis-aspirated] and [lenis-fortis] contrasts.

Table 10, Table 11 and Table 12 show results of the statistical analysis of the discrimination of the three-way contrasts in relation to L2 experience and condition in more detail.

Table 10. Results of the mixed-effects model for the Korean three-way contrast in word-initial stops and affricates in the AX discrimination task

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>T-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>4.56128</td>
<td>0.17061</td>
<td>26.735</td>
<td>&lt; 2e-16 ***</td>
</tr>
<tr>
<td>Native Korean vs. Beginner Mandarin</td>
<td>-0.83211</td>
<td>0.21916</td>
<td>95.42719</td>
<td>0.000257 ***</td>
</tr>
<tr>
<td>Native Korean vs. Intermediate Mandarin</td>
<td>-0.68337</td>
<td>0.23693</td>
<td>-2.884</td>
<td>0.004850 **</td>
</tr>
<tr>
<td>Condition</td>
<td>t-value</td>
<td>df</td>
<td>p-value</td>
<td>Significance</td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>----</td>
<td>---------</td>
<td>--------------</td>
</tr>
<tr>
<td>Native Korean vs. Advanced Mandarin</td>
<td>-0.53336</td>
<td>0.25875</td>
<td>-2.061</td>
<td>0.041991 *</td>
</tr>
<tr>
<td>Aspirated-Fortis vs. Aspirated-Lenis</td>
<td>-0.03566</td>
<td>0.15690</td>
<td>-0.227</td>
<td>0.820630</td>
</tr>
<tr>
<td>Aspirated-Fortis vs. Fortis-Lenis</td>
<td>-0.25934</td>
<td>0.15690</td>
<td>-1.653</td>
<td>0.101309</td>
</tr>
<tr>
<td>Experience (Native Korean vs. Beginner Mandarin) * Condition (Aspirated-Fortis vs. Aspirated-Lenis)</td>
<td>-1.84105</td>
<td>0.20154</td>
<td>-9.135</td>
<td>4.96e-15 ***</td>
</tr>
<tr>
<td>Experience (Native Korean vs. Intermediate Mandarin) * Condition (Aspirated-Fortis vs. Aspirated-Lenis)</td>
<td>-1.52173</td>
<td>0.21789</td>
<td>-6.984</td>
<td>2.62e-10 ***</td>
</tr>
<tr>
<td>Experience (Native Korean vs. Advanced Mandarin) * Condition (Aspirated-Fortis vs. Aspirated-Lenis)</td>
<td>-1.11261</td>
<td>0.23795</td>
<td>-4.676</td>
<td>8.66e-06 ***</td>
</tr>
<tr>
<td>Experience (Native Korean vs. Beginner Mandarin) * Condition (Aspirated-Fortis vs. Fortis-Lenis)</td>
<td>-1.27740</td>
<td>0.20154</td>
<td>-6.338</td>
<td>5.78e-09 ***</td>
</tr>
<tr>
<td>Experience (Native Korean vs. Intermediate Mandarin) * Condition (Aspirated-Fortis vs. Fortis-Lenis)</td>
<td>-1.19161</td>
<td>0.21789</td>
<td>-5.469</td>
<td>3.04e-07 ***</td>
</tr>
<tr>
<td>Experience (Native Korean vs. Advanced Mandarin) * Condition (Aspirated-Fortis vs. Fortis-Lenis)</td>
<td>-1.00088</td>
<td>0.23795</td>
<td>-4.206</td>
<td>5.45e-05 ***</td>
</tr>
</tbody>
</table>

*Note.* Significance codes: <0.001 ‘***’, <0.01 ‘**’, <0.05 ‘*’. Reference levels for categorical variables are underlined. All factors are dummy coded. Significant results are in bold.

The results in Table 10 show a significant main effect of EXPERIENCE and a significant interaction of L2 EXPERIENCE and CONDITION, indicating that the effect of experience differs depending on the contrast pair. Post-hoc analyses of the interaction between the two factors are shown in Table 11, which focuses on the effect of L2 experience, and Table 12, focusing on the effect of condition.
Table 11. Post-hoc analysis of interaction between CONDITION and L2 EXPERIENCE

<table>
<thead>
<tr>
<th>Condition</th>
<th>L2 Experience</th>
<th>Value</th>
<th>Df</th>
<th>Chisq</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirated-Fortis</td>
<td>Native Korean-Beginner Mandarin</td>
<td>0.83211</td>
<td>1</td>
<td>14.4165</td>
<td>&lt; 0.001***</td>
</tr>
<tr>
<td>Aspirated-Fortis</td>
<td>Native Korean-Intermediate Mandarin</td>
<td>0.68337</td>
<td>1</td>
<td>8.3187</td>
<td>&lt; 0.01 **</td>
</tr>
<tr>
<td>Aspirated-Fortis</td>
<td>Native Korean-Advanced Mandarin</td>
<td>0.53336</td>
<td>1</td>
<td>4.2491</td>
<td>0.05*</td>
</tr>
<tr>
<td>Aspirated-Fortis</td>
<td>Beginner Mandarin-Intermediate Mandarin</td>
<td>-0.14874</td>
<td>1</td>
<td>0.4815</td>
<td>0.51</td>
</tr>
<tr>
<td>Aspirated-Fortis</td>
<td>Beginner Mandarin-Advanced Mandarin</td>
<td>-0.29875</td>
<td>1</td>
<td>1.5724</td>
<td>0.25</td>
</tr>
<tr>
<td>Aspirated-Fortis</td>
<td>Intermediate Mandarin-Advanced Mandarin</td>
<td>-0.15001</td>
<td>1</td>
<td>0.3469</td>
<td>0.55</td>
</tr>
<tr>
<td>Aspirated-Lenis</td>
<td>Native Korean-Beginner Mandarin</td>
<td>2.67316</td>
<td>1</td>
<td>148.7811</td>
<td>&lt; 0.001 ***</td>
</tr>
<tr>
<td>Aspirated-Lenis</td>
<td>Native Korean-Intermediate Mandarin</td>
<td>2.20509</td>
<td>1</td>
<td>86.6162</td>
<td>&lt; 0.001 ***</td>
</tr>
<tr>
<td>Aspirated-Lenis</td>
<td>Native Korean-Advanced Mandarin</td>
<td>1.64597</td>
<td>1</td>
<td>40.4666</td>
<td>&lt; 0.001 ***</td>
</tr>
<tr>
<td>Aspirated-Lenis</td>
<td>Beginner Mandarin-Intermediate Mandarin</td>
<td>-0.46807</td>
<td>1</td>
<td>4.7680</td>
<td>0.04*</td>
</tr>
<tr>
<td>Aspirated-Lenis</td>
<td>Beginner Mandarin-Advanced Mandarin</td>
<td>-1.02719</td>
<td>1</td>
<td>18.5888</td>
<td>&lt; 0.001 ***</td>
</tr>
<tr>
<td>Aspirated-Lenis</td>
<td>Intermediate Mandarin-Advanced Mandarin</td>
<td>-0.55912</td>
<td>1</td>
<td>4.8192</td>
<td>0.04 *</td>
</tr>
<tr>
<td>-------------</td>
<td>------------------------------------</td>
<td>---------------------------------</td>
<td>----------------------------------------</td>
<td>------------------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td></td>
<td>2.10951 1 92.6529 &lt; 0.001 ***</td>
<td>1.87498 1 62.6237 &lt; 0.001 ***</td>
<td>-0.23453 1 1.1970 0.30</td>
<td>-0.57527 1 5.8303 0.02 *</td>
<td>-0.34074 1 1.7898 0.23</td>
</tr>
</tbody>
</table>

*Note. Significance codes: <0.001 ‘***’, <0.01 ‘**’, <0.05 ‘*’. Reference levels for categorical variables are underlined. All factors are dummy coded. Significant results are in bold.*

Post-hoc comparisons shown in Table 11 reveal that all Mandarin learners differed significantly from the Korean native group for the [aspirated-fortis], [aspirated-lenis] and [fortis-lenis] contrasts. This suggests that overall, Mandarin-speaking learners, irrespective of Korean experience, did not reach the same level of discrimination accuracy as native Korean speakers. For the [aspirated]-[fortis] contrast, there was no significant difference among the three Mandarin groups. This result shows that adult learners’ L2 experience did not play a role in the perception of the [aspirated]-[fortis] contrast in Korean, since this pair is likely discriminable for Mandarin listeners, with the primary cue being VOT in both languages; however, for the [aspirated]-[lenis] contrast, a significant effect of L2 experience was found, suggesting that the greater their L2 experience, the better their perception of this particular contrast. Finally, for the [fortis-lenis] contrast, there was a significant difference between the beginner and advanced groups, confirming that Mandarin advanced learners were better at discriminating the contrast than beginning learners. However, no significant differences were found between the beginner and the intermediate groups or the intermediate and advanced groups.

Table 12. Post-hoc analysis of interaction between L2 EXPERIENCE and CONDITION
<table>
<thead>
<tr>
<th>L2 Experience</th>
<th>Condition</th>
<th>Value</th>
<th>Df</th>
<th>Chisq</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginner</td>
<td>Aspirated-Lenis vs. Fortis-Lenis</td>
<td>-0.33998</td>
<td>1</td>
<td>7.2235</td>
<td>0.007195 **</td>
</tr>
<tr>
<td>Beginner</td>
<td>Aspirated-Lenis vs. Aspirated-Fortis</td>
<td>1.87671</td>
<td>1</td>
<td>220.1110</td>
<td>&lt;2.2e-16 ***</td>
</tr>
<tr>
<td>Beginner</td>
<td>Fortis-Lenis vs. Aspirated-Fortis</td>
<td>1.53674</td>
<td>1</td>
<td>47.5855</td>
<td>&lt;2.2e-16 ***</td>
</tr>
<tr>
<td>Intermediate</td>
<td>Aspirated-Lenis vs. Fortis-Lenis</td>
<td>-0.10643</td>
<td>1</td>
<td>0.4956</td>
<td>0.481450</td>
</tr>
<tr>
<td>Intermediate</td>
<td>Aspirated-Lenis vs. Aspirated-Fortis</td>
<td>1.55739</td>
<td>1</td>
<td>106.1054</td>
<td>&lt;2.2e-16 ***</td>
</tr>
<tr>
<td>Intermediate</td>
<td>Fortis-Lenis vs. Aspirated-Fortis</td>
<td>1.45095</td>
<td>1</td>
<td>92.0981</td>
<td>&lt;2.2e-16 ***</td>
</tr>
<tr>
<td>Advanced</td>
<td>Aspirated-Lenis vs. Fortis-Lenis</td>
<td>0.11195</td>
<td>1</td>
<td>0.3916</td>
<td>0.531461</td>
</tr>
<tr>
<td>Advanced</td>
<td>Aspirated-Lenis vs. Aspirated-Fortis</td>
<td>1.14827</td>
<td>1</td>
<td>41.2006</td>
<td>1.374e-10 ***</td>
</tr>
<tr>
<td>Advanced</td>
<td>Fortis-Lenis vs. Aspirated-Fortis</td>
<td>1.26022</td>
<td>1</td>
<td>49.6256</td>
<td>1.861e-12 ***</td>
</tr>
<tr>
<td>Native Korean</td>
<td>Aspirated-Lenis vs. Fortis-Lenis</td>
<td>0.22368</td>
<td>1</td>
<td>2.0324</td>
<td>0.153981</td>
</tr>
<tr>
<td>Native Korean</td>
<td>Aspirated-Lenis vs. Aspirated-Fortis</td>
<td>0.03566</td>
<td>1</td>
<td>0.0517</td>
<td>0.820193</td>
</tr>
<tr>
<td>Native Korean</td>
<td>Fortis-Lenis vs. Aspirated-Fortis</td>
<td>0.25934</td>
<td>1</td>
<td>2.7321</td>
<td>0.098350</td>
</tr>
</tbody>
</table>

**Note.** Significance codes: <0.001 ‘***’, <0.01 ‘**’, <0.05 ‘*’. Significant results are in bold.

The post-hoc analysis shown in Table 12 focuses on the effects of CONDITION for each group and demonstrates that there was significant difference among the three contrast pairs for the beginner.
group. This result indicates that beginning learners have different degrees of discrimination sensitivity to the three different contrast pairs. More specifically, discriminating [aspirated-lenis] is more difficult than [fortis-lenis] and [aspirated-fortis], and discriminating [fortis-lenis] is more difficult than [aspirated-fortis] for Mandarin beginner learners of Korean. However, for both the intermediate and advanced learners, there was no significant difference between [aspirated-lenis] and [fortis-lenis] contrast pairs, suggesting that learners with more experience are equally sensitive in discriminating these two contrast pairs. Finally, no significant differences were found among the three contrast pairs for the native Korean listeners. Taken together, the post-hoc analyses reveal that L2 experience effects depend on the particular contrasts.

3.4.2 Identification task

Figure 6 provides the results of the identification task, focusing on the effect of Korean language experience on the perceptual accuracy of Korean stops and affricates.

![Figure 6. Identification accuracy of the Korean three-way categories by L2 experience](image)

*Note.* Error bars represent standard error.

Figure 6 illustrates the influence of Mandarin-speaking learners’ Korean experience on their perception of the Korean contrasts based on real and nonce words containing stops and affricates.
in word-initial position\(^7\). Native Korean listeners were extremely good at identifying the categories, while Mandarin listeners did not reach similar levels of accuracy; however, Mandarin listeners’ identification accuracy for the Korean three-way categories appeared to increase with their L2 proficiency levels. The advanced learners were more likely to correctly identify the Korean contrasts compared to the intermediate and beginner groups, suggesting that the more proficient non-native listeners become in their L2, the more their perception accuracy in the L2 improves. In particular, advanced Mandarin listeners achieved 83.5% accuracy, intermediate Mandarin listeners achieved 77.5% accuracy, and the beginner learners performed the worst in the identification task, achieving only 64% accuracy. This result revealed an improvement in identification accuracy of the Korean contrasts across the levels of L2 experience.

Turning to the perception of the Korean laryngeal types, identification accuracy results for Mandarin learners are illustrated in Figure 7.\(^8\)

![Figure 7. Distribution of response patterns of the Korean three-way contrast by Mandarin listeners](image)

As expected, aspirated sounds were most accurately identified (84.4%), whereas lenis sounds were least accurately identified (68.5%). When Mandarin learners heard lenis sounds, 28% of the time they were misidentified as aspirated and only around 3% of the time as fortis. This uneven

---

\(^7\) Although all Mandarin groups identified real words more accurately than nonce words, the chi-square test showed that this difference was not statistically significant ($\chi^2 = 2.69$, df = 1, $p = 0.10$).

\(^8\) These include both real and nonce words.
division between aspirated and fortis responses shows that Mandarin listeners perceived a
distinction between aspirated and fortis sounds, most likely due to the VOT difference between
them.

Figure 8. Distribution of response patterns of the Korean three-way contrast by L2 experience

As seen in Figure 8, Mandarin-speaking learners’ identification accuracy of the three Korean
phonation types showed improvement with more Korean experience. For example, for aspirated
sounds, 73.4% of beginner learners, 82.1% of intermediate learners and 91.9% of advanced
learners responded correctly, showing that learners with the most Korean language experience
(over three years) most accurately identified the target sounds. In terms of identification accuracy
across the three laryngeal categories, Mandarin-speaking learners identified lenis consonants
with the least accuracy, showing that they have the most difficulty distinguishing lenis from
other phonation types. Lenis consonants were misidentified as aspirated by beginner,
intermediate, and advanced learners, 42.3%, 30.4% and 26.6% of the time, respectively. Lenis
consonants were also misidentified by the groups as fortis 7%, 2.7%, and 2.5% of the time,
respectively. Mandarin-speaking learners confused lenis sounds more often with aspirated rather
than fortis sounds, suggesting that Mandarin learner groups, especially the beginner group, relied
consistently on VOT duration for perceiving the Korean contrast.

Table 13. Results of the mixed-effects model of the Korean three-way contrast in word-initial
stops and affricates in the identification task
<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Z-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>5.1380</td>
<td>0.6843</td>
<td>7.508</td>
<td>6.01e-14 ***</td>
</tr>
<tr>
<td><strong>Native Korean vs. Beginner Mandarin</strong></td>
<td>-3.6546</td>
<td>0.7585</td>
<td>-4.818</td>
<td>1.45e-06 ***</td>
</tr>
<tr>
<td><strong>Native Korean vs. Intermediate Mandarin</strong></td>
<td>--2.7395</td>
<td>0.8020</td>
<td>-3.416</td>
<td>0.000635 ***</td>
</tr>
<tr>
<td><strong>Native Korean vs. Advanced Mandarin</strong></td>
<td>-1.4507</td>
<td>0.8873</td>
<td>-1.635</td>
<td>0.102049</td>
</tr>
<tr>
<td>Aspirated vs. Fortis</td>
<td>-0.1210</td>
<td>0.8827</td>
<td>-0.137</td>
<td>0.890955</td>
</tr>
<tr>
<td>Aspirated vs. Lenis</td>
<td>-1.3215</td>
<td>0.7910</td>
<td>-1.671</td>
<td>0.094792</td>
</tr>
<tr>
<td>Experience (Native Korean vs. Beginner Mandarin) * Laryngeal type (Aspirated vs. Fortis)</td>
<td>-0.3680</td>
<td>0.9554</td>
<td>-0.385</td>
<td>0.700103</td>
</tr>
<tr>
<td>Experience (Native Korean vs. Intermediate Mandarin) * Laryngeal type (Aspirated vs. Fortis)</td>
<td>-0.2903</td>
<td>1.0041</td>
<td>-0.289</td>
<td>0.772524</td>
</tr>
<tr>
<td>Experience (Native Korean vs. Advanced Mandarin) * Laryngeal type (Aspirated vs. Fortis)</td>
<td>-0.8079</td>
<td>1.0963</td>
<td>-0.737</td>
<td>0.461203</td>
</tr>
<tr>
<td>Experience (Native Korean vs. Beginner Mandarin) * Laryngeal type (Aspirated vs. Lenis)</td>
<td>-0.1651</td>
<td>0.8733</td>
<td>-0.189</td>
<td>0.850023</td>
</tr>
<tr>
<td>Experience (Native Korean vs. Intermediate Mandarin) * Laryngeal type (Aspirated vs. Lenis)</td>
<td>-0.1978</td>
<td>0.9219</td>
<td>-0.215</td>
<td>0.830146</td>
</tr>
</tbody>
</table>
Table 14. Post-hoc analysis of interaction between LARYNGEAL CATEGORY and L2 EXPERIENCE

<table>
<thead>
<tr>
<th>Laryngeal category</th>
<th>L2 Experience</th>
<th>Value</th>
<th>Df</th>
<th>Chisq</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Aspirated</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Native Korean - Beginner Mandarin</td>
<td>0.97478</td>
<td>1</td>
<td>23.2119</td>
<td>1.451e-06 ***</td>
<td></td>
</tr>
<tr>
<td>Native Korean – Intermediate Mandarin</td>
<td>0.93932</td>
<td>1</td>
<td>11.6691</td>
<td>0.0006355 ***</td>
<td></td>
</tr>
<tr>
<td>Native Korean – Advanced Mandarin</td>
<td>0.81010</td>
<td>1</td>
<td>2.6732</td>
<td>0.1020488</td>
<td></td>
</tr>
<tr>
<td>Beginner Mandarin – Intermediate Mandarin</td>
<td>0.28597</td>
<td>1</td>
<td>2.3022</td>
<td>0.1291922</td>
<td></td>
</tr>
<tr>
<td>Beginner Mandarin – Advance Mandarin</td>
<td>0.09940</td>
<td>1</td>
<td>9.1586</td>
<td>0.0024755 **</td>
<td></td>
</tr>
<tr>
<td>Intermediate Mandarin - Advanced Mandarin</td>
<td>0.21604</td>
<td>1</td>
<td>2.7509</td>
<td>0.0971978</td>
<td></td>
</tr>
<tr>
<td><strong>Fortis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Native Korean- Beginner Mandarin</td>
<td>0.98241</td>
<td>1</td>
<td>39.7382</td>
<td>2.904e-10</td>
<td></td>
</tr>
<tr>
<td>Native Korean – Intermediate Mandarin</td>
<td>0.95390</td>
<td>1</td>
<td>20.5715</td>
<td>5.745e-06</td>
<td></td>
</tr>
<tr>
<td>Native Korean – Advanced Mandarin</td>
<td>0.90538</td>
<td>1</td>
<td>9.9578</td>
<td>0.0016017</td>
<td></td>
</tr>
<tr>
<td>Beginner Mandarin – Intermediate Mandarin</td>
<td>0.27037</td>
<td>1</td>
<td>5.0880</td>
<td>0.0240914</td>
<td></td>
</tr>
<tr>
<td>Beginner Mandarin – Advance Mandarin</td>
<td>0.14629</td>
<td>1</td>
<td>11.5822</td>
<td>0.0006659 ***</td>
<td></td>
</tr>
</tbody>
</table>

Significance codes: <0.001 ‘***’, <0.01 ‘**’, <0.05 ‘*’. Reference levels for categorical variable are underlined. Significant results are in bold.
<table>
<thead>
<tr>
<th></th>
<th>Intermediate Mandarin - Advanced Mandarin</th>
<th>0.31620</th>
<th>1</th>
<th>1.9331</th>
<th>0.1644213</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native Korean- Beginner Mandarin</td>
<td></td>
<td>0.97854</td>
<td>1</td>
<td>77.2323</td>
<td>&lt; 2.2e-16 ***</td>
</tr>
<tr>
<td>Native Korean – Intermediate Mandarin</td>
<td></td>
<td>0.94966</td>
<td>1</td>
<td>41.4947</td>
<td>1.182e-10 ***</td>
</tr>
<tr>
<td>Native Korean – Advanced Mandarin</td>
<td></td>
<td>0.93001</td>
<td>1</td>
<td>28.1502</td>
<td>1.123e-07 ***</td>
</tr>
<tr>
<td>Beginner Mandarin – Intermediate Mandarin</td>
<td></td>
<td>0.29268</td>
<td>1</td>
<td>6.7665</td>
<td>0.0092885 **</td>
</tr>
<tr>
<td>Beginner Mandarin – Advanced Mandarin</td>
<td></td>
<td>0.22568</td>
<td>1</td>
<td>10.2778</td>
<td>0.0013464 **</td>
</tr>
<tr>
<td>Intermediate Mandarin - Advanced Mandarin</td>
<td></td>
<td>0.41327</td>
<td>1</td>
<td>0.7278</td>
<td>0.3936096</td>
</tr>
</tbody>
</table>

**Note.** Significance codes: <0.001 ‘***’, <0.01 ‘**’, <0.05 ‘*’. Reference levels for categorical variables are underlined. Significant results are in bold.

In this statistical analysis, NATIVE KOREAN and ASPIRATED were used as reference levels. The results in Table 13 show that there was a significant difference between native Korean listeners and Mandarin beginner listeners as well as between native Korean listeners and Mandarin intermediate listeners for Korean aspirated stops and affricates. However, no significant difference between native Korean listeners and Mandarin advanced listeners was found for the Korean aspirated sounds. These results suggest that in terms of L2 experience, beginner and intermediate Mandarin listeners significantly differed from native Korean listeners, but advanced Mandarin listeners were as accurate as Korean listeners. In Table 14, a post-hoc analysis of interaction between L2 EXPERIENCE and LARYNGEAL CATEGORY shows that the effect of Korean-language experience differed by Korean laryngeal category with respect to identification. For example, there was a significant difference between Mandarin beginner- and advanced listeners for the identification of all Korean aspirated, fortis and lenis sounds, indicating that advanced L2 learners show higher identification accuracy than beginner L2 learners in terms of their target sounds.
3.5 Summary and discussion

This study was concerned with the effects of Korean language experience on the perception of the Korean three way-contrast in stops and affricates by native speakers of Mandarin in an FLA context. Specifically, through two perceptual experiments, I investigated how accurately Mandarin listeners, whose L1 has only a binary laryngeal contrast, are able to perceive the Korean three-way laryngeal contrast, focusing on their Korean language experience.

Both experiments showed that most Mandarin-speaking learner groups did not attain native levels of perception accuracy for the Korean three-way contrast; however, it was also observed that L2 experience influenced the listeners’ perception of non-native contrasts. The advanced Mandarin learners were better at distinguishing the three Korean categories compared to the beginner learners with the exception of [aspirated-fortis] for which the accuracy was already high, although not-native like. That is, the greater the adult learners’ experience in the L2, the better their identification and discrimination accuracy. In addition, the results of both experiments provide empirical evidence of L1 influence on L2 perception with respect to cue-weighting strategies. More specifically, in the discrimination task, Mandarin listeners had difficulty discriminating [lenis-aspirated] and [lenis-fortis] contrasts, suggesting that they rely primarily on VOT differences for discriminating laryngeal contrasts.

In the identification task, aspirated consonants were most correctly identified, while lenis stops were least correctly identified; Mandarin learners had a tendency to misidentify lenis tokens as aspirated. Since Mandarin listeners in all L2 experience groups are sensitive to VOT, the primary cue to making phonemic stop contrast distinctions in their L1, these results suggest that L1 background plays a role in L2 adult learners’ perception when distinguishing Korean laryngeal contrasts. In the current study, the perceptual difficulty from least to most difficult was aspirated, fortis, then lenis sounds, regardless of Korean experience.

This study showed that Mandarin learners of Korean are more likely to accurately perceive Korean aspirated sounds, since Korean aspirated stops and affricates are similar to Mandarin aspirated stops in which VOT is a primary cue; however, they achieved relatively lower perception accuracy for Korean lenis sounds. This could be due to the fact that Mandarin speakers do not use f0 as a cue to distinguish stops in Mandarin, therefore leading to an inability to distinguish between Korean lenis sounds and aspirated and fortis sounds. As such, it appears
that phonetic similarities and differences between the contrasts in the participants’ L1 and L2 have considerable impact on their performance in identifying and discriminating the contrasts used in Experiment 1.
Chapter 4
Experiment 2: Effects of L1 inventory size on the perception of Korean vowels and codas by native Mandarin and English listeners

4.1 Introduction

Adult learners often struggle to acquire L2 sounds to a native-like performance level. Many studies have claimed that the L1 structure in particular is a prominent factor affecting L2 sound acquisition (Broselow, 1984; Durgunoğlu, Nagy, & Hancin-Bhatt, 1993; Flege, 1995). However, more awareness of the limitations of L1 transfer in L2 acquisition has emerged over the last two decades. That is, L2 learners’ difficulty in perception and production of non-native sounds cannot always be accounted for based on an L1-L2 transfer perspective. More specifically, there are mixed findings about the impact of L1 inventory size on the performance of L2 sounds.

As mentioned in section 2.3, some studies have observed that L1 vowel inventory size plays a significant role in L2 speech perception (Bradlow, 1995; Bundgaard-Nielsen et al., 2011; Escudero et al., 2014; Flege & Wang; 1989; Fox et al., 1995; Iverson & Evans, 2007, 2009). On the other hand, a number of studies have shown that L1 inventory size is not always a good predictor of L2 perceptual difficulty (Alispahic et al., 2017; de Jong & Hao, 2015; Elvin et al., 2014). These divergent findings on the effects of L1 phoneme inventory size on the perception of L2 sounds provide motivation to examine learners’ perceptual accuracy with Korean vowels and codas in more detail.

To examine the effects of L1 inventory size on L2 perceptual learning, this study selected Korean as a target language and both Mandarin and Canadian English as L1 languages for comparison due to the following reasons: First, Standard Korean has a seven monophthong system /a, e, i, o, u, ɨ, ʌ/ including a central-close unrounded vowel /ɨ/ which Mandarin and Canadian English vowel systems do not have. Thus, it would be interesting to see to what extent both Mandarin and English listeners perceive the new vowel /ɨ/. Second, the Korean vowel inventory falls between the two languages. Mandarin has a small vowel inventory with five monophthongs /i, y, a, u, a/ (Duanmu, 2007; Lin, 2007), whereas Canadian English has a rich vowel inventory with 10 monophthongs /i, ɪ, æ, e, ə, ʌ, o, u/ (Boberg, 2008). By comparing Mandarin and Canadian English listeners’ performance of Korean vowels, this can show whether
or not Canadian English as an L1, whose vowel inventory is larger than that of the target language, Korean, facilitates the perception of Korean vowels more than Mandarin as an L1, whose vowel inventory is smaller than that of Korean. In addition to these differences in vowel inventory size, the three languages differ in the range of permissible consonants and phonotactic restrictions in coda position. The Korean consonant system consists of fifteen obstruents and four sonorants. Among the consonants, only three voiceless stops, three nasals and a liquid /p, t, k, m, n, ŋ, l/ appear at the end of a syllable. In addition, the Korean stops /p, t, k/ are obligatorily unreleased in coda position, thus it might be difficult for L2 learners to perceive stop codas accurately. Compared to the target language, Korean, Canadian English permits all obstruents, except /h/, in coda position. On the other hand, Mandarin has a very limited coda inventory having only two nasal consonants /n, ŋ/ in final position.

The goal of Experiment 2 is to determine the role of L1 inventory size in L2 perception, including both vowels and codas, which differs from previous studies that only include either vowels or consonants. In this study, an identification task is used to examine Mandarin and English learners’ perceptual accuracy of the Korean vowels and codas. The inclusion both vowels and codas can provide greater insights on the impact of L1 inventory size to the perception of L2 sounds.

4.2 Research questions and predictions

Question: Does L1 inventory size affect the perception of L2 sounds?

Prediction: As a result of the inventory size differences among Canadian English, Mandarin and Korean, I expect the Mandarin and English L1 groups to perform differently in identifying Korean vowels and codas. Based on findings supporting the predictive role of L1 inventory size in L2 perception (Bradlow, 1995; Bundgaard-Nielsen et al., 2011; Escudero et al., 2014; Flege & Wang, 1989; Fox et al., 1995; Iverson & Evans, 2007, 2009; Scholes, 1968), I hypothesize that native Canadian English listeners, whose vowel and coda inventories are larger than those of Korean, will be more successful in perceiving Korean vowels and codas than native Mandarin listeners, whose vowel and coda inventories are smaller than those of Korean.
4.3 Method

4.3.1 Overview

The identification task was designed to assess the perception accuracy of Korean vowels and codas by native English and Mandarin listeners and to investigate whether L1 inventory size influences L2 perception performance.

4.3.2 Participants

For this study, a total of 56 subjects (28 native Mandarin and 28 native Canadian English speakers) were recruited from undergraduate Korean language courses at the University of Toronto, Canada (48 females, 8 males, mean age = 21.2 years, range = 18 - 27 years). Subjects were divided into two Korean language experience groups, depending on their length of study of Korean at the University of Toronto. The inexperienced group consisted of students enrolled in either a first- or second-year Korean course and the experienced group consisted of students enrolled in either a third- or fourth-year Korean course. The Mandarin speakers were born and raised in China and moved to Canada later in their lives (26 females, 2 males, mean age = 21.1 years). Their average length of residency in Canada is 5.9 years. The native English speakers were born and raised in Canada (22 females, 6 males, mean age = 21.3 years). A comparison group included 10 native Korean speakers who were born and raised in Seoul or Gyeonggi province in South Korea but were living in Toronto, Canada at the time of the experiment (6 females, 4 males, mean age = 22.8 years). All participants completed consent forms, reported no hearing disorders, and were compensated $10 (CAD) for their participation. Participant information is provided in Table 15.

Table 15. Information about participants in the study

<table>
<thead>
<tr>
<th>L1 group</th>
<th>L2 experience</th>
<th>No. of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Inexperienced</td>
<td>18 (15 female, 3 males)</td>
</tr>
<tr>
<td></td>
<td>Experienced</td>
<td>10 (7 females, 3 males)</td>
</tr>
<tr>
<td>Mandarin</td>
<td>Inexperienced</td>
<td>18 (16 females, 2 males)</td>
</tr>
<tr>
<td></td>
<td>Experienced</td>
<td>10 (10 females)</td>
</tr>
<tr>
<td>Korean</td>
<td>NA</td>
<td>10 (6 females, 4 males)</td>
</tr>
</tbody>
</table>
4.3.3 Stimuli

The vowel stimuli consisted of 96 monosyllabic Korean words including the four vowels /i, o, u, \( \lambda \)/, which are relatively difficult to acquire for Mandarin learners of Korean (Kim, 2010, 2016). For the coda auditory stimuli, 150 monosyllabic Korean words were used: 105 stimuli containing three stop codas /p, t, k/ and 45 stimuli containing three nasal codas /n, m, \( \eta \)/. The full sets of stimuli for vowels as well as stop and nasal codas are presented in Appendix C, D and E, respectively. Table 16 summarizes all stimuli used in the study for the vowel and coda identification tasks.

<table>
<thead>
<tr>
<th>Target sounds</th>
<th>Monosyllabic words (C(_1)VC(_2))</th>
<th>No. of stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowel identification</td>
<td>/i, o, u, ( \lambda )/</td>
<td>96 (3 * 4 *8)</td>
</tr>
<tr>
<td></td>
<td>C(_1)/k, s, no onset/ + V /o, u, ( \lambda ), i/ + C(_2)/p, t, k, n, m, ( \eta ), l, no coda/</td>
<td></td>
</tr>
<tr>
<td>Stop coda identification</td>
<td>/p, t, k/</td>
<td>105 (5 * 7 * 3)</td>
</tr>
<tr>
<td></td>
<td>C(_1)/k, t, p, s, tj/ + V /a, e, i, o, u, i, ( \lambda )/ + C(_2)/p, t, k/</td>
<td></td>
</tr>
<tr>
<td>Nasal coda identification</td>
<td>/n, m, ( \eta )/</td>
<td>45 (5 * 3 *3)</td>
</tr>
<tr>
<td></td>
<td>C(_1)/k, t, p, s, tj/ + V/a, i, u/ + C(_2)/n, m, ( \eta )/</td>
<td></td>
</tr>
</tbody>
</table>

All stimuli for the Korean vowel and coda identification tasks were recorded by a male native Korean speaker (45 years old). Each item was recorded five times with a DPA 4011 cardioid microphone at a sampling rate of 44.1 kHz. From these recordings, the most natural and clear tokens were selected for the stimuli by the author.

4.3.4 Procedure

All participants completed both the vowel and coda identification tasks. The coda identification task was composed of two separate sub-tasks for stop codas and nasal codas. All participants completed stop coda identification followed by nasal coda identification. The order of presentation between the vowel and coda identification tasks was counterbalanced across participants.

Participants were presented with numbered Korean monosyllabic words on a computer screen. They were then instructed to listen to a Korean stimulus over headphones and identify the sound
they heard by pressing the corresponding number on the keyboard. For example, in the stop coda identification task, a participant would hear ‘sip’ [십]. The task was then to choose from three visually presented stimuli, e.g., sip [십], sit [실피], sik [식], by pressing the corresponding number, 1, 2, or 3 on the keyboard, as seen in Figure 9. They were asked to respond as accurately and quickly as possible. All instructions were provided in Mandarin for the Mandarin speakers, in English for the English speakers and in Korean for the Korean speakers. The experiment was deployed using PsychoPy (Peirce, 2007) and conducted in a sound attenuated booth located in the University of Toronto Phonetics Laboratory.

![PsychoPy](image.png)

Figure 9. Example of stop coda identification trial

Before the actual experiment, each participant completed a practice session to become familiarized with the identification task. All trials were randomly presented for each participant. There was no time-out and participants had to choose a stimulus in order to move on to the next stimulus.

### 4.3.5 Statistical analysis

To determine whether L1 inventory size influences L2 perception, statistical analysis was performed using a mixed-effects logistic regression model (Baayen et al., 2008) including the
lme4 package (Bates et al., 2017) in R (R Core Team, 2017). Separate analyses were conducted for vowel and coda perception. For vowel perception, the model included vowel response pattern (correct vs. incorrect response) as the dependent variable and L1 (Korean, Mandarin, English) as the fixed effects predictor. For coda perception, coda response pattern (correct vs. incorrect response) was used as the dependent variable. The fixed effects structure of the model contained the factors of L1 (Korean, Mandarin, English), CODA TYPE (nasal codas, stop codas) and their interaction. The L1 was Helmert coded to compare native Korean listeners against non-Korean listeners and English against Mandarin listeners. Both models also included random intercepts for items and subjects.

4.4 Results

4.4.1 Effects of L1 inventory size on L2 vowel perception

Figure 10 illustrates identification accuracy of Korean vowels by native Mandarin, English and Korean listeners.

![Boxplots of identification accuracy of Korean vowels by L1 group](image)

Figure 10. Boxplots of identification accuracy of Korean vowels by L1 group

As seen in Figure 10, both English and Mandarin listeners did not reach the same accuracy level of Korean vowel identification as native Korean listeners; however, within the non-Korean listeners, English listeners perceived the vowels more accurately than Mandarin listeners. Specifically, English listeners identified Korean vowels with 84.05% accuracy, whereas
Mandarin listeners identified them with 79.25% accuracy, showing approximately a 5% accuracy difference between the two learner groups.

Table 17. Results of the mixed-effects model of perception accuracy of Korean vowels

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Z-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(intercept)</td>
<td>2.5797</td>
<td>0.1721</td>
<td>14.988</td>
<td>&lt; 2e-16 ***</td>
</tr>
<tr>
<td>Korean vs. non-Korean</td>
<td>-2.0014</td>
<td>0.3803</td>
<td>-5.263</td>
<td>1.42e-07 ***</td>
</tr>
<tr>
<td>Mandarin vs. English</td>
<td>-0.5992</td>
<td>0.2630</td>
<td>-2.279</td>
<td>0.0227 *</td>
</tr>
</tbody>
</table>

*Note. Significance codes: <0.001 ‘***’, <0.01 ‘**’, <0.05 ‘*’. Reference levels for categorical variables are underlined. Significant results are in bold.*

Results in Table 17 revealed that there was a main effect of L1 on L2 vowel perception. A significant difference between native Korean listeners and non-Korean listeners was found, indicating that Korean listeners were extremely accurate at identifying Korean vowels, while neither L2 listener groups achieved the native-like vowel perception accuracy. More importantly, between the non-native listener groups, English listeners outperformed Mandarin listeners, suggesting that L2 learners who have a larger vowel inventory are likely to perceive L2 vowels more accurately than L2 learners with a smaller vowel inventory. This finding is in line with my prediction that L1 vowel inventory size plays a significant role in the identification of L2 vowels.

Next, let us turn to the identification accuracy of individual Korean vowels by the three L1 groups. The mean percentages of correct identification of Korean vowels by L1 group are presented in Figure 11.
Figure 11. Identification accuracy of individual Korean vowels by L1 group

Note. Error bars represent the standard error.

Figure 11 illustrates that native Korean listeners reached 99% identification accuracy for all Korean vowels except for /u/ (89%), which was misidentified as /o/ (11%). A possible explanation of this perceptual difficulty by native Korean speakers is that /o/ is becoming increasingly approximated with /u/ by young Korean speakers (Han & Kang, 2013; Seong, 2004; Ryu, 2018a). In a production analysis of Korean vowels, Ryu (2018a) observed an extensive overlap between /o/ and /u/ by young female native Korean speakers. These findings suggest that the contrast between /o/ and /u/ poses some difficulty even for native Korean listeners due to the increasing overlap in their distribution in production.

The perceptual difficulty from the least to most difficult vowels for both English and Mandarin listeners was: ɨ > ʌ > u > o. Both English and Mandarin listeners were most likely to accurately identify Korean /i/, which is a new L2 phone for both groups. On the other hand, the back vowels /ʌ, u, o/ were least accurately identified. Between English and Mandarin listeners, English listeners were more successful at identifying them than Mandarin listeners, although they were less accurate at identifying /i/ than Mandarin listeners.
Figure 12. Distribution of response patterns for Korean vowels by English and Mandarin listeners

Figure 12 shows the response patterns for each target vowel by English and Mandarin listeners. For the target vowel /ɨ/, as mentioned earlier, Mandarin listeners were more successful in perception than English listeners\(^9\), despite the fact that Mandarin has a smaller vowel inventory than English. This might be due to the fact that L2 listeners with a smaller L1 vowel inventory have more unoccupied vowel space available to establish a new L2 phonological or phonetic category, resulting in accurate perception. For the target vowel /ɨ/, English listeners misidentified it as /u/ (8.7%) which might be explained by the acoustic similarity between the two languages. Ryu (2018a), in a cross-language acoustic comparison of Korean and English, found that English /u/ is located in the middle of the Korean /ɨ/ and /u/ in F2 in a vowel space. It is also important to note that the fronting of back vowels is an active change in process in Canadian English (Boberg, 2011). Thus, due to the fronting of Canadian English /u/, Canadian English listeners might perceive Korean /ɨ/ as /u/ and vice versa.

The two groups misidentified the target vowels /ʌ/ and /u/ as /o/ most often in comparison to the other response options. Mandarin listeners misidentified /o/ as either /u/ (11.1%) or /ʌ/ (18.7%),

\(^9\) It should be noted that for the perceptual accuracy of Korean vowel /ɨ/, there was no statistically significant difference between English and Mandarin listener groups. A post-hoc analysis of interaction between Korean individual vowels and L1 for the perception of Korean vowels is provided in Appendix I.
while English listeners misidentified /o/ mostly as /ʌ/ (18.2%), showing that misidentification patterns for /o/ differed by L1 background.

4.4.2 Effects of L1 inventory size on L2 coda perception

The mean accuracy scores for Korean coda identification by native Mandarin, English and Korean listeners are shown in Figure 13.

![Figure 13. Boxplots of identification accuracy of Korean codas by L1 group](image)

As seen in Figure 13, English listeners’ perception accuracy for both Korean nasal and stop codas was almost equivalent to that of the native Korean listeners. In addition, between Mandarin and English listeners, English listeners scored higher for both stop and nasal codas than Mandarin listeners, suggesting that L1 coda inventory size has an effect on the perception of L2 codas. Between Korean nasal and stop codas, both English and Mandarin listeners achieved higher accuracy with nasal codas than stop codas, indicating that Korean nasal codas are relatively easier for the learners to identify than stop codas. Interestingly, native Korean listeners also showed the same patterns for coda perception.

Now let’s turn to the groups’ performance in identifying individual nasal codas to further examine the perception patterns of Korean nasal codas.
Figure 14. Identification accuracy of Korean nasal codas by L1 group

Note. Error bars represent the standard error.

Figure 14 displays the three L1 groups’ accuracy rates of identifying individual nasal codas. English listeners’ accuracy rate for all nasal consonants in coda position reached native Korean listeners’ perceptual accuracy levels (above 97%). However, Mandarin listeners’ performance in nasal coda perception was strikingly lower than the English listeners. Interestingly, each nasal coda was perceived at different levels by Mandarin listeners. The Korean alveolar /n/ (89%) and bilabial /m/ (87%) were relatively easier for Mandarin listeners to perceive than the Korean velar nasal /ŋ/ (77%). The hierarchy of perception accuracy by Mandarin L2 learners is as follows: Korean coronal /n/ was the most accurately perceived, followed by /m/, which is in turn followed by /ŋ/. Interestingly, although Mandarin does not have /m/ in coda position and does have /ŋ/ in coda position, the identification accuracy rate for /m/ was higher than for /ŋ/.

Next, I turn to the identification accuracy of Korean stop codas by the three language groups.
As seen in Figure 15, Korean listeners showed different perception patterns for stop codas. The coronal consonant /t/ was the most accurately perceived, followed by /p/, then /k/. Interestingly, English listeners had similar perception patterns to Korean listeners. That is, the perception of codas was most accurate for coronal consonants followed by labials, then velars. Among the stop codas, Mandarin listeners identified Korean labial /p/ most accurately (83%) and Korean velar /k/ least accurately (56%).

Figure 16 and Figure 17 display the distribution of response patterns by English and Mandarin listeners of Korean nasal, and stop codas, respectively.
In Figure 16, it is apparent that English listeners were accurate at identifying Korean nasal sounds, but Mandarin listeners show some confusion among the three nasal sounds in syllable-final position. Mandarin listeners misidentified /m/ as /n/ approximately 7% of the time and as /ŋ/ approximately 6% of the time. In addition, the misidentification pattern for the target codas /n/ and /ŋ/ show that they were misperceived as one another.

In Figure 17, distribution of response patterns of Korean stop codas by English and Mandarin listeners.
Figure 17 shows that English listeners misidentified the target sound /p/ as either /t/ or /k/ at equal rates. Additionally, English listeners confused /k/ more with /p/ (10.8%) than with /t/ (4.7%). This /k/-misidentification pattern appears to contrast with the Mandarin listeners’ responses. Mandarin listeners misidentified /k/ more as /t/ (25.4%) than /p/ (19.3%), showing that while English listeners were most confused between Korean /k/ and /p/, Mandarin listeners were most confused between Korean /k/ and /t/.

To confirm whether the listeners’ L1 coda inventory and coda types play a fundamental role in L2 speech perception, a mixed-effects logistic regression model was conducted. The results are shown in Table 18.

Table 18. Results of the mixed-effects model of perception accuracy of Korean codas

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Z-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>3.9175</td>
<td>0.2559</td>
<td>15.311</td>
<td>&lt; 2e-16 ***</td>
</tr>
<tr>
<td>Korean vs. Non-Korean</td>
<td>-2.1162</td>
<td>0.5382</td>
<td>-3.932</td>
<td>8.44e-05 ***</td>
</tr>
<tr>
<td>Mandarin vs. English</td>
<td>-2.1759</td>
<td>0.2587</td>
<td>-8.409</td>
<td>&lt; 2e-16 ***</td>
</tr>
<tr>
<td>Nasal codas vs. Stop codas</td>
<td>-1.5112</td>
<td>0.2667</td>
<td>-5.666</td>
<td>1.46e-08 ***</td>
</tr>
<tr>
<td>L1 (Korean vs. Non-Korean) * Coda type (Nasal codas vs. Stop codas)</td>
<td>1.0225</td>
<td>0.5122</td>
<td>1.996</td>
<td>0.045922 *</td>
</tr>
<tr>
<td>L1 (Mandarin vs. English) * Coda type (Nasal codas vs. Stop codas)</td>
<td>0.7317</td>
<td>0.2173</td>
<td>3.367</td>
<td>0.000759 ***</td>
</tr>
</tbody>
</table>

Note. Significance codes: <0.001 ‘***’, <0.01 ‘**’, <0.05 ‘*’. Reference levels for categorical variables are underlined. Significant results are in bold.

The results in Table 18 show main effects of L1 and CODA TYPE on L2 coda perception. There was a significant difference in identification accuracy between Korean listeners and non-Korean listeners, such that Korean listeners were more accurate at identifying Korean codas than non-
Korean listener groups as a whole. In addition, English listeners, whose language has a rich coda system, outperformed Mandarin listeners, whose language has a limited coda inventory. This finding is consistent with the hypothesis that L1 coda inventory size plays an important role in the identification of L2 codas. In terms of the effects of coda type on L2 perception, a significant difference between nasal and stop codas was found. This indicates that stop codas are more difficult for listeners to identify than nasal codas. There were also significant interactions between CODA TYPE and L1. Post-hoc analysis of this interaction is shown in Table 19.

Table 19. Post-hoc analysis of interaction between CODA TYPE and L1

<table>
<thead>
<tr>
<th>Coda type</th>
<th>L1</th>
<th>Value</th>
<th>Df</th>
<th>Chisq</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasal codas</td>
<td>Korean vs. Mandarin</td>
<td>0.96099</td>
<td>1</td>
<td>34.9394</td>
<td>3.401e-09 ***</td>
</tr>
<tr>
<td></td>
<td>English vs. Korean</td>
<td>0.73657</td>
<td>1</td>
<td>3.3134</td>
<td>0.06872</td>
</tr>
<tr>
<td></td>
<td>English vs. Mandarin</td>
<td>0.10193</td>
<td>1</td>
<td>70.7173</td>
<td>&lt;2.2e-16 ***</td>
</tr>
<tr>
<td>Stop codas</td>
<td>Korean vs. Mandarin</td>
<td>0.86006</td>
<td>1</td>
<td>48.1828</td>
<td>3.883e-12 ***</td>
</tr>
<tr>
<td></td>
<td>English vs. Korean</td>
<td>0.59185</td>
<td>1</td>
<td>1.9634</td>
<td>0.16115</td>
</tr>
<tr>
<td></td>
<td>English vs. Mandarin</td>
<td>0.19090</td>
<td>1</td>
<td>63.1462</td>
<td>1.919e-15 ***</td>
</tr>
</tbody>
</table>

*Note.* Significance codes: <0.001 ‘***’, <0.01 ‘**’, <0.05 ‘*’. Reference levels for categorical variables are underlined. Significant results are in bold.

For the perception of both Korean nasal and stop codas, there were significant differences between native Korean and Mandarin listeners as well as between Mandarin and English listeners as shown in Table 19. These results confirm that native Korean listeners outperformed Mandarin listeners in the identification of Korean nasal and stop codas, as expected. Between Mandarin-speaking and English-speaking learners, the analysis revealed that English-speaking learners showed better identification abilities for both Korean nasal and stop codas than Mandarin-speaking learners. On the other hand, there was no significant difference between native Korean and English listeners, showing that English listeners identify Korean codas as accurately as Korean listeners.
4.5 Summary and discussion

This study focused on whether L1 phoneme inventory size affects L2 perception by comparing English and Mandarin listeners’ identification patterns for Korean vowels and codas. With respect to the influence of L1 vowel inventory, English listeners were better at identifying Korean vowels /ʌ, u, o/ than Mandarin listeners, suggesting that individuals with a larger L1 vowel inventory have an advantage in L2 vowel perception. However, for the perception accuracy of Korean vowel /ɨ/ which is a new phone to both Mandarin and English listeners, there was so significant difference between the two language groups. For this interesting perception pattern, there are two possible explanations. First of all, as the SLM predicts, this may be due to the lack of space within a larger L1 vowel inventory to adapt new categories. New phones can be relatively easy to incorporate into the L1 vowel inventory if the inventory is small, and therefore can be accurately perceived by learners with a small vowel inventory. Another possible explanation is the acoustic similarity between the L1 and L2 sounds. Specifically, Korean /ɨ/ and /u/ are acoustically close to Canadian English /u/ so that English listeners might be confused between Korean /ɨ/ and /u/, while Mandarin /u/ is far from Korean /ɨ/ in F2. Therefore,

The findings from the Korean vowel identification task showed that the order of perceptual difficulty for both English and Mandarin listeners is as follows: i > ɨ > u > o. Both groups of listeners most accurately perceived Korean /ɨ/, while they least accurately perceived Korean /o/. This finding shows that not all non-native vowels are equally difficult for L2 learners. In other words, there are vowel-specific differences in L2 perceptual learning. It was also found that the Korean /o/ was mostly misidentified as /ɨ/ and the Korean /u/ was mainly confused with /o/. These misidentification patterns, as mentioned previously, may be attributed to acoustic similarity between the two close vowels in the L2, as well as the influence of L1 sound assimilation to L2 sounds (Flege, 1993, 1995).

Given adult L2 learners’ perceptual difficulty with certain Korean vowels, high variability phonetic training, which provides intensive input, would be effective for L2 perception accuracy (Bradlow et al., 1997; Lively et al., 1993; Rato, 2013; Rochet & Chen, 1992; Trapp & Bohn 2002; many others).

With respect to Korean coda perception, English listeners performed on par with native speakers of Korean. Mandarin listeners, who lack voiceless obstruents and /m/ codas in their L1, had more
perceptual difficulties with Korean codas than the English L2 learners, whose L1 is much less restricted in the segments that can appear syllable finally. These results support the claim that L1 coda inventory size has a significant effect on L2 coda perception. In terms of the effects of coda type, all three groups of listeners more accurately identified Korean nasal than stop codas, which might be due to the fact that the more marked stops in coda position are more difficult to identify than less marked nasals based on the Sonority Hierarchy (Clements, 1990; Ladefoged, 1982). Another possible explanation for perceptual difficulties of Korean coda types are the effects of acoustic-perceptual salience. Jun (1995, 2004) showed that the place contrasts in nasals are perceptually weaker than the place contrasts in oral stops. Thus, one may argue that adult L2 learners are better at identifying oral stop place contrasts than nasal place contrasts, because stops are more salient than nasals. However, it should be considered that Korean final stops are not released so it may not be more salient than final nasals. As Kochetov and So (2007) found in a cross-linguistic perceptual study, released syllable-final stops are considerably more salient than unreleased stops, confirming that presence or absence of release is perceptually important for identifying consonants.

With regard to place of articulation, each nasal and stop coda was perceived at different accuracy rates by Mandarin listeners. Korean coronal /n/ was relatively easy for Mandarin listeners to perceive (around 89% accurate) compared to Korean velar nasal /ŋ/ (77%). Similarly, for the perception of Korean stop codas, Mandarin learners had the highest accuracy rate for Korean labial /p/ (83%) and the lowest accuracy rate for Korean velar stop /k/ (56%). Taken together, these results show that the Korean coronal codas /n, t/ are more likely to be perceived accurately than the Korean dorsal codas /ŋ, k/ by Mandarin listeners learning Korean as a foreign language. The same pattern regarding place of articulation was found for English listeners. This acquisition pattern shows an effect of markedness in place of articulation on L2 coda perception, which is in line with previous evidence that dorsal codas are more difficult to perceive accurately than other codas (Cardoso, 2011; Hume, 2003; Woods, 2005).

This study concludes that Mandarin L2 learners have more difficulty with the perception of Korean vowels and codas than English L2 learners, suggesting that L1 phoneme inventory size significantly influences L2 perception; however, not only L1 inventory size, but other factors, such as perceptual similarity or acoustic similarity between L1 and L2 sounds may be good predictors for L2 perceptual difficulty. Thus, to account for perceptual difficulty for Korean
vowels and consonants, it would be useful to investigate how adult English and Mandarin listeners perceive Korean sounds on the bases of perceived phonetic distance between their L1 and Korean. In addition, more evidence would be needed from a variety of languages to support the effects of L1 phoneme inventory size on the perception of L2 segments.
Chapter 5
Experiment 3: Effects of web-based auditory training on the perception of Korean vowels and codas by native Mandarin listeners

5.1 Introduction

According to the Speech Learning Model (SLM, Flege, 1995), learners’ ability to perceive L2 sounds can improve through naturalistic exposure; however, there is often not enough natural L2 exposure available for learners in a classroom setting, particularly for those who reside outside of country where the target language is spoken. Since the SLM was introduced, a variety of phonetic/perceptual training programs have been proposed to enhance L2 learners’ perception and/or production accuracy of non-native sounds. Such programs aim to provide learners with dense and targeted L2 input, and have proved to be efficient tools to improve the acquisition of L2 vowels and consonants (Bradlow et al., 1997, 1999; Hazan et al., 2005; Iverson & Evans, 2007; Pisoni et al., 1994; Strange & Dittmann, 1984; among others for L2 consonants; Aliaga-García, 2010; Iverson & Evans, 2007; Rato, 2013; among others for L2 vowels). However, these perceptual training studies have been mostly laboratory-based. Consequently, training was carried out according to a fixed schedule under controlled conditions. This limits exploration of the potential benefits of a student-centered approach, which aims to develop learner autonomy and independence by having responsibility for self-guided or self-paced learning. Moreover, although traditional language teaching emphasizes practices of L2 learning centered around the classroom and the laboratory, evolving technology continues to transform the learning experience, not only for language instruction but for pedagogy in general.

Experiments 1 and 2 indicated that Mandarin-speaking learners of Korean have particular difficulty perceiving Korean vowels /o, u, ʌ/ and velar codas /k, ƞ/. Thus, to help Mandarin learners of Korean improve their perceptual accuracy of Korean vowels and codas, I developed a web-based perceptual training program. This study explores the effects of this web-based auditory training on L2 segmental speech perception across two conditions: implicit and explicit attention.

The web-based perceptual training program developed for this dissertation was inspired by Motohashi-Saigo and Hardison (2009)’s web-based production training, in which native English
speakers learning Japanese were given audio-visual training involving waveform displays to improve Japanese geminate production accuracy. Results revealed that the audio-visual training group showed significant improvement in geminate identification, while the control group did not show any improvement. After the training, Motohashi-Saigo and Hardison (2009) interviewed their participants to determine how effective a web-based platform is for L2 learning and found that participants appreciated the flexibility of the web-based format, as well as the immediate feedback they received during training. Due to these benefits of web-based production training, it would be worthwhile to develop a web-based perceptual training program for L2 learners.

The online perceptual training program was implemented using jsPsych (De Leeuw, 2015). The program aimed to improve perception accuracy of Korean vowels and codas by L2 learners of Korean. To my knowledge, it is the first study employing web-based perceptual training on L2 segments, and there is no previous research on the perceptual training of Korean vowels and codas by Mandarin learners of Korean. The first goal of this study, therefore, is to examine whether the positive learning effects associated with laboratory-based training can also be observed in web-based training by Mandarin learners of Korean. In web-based training, the training conditions are less controlled in terms of duration and frequency of training, with the only requirement being that participants complete eight training sessions within two weeks with no more than a session each day. Otherwise, learners can complete the sessions in a comfortable environment and at a convenient time, allowing them to maintain a higher level of attention. That is, the online nature of the training tool provides a high degree of flexibility for learners, who are able to log in and complete their training sessions at any time, from anywhere.

The second goal of this study is to determine whether explicit attention is significantly superior to implicit attention in L2 perceptual learning. Following Dekeyser (1995) and Norris and Orega (2000), the present study uses the term *explicit attention* to mean that learners are instructed to attend to target sounds and that they have a conscious awareness of what is being learned during perceptual training. Conversely, the term *implicit attention* is used to mean that learners are passively exposed to target sounds and instead asked to focus on non-target sounds so that they are not aware of what is being learned during perceptual training.
For the web-based program, I utilized the HVPT method including multiple speakers to increase learners’ ability to perceive Korean sounds. The study targeted the full set of standard Korean vowels and codas: the seven monophthongs /a, e, i, o, u, ɨ, ʌ/ and the seven consonants /t, k, p, n, m, ŋ, l/ in coda position; these decisions were based on previous training studies and recommendations by Sakai and Moorman (2017)’s meta-analysis of L2 perception training studies.

The final goal of this study is to investigate whether L2 learners are able to generalize the knowledge acquired through training to novel items. As mentioned in the literature review on the effects of perceptual training, a number of lab-based training studies have found that training effects can be transferred to sounds in new phonetic contexts and this study will see whether or not the positive effect of generalization can be found in web-based HVPT training.

Taken together, this experiment shows how and to what extent learners’ identification accuracy of L2 vowels and codas improves and whether learners are able to transfer their knowledge to new sounds as a result of web-based perceptual training. In addition, it provides evidence for the relative effectiveness of implicit and explicit attention on the perception of Korean vowels and codas by Mandarin learners of Korean, allowing direct comparisons between the two different attention conditions.

5.2 Research questions and predictions

The following are the research questions and hypotheses addressed in Experiment 3.

Question 1: Do Mandarin learners of Korean improve their perceptual accuracy in identifying Korean vowels and codas as a result of web-based auditory training?

Prediction 1: I hypothesize that native Mandarin speakers who receive web-based training will show improved performance in the perception of Korean vowels and codas, compared to a control group that does not receive web-based training.

Question 2: Does explicit attention in trainings lead to greater improvement in the perception of Korean vowels and codas than implicit attention in training, even when learners are exposed to identical L2 input during training?
Prediction 2: I hypothesize with respect to vowel perception performance that the vowel-trained group with explicit attention will show more improvement than the coda-trained group that is implicitly exposed to Korean vowels during training. Similarly, the coda-trained group with explicit attention will display more improvement for Korean coda perception than the vowel-trained group which is passively exposed to Korean codas during training. In other words, improvement on trained sounds with explicit attention is expected to be greater than improvement on untrained sounds with passive exposure.

Question 3: Can learning effects gained from web-based training be successfully transferred to novel stimuli?

Prediction 3: I hypothesize that perceptual improvement obtained through web-based auditory training will transfer to novel words.

This training study will shed light not only on the role of web-based auditory training in the perception of L2 segments, but also on the effects of implicit and explicit attention on L2 speech perception. Findings obtained through this research will have valuable implications for L2 pedagogy.

5.3 Method

5.3.1 Overall design

The design of this study consists of four components: pre-test, training, post-test, and generalization test. To explore how attention facilitates the perception of L2 sounds, participants were randomly assigned to one of three groups: a vowel-trained group, a consonant-trained group, or a control group. The vowel-trained group was instructed to directly attend to Korean vowels in an identification task. Conversely, the coda-trained group was exposed to identical auditory input but was instructed to focus on Korean codas. Finally, the control group did not receive any training at all. It is important to note that the two training groups were trained on the same stimuli therefore the vowel-trained group acted as both an explicitly trained group (i.e., when they were tested on vowels) and an implicitly trained group (i.e., when they were tested on codas). Similarly, the coda-trained group acted as both an explicitly trained group (i.e., when they were tested on codas) and an implicitly trained group (i.e., when they were tested on
vowels). All participants completed the pre-test, post-test and generalization test within a 2- to 3-week window. A summary of each component in this study is presented summarized in Table 20.

Table 20. Design of the current experiment

<table>
<thead>
<tr>
<th>Phase</th>
<th>Group</th>
<th>Task</th>
<th>Feedback</th>
<th>Platform</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>all groups</td>
<td>vowel &amp; coda identification</td>
<td>no</td>
<td>PsychoPy</td>
<td>lab</td>
</tr>
<tr>
<td>Training</td>
<td>vowel-trained group</td>
<td>vowel identification</td>
<td>immediate feedback</td>
<td>jsPysch</td>
<td>online</td>
</tr>
<tr>
<td></td>
<td>coda-trained group</td>
<td>coda identification</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test</td>
<td>all groups</td>
<td>same as pre-test</td>
<td>no</td>
<td>PsychoPy</td>
<td>lab</td>
</tr>
<tr>
<td>Generalization</td>
<td>vowel-trained group &amp; control group</td>
<td>vowel identification with new stimuli</td>
<td>no</td>
<td>PsychoPy</td>
<td>lab</td>
</tr>
<tr>
<td>test</td>
<td>coda-trained group &amp; control group</td>
<td>coda identification with new stimuli</td>
<td>no</td>
<td>PsychoPy</td>
<td>lab</td>
</tr>
</tbody>
</table>

5.3.2 Participants

This experiment included native Mandarin listeners who were enrolled in beginner-level Korean language courses at Canadian universities at the time of the study, because beginning Mandarin learners of Korean had more difficulty perceiving Korean sounds than advanced Mandarin learners of Korean as shown in Experiment 1.

A total of 63 native Mandarin listeners took part in the pre-tests. Of these subjects, 11 who identified Korean vowels or codas with more than 90% accuracy at pre-test were excluded. This measure was taken in order to avoid the risk of a ceiling effect, so that all participants had room to improve through training. Additionally, 7 subjects were not retained because they did not finish all tests. The data of the remaining 45 native Mandarin learners of Korean were analyzed. This group consisted of 39 females and 6 males, ranging in age from 17 to 28, with an average age of 20.62 years old (SD = 2.41 years). These students had completed 1 to 2 years of Korean courses. All were born and raised in China (dialect not controlled). The average length of stay in
Canada was 4.66 years (SD = 2.59 years). The three groups did not differ significantly in terms of age, length of stay in Canada, or number of hours of exposure to Korean in a week (all p > 0.05). A summary of participant characteristics is shown in Table 21.

Table 21. Information about participants in the study

<table>
<thead>
<tr>
<th>Group</th>
<th>Age</th>
<th>Length of stay in Canada</th>
<th>Weekly exposure hours to Korean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowel-trained group (n = 15)</td>
<td>20.53 (2.83)</td>
<td>5.07 (3.11)</td>
<td>7.07 (5.91)</td>
</tr>
<tr>
<td>Consonant-trained group (n = 15)</td>
<td>21.33 (2.61)</td>
<td>4.73 (2.60)</td>
<td>9.13 (6.05)</td>
</tr>
<tr>
<td>Control group (n = 15)</td>
<td>20.00 (1.60)</td>
<td>4.18 (2.06)</td>
<td>11.43 (6.94)</td>
</tr>
</tbody>
</table>

*Note.* Standard deviation is in parentheses.

It is important to note that all participants may have been exposed to Korean either in or outside of the classroom during the experimental period. Participants in the two experimental groups were paid $65 (CAD) and participants in the control group received $25 (CAD) for their time. None of the participants had training in phonetics or reported hearing impairments.

5.3.3 Stimuli

For the pre-test, post-test, and training, the stimuli consisted of 49 monosyllabic consonant-vowel-consonant (CVC) words where the initial consonant was /h/, the vowel was one of the seven Korean monophthongs /a, e, i, o, u, ɨ, ʌ/, and the final consonant was one of the seven consonants /t, k, p, n, m, ŋ, l/ allowed in coda position (e.g., /han/).

For the pre-test and post-test, 98 tokens (49 stimuli * 2 speakers) were produced by two native Korean speakers (1 female, 1 male), with no repetition. The stimuli for training were recorded by four native Korean speakers (2 females, 2 males) for a total of 196 tokens (49 stimuli * 4 speakers) to create a high-variability training condition. None of these four speakers were involved in the production of the pre-test and post-test stimuli. This ensured that the perception improvement gained in training would not be specific to the tokens or speakers that participants were exposed to in training.

Stimuli for the generalization condition were made of 98 tokens (49 stimuli * 2 speakers) recorded by the same native Korean speakers from the pre-test and post-test. To investigate the
effects of generalization, the stimuli consisted of untrained Korean monosyllabic CVC words in which the initial consonant was /k/, rather than /h/. The stimuli used at pre-test, post-test and training and the stimuli used in the generalization test are provided in Appendix F and G, respectively. Table 22 summarizes the stimuli used in this study.

Table 22. Information on stimuli at pre-test, training, post-test, and generalization test

<table>
<thead>
<tr>
<th>Component</th>
<th>Stimuli</th>
<th>No. of native Korean speakers</th>
<th>No. of stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>/h-V-C /</td>
<td>4 speakers</td>
<td>196 tokens (49 × 4)</td>
</tr>
<tr>
<td>Pre-test</td>
<td>/h-V-C /</td>
<td>2 speakers</td>
<td>98 tokens (49 × 2)</td>
</tr>
<tr>
<td>Post-test</td>
<td>/h-V-C /</td>
<td>2 speakers</td>
<td>98 tokens (49 × 2)</td>
</tr>
<tr>
<td>Generalization test</td>
<td>/k-V-C /</td>
<td>2 speakers</td>
<td>98 tokens (49 × 2)</td>
</tr>
</tbody>
</table>

*Note.* $V \in \{a, \, e, \, i, \, o, \, u, \, ɨ, \, ʌ\}$, $C \in \{t, \, k, \, p, \, n, \, m, \, ŋ, \, l\}$

In total, six native Korean speakers (3 females, 3 males) recorded the stimuli for the experiment. They were all in their 20s (mean age = 23.4 years) and were either from Seoul or Gyeonggi province in South Korea. They had lived in Toronto, Canada for an average of 8 months (SD = 1.67 months) at the time of the recordings. They were asked to read a list of stimuli five times in a natural fashion and the stimuli were recorded in a sound-attenuated booth at the University of Toronto. Recordings were made with a 48-kHz sampling rate using a Sound Device 722 digital audio recorder and were normalized to a mean intensity of 60dB SPL in Praat (Boersma & Weenink, 2016). Two native speakers of Korean (1 female, the author, and 1 male, 45 years old) assessed the intelligibility of the stimuli by identifying which segment was produced. The accuracy score was 100%.

5.3.4 Procedure

Participants signed a consent form and filled out a language background questionnaire\(^{10}\). All participants took the pre-test and were then scheduled for subsequent training sessions and the post-test.

\(^{10}\) The language background questionnaire is presented in Appendix J.
5.3.4.1  Pre-test and post-test

The aim of the pre-test and post-test design is to gauge Mandarin learners’ initial identification accuracy of Korean vowels and codas, as well as to observe improvements in perception as a result of training. The pre-test and post-tests were identical, consisting of a seven-alternative forced choice task for Korean vowels and codas. Both tests were carried out individually at the phonetics laboratory of the University of Toronto. The experiment began with a short practice session of four trials, during which participants received feedback to familiarize themselves with the task. Participants were given no feedback regarding the accuracy of their response during the pre-test and post-test. The stimuli for these practice trials (/a, i, u, i/ for vowels, /ak, ap, am, al/ for codas) were distinct from those used in the main test.

The seven-alternative forced choice task was deployed using PsychoPy. The participants were instructed to listen to words through headphones and identify one of seven target phonemes that matched the auditory stimulus. The phoneme choices were presented in Hangul, the Korean alphabet, on a computer screen. Participants pressed the corresponding numbers 1, 2, 3, 4, 5, 6 or 7 on the keyboard. For example, participants in an explicitly trained group for Korean vowels heard [han] and were asked to identify the vowel (target response: <a>), while participants in an explicitly trained group for Korean codas were asked to identify the coda (target response: <n>). There was no time limit and participants had to respond in order to proceed to the next trial. In total, 96 trials were presented to each participant in random order to prevent ordering effects. These tests took about 10 minutes to complete. In order to evaluate the degree of improvement, the three groups were given the same tests with identical stimuli in both the pre-test and post-test.

5.3.4.2  Training

The training program consisted of eight sessions of high-variability phonetic training. The training was delivered through an online format programmed with jsPsych (De Leeuw, 2015), which is a JavaScript library for running experiments in a web browser. This web-based training was designed for learners to improve their perception of Korean sounds through immediate feedback, by accessing the training website at any time of the day in a comfortable setting where they could pay attention.
After completing the pre-test in the lab, learners in the two training groups were given both verbal and written instructions in English by the experimenter (the author) or two research assistants who are also native speakers of Korean (1 female, 1 male). They were told how to securely access and conduct the online training website using their own ID and password. The entire 8 training sessions were completed within two weeks. There is no more than one training session per day, following Earle and Myers (2015) which suggested that sleep promotes generalization in identification of non-native speech sounds.

As mentioned previously, the two experimental training phases were identical with respect to the amount of training time, the auditory stimuli input, and the platform; however, the phases crucially differed with regard to the emphasis on their target sounds during training. This was done to explore the effects of attention type. Specifically, the vowel-trained group was explicitly instructed to attend to Korean vowels while receiving exposure to the training set, whereas the coda-trained group was asked to focus on Korean codas, as seen in Figure 18\textsuperscript{11}.

![Figure 18. Examples of web-based training. Left side: Korean coda identification task, right side: Korean vowel identification task](image)

\textit{Note:} IPA symbols were not included in the actual experiment.

Both experimental groups followed the same procedure in each training session: (1) Read instructions written in Mandarin on the screen; (2) Listen to seven monosyllabic sounds which contain the target sounds for each group (/ɑ, e, i, o, u, ɨ, ʌ/ for the vowel-trained group, /am, an, an, ap, ak, al/ for the coda-trained group) associated with the corresponding written form in Korean.

\textsuperscript{11} It should be note that there is a one-to-one correspondence between phonems and graphemes in Korean.
Korean (/a, e, i, o, u, ɨ/ for the vowel-trained group, /m, n, ŋ, p, t, k, l/ for the coda-trained group) on the screen prior to the beginning of each session; (3) Complete an identification task.

During training, trial-by-trial performance feedback was provided, which is known to facilitate learning (Logan & Pruitt, 1995). Empirical evidence from training studies by Ellis (2005) reveals that feedback is an efficient way of promoting rapid learning. As such, this study adopted an immediate feedback approach to enable learners to consistently attend to the target stimuli and improve their performance. More specifically, participants received immediate feedback in Mandarin after each trial as to whether their answer was “correct” or “incorrect”. If the answer they clicked on was correct, the message “correct” appeared in green, whereas if the answer was incorrect, the message “incorrect” appeared in red on the screen as seen in Figure 19. They were required to make a selection again until they answered correctly.

Figure 19. Examples of immediate feedback on the training website

To keep participants engaged and motivated in undertaking the training, they were asked to write down the token number that appeared on the right corner of the screen for all items that they answered incorrectly in an instruction form. The participants submitted the instruction form before they took part in the post-test. See a sample instruction form for online perceptual training in Appendix H.

The order of items was randomized for each participant and for each session. Each session was self-paced and typically took approximately 20 minutes to complete, with little variation between participants. According to their self-reported forms, the participants completed all sessions in a quiet place, such as at home or a library, using their earphones.
5.3.4.3 Generalization test

The generalization test was designed to assess the degree to which the effects of training generalized to novel /kVC/ items. It was administered immediately after the post-test which used /hVC/ items. It should be noted that the vowel-trained group took a generalization test for Korean vowels, while the coda-trained group took a generalization test for Korean codas, and no feedback was provided.

5.3.5 Statistical analysis

Improvement in vowel and coda perception through web-based perceptual training was analyzed based on the rates of correct identification in the pre-test, post-test, online training, and generalization test. Separate analyses were conducted for vowel and coda perception.

A mixed-effects logistic model with the `glmer` function of the `lme4` package (Bates et al., 2017) in R (R Core Team, 2017) was used with response pattern (correct vs. incorrect response) as the dependent variable. TEST (pre-test, post-test, generalization), GROUP (vowel-trained group, coda-trained group, control group), and their interaction were used as fixed effects. TEST was dummy-coded and GROUP was Helmert-coded to (i) compare the vowel-trained and the coda-trained groups and (ii) compare the two experimental groups against the control group. Intercepts for subjects and items, as well as by-subject random slopes for TEST, were added as random effects.

Another model was built to examine how Mandarin-speaking learners of Korean improved their identification accuracy throughout the eight online training sessions. TRAINING SESSION was used as a fixed effect and the response pattern was used as the dependent variable.

Finally, in order to test whether training with different types of attention affects performance differently for individual vowels and codas, I built another model including INDIVIDUAL VOWELS, INDIVIDUAL CODAS, TEST, and GROUP as fixed effects and response pattern as the dependent variable.

Factors resulting in p-values of less than 0.05 were considered significant. In the case of significant interactions, follow-up Wald Chi-square tests were performed using the `testInteractions` function in the `phia` package (De Rosario-Martinez, 2015).
5.4 Results

5.4.1 Effects of web-based auditory training on L2 perception

The effects of web-based HVPT were assessed by comparing the correct identification rates obtained during pre-test and post-test for Korean vowels and codas, as seen in Figure 20 and Table 23.

![Figure 20. Identification accuracy for Korean vowels (left side) and codas (right side) across three groups at pre-test and post-test](image)

*Note.* Error bars represent the standard error.

Table 23. Mean % scores for Korean vowels and codas at pre-test and post-test and mean % of improvement by group

<table>
<thead>
<tr>
<th></th>
<th>Vowels</th>
<th>Codas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
<td>Post-test</td>
</tr>
<tr>
<td>Vowel-trained group</td>
<td>75.03</td>
<td>88.37</td>
</tr>
<tr>
<td></td>
<td>(43.30)</td>
<td>(32.07)</td>
</tr>
<tr>
<td>Coda-trained group</td>
<td>78.71</td>
<td>82.11</td>
</tr>
<tr>
<td></td>
<td>(40.95)</td>
<td>(38.34)</td>
</tr>
<tr>
<td>Control group</td>
<td>79.86</td>
<td>80.61</td>
</tr>
<tr>
<td></td>
<td>(40.12)</td>
<td>(39.55)</td>
</tr>
</tbody>
</table>
As shown in Figure 20 and Table 23, the vowel-trained group responded correctly for 75.03% of Korean vowels at pre-test and 88.37% at post-test, indicating an improvement in their ability to identify Korean vowels by approximately 13.34%. The coda-training group, which was passively exposed to Korean vowels, identified 78.71% of the vowel tokens correctly at pre-test and 82.11% at post-test, showing a modest 3.4% improvement in their vowel perception. An increase of less than 1% in Korean vowel perception accuracy was found in the control group. Turning to the results of the coda identification task, the coda-trained group, which consciously attended to codas, accurately identified 75.71% at pre-test and 88.30% at post-test for Korean codas, improving their Korean coda perception accuracy by approximately 12.59%, whereas the vowel-trained group, which was passively exposed to Korean codas, showed approximately 2.52% improvement in coda perception. The control group showed improvement by 2.45% between pre-test and post-test.

To summarize, Mandarin-speaking learners’ Korean vowel and coda identification accuracy improved by around 13% through conscious attention to the target sounds during training, indicating that web-based perceptual training is a useful tool to improve perception of L2 target sounds. In addition, although Mandarin-speaking learners were not asked to focus on Korean vowels while conducting the coda identification task and vice versa, results showed that they nonetheless improved their perception of unattended sounds by around 3% for both tasks.

### 5.4.2 Effects of implicit and explicit training on L2 perception

To analyze whether and to what extent web-based training with different types of attention affects Mandarin-speaking learners’ perceptual ability of Korean vowels and codas, a logistic mixed-effects model was used.

<table>
<thead>
<tr>
<th>Table 24. Results of the mixed-effects model of perception accuracy of Korean vowels</th>
<th>β</th>
<th>SE</th>
<th>Z-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.91423</td>
<td>0.19675</td>
<td>9.729</td>
<td>&lt;2e-16 ***</td>
</tr>
<tr>
<td>Test (Pre-test vs. Post-test)</td>
<td>0.54022</td>
<td>0.06904</td>
<td>7.825</td>
<td>5.09e-15 ***</td>
</tr>
</tbody>
</table>
Control group vs. Two training groups  
-0.22197  0.22403  -0.991  0.322

Vowel-trained group vs. Coda-trained group  
-0.33980  0.25796  -1.317  0.188

Test (Pre-test vs. Post-test) * Group (Control group vs. Two training groups vs.)  
0.73045  0.14171  5.154  2.54e-07 ***

Test (Pre-test vs. Post-test) * Group (Vowel-trained group vs. Coda-trained)  
0.99874  0.16824  5.936  2.92e-09 ***

*Note*.
Significance codes: <0.001 ‘***’, <0.01 ‘**’, <0.05 ‘*’. Reference levels for categorical variables are underlined. Significant results are in bold.

In the statistical analyses, TEST was dummy-coded, with PRE-TEST as the reference level. This means that the non-significant effect of training group in Table 24 indicates no difference at pre-test. In other words, at pre-test, the control group and the two training groups did not differ significantly in their vowel identification accuracy. This shows that participants had a similar proficiency in perceiving Korean vowels before training. However, there was a significant main effect of TEST and a significant interaction between TEST and GROUP, indicating that the main effect of test differed by L1 group. As expected, together the vowel-trained group and coda-trained group as a whole significantly improved their perception of Korean vowels from pre-test to post-test compared to the control group, indicating that the web-based training played a significant role in producing gains in Mandarin-speaking learners’ ability to identify Korean vowels. Between the two experimental groups, the vowel-trained group improved more than the coda-trained group, indicating that training Korean vowels with explicit attention was more effective than incidental exposure to Korean vowels during training. In other words, this finding indicates that the amount of improvement in identification accuracy of L2 vowels differs by different attention conditions: the more learners are explicitly instructed to attend to their L2 vowels, the more they successfully enhance their perception of L2 vowels through training.

To check whether the perceptual improvement in Korean vowels between pre-test and post-test across the three groups is statistically significant, a post-hoc analysis was conducted. Table 25 represents the results of a post-hoc analysis of the interaction between TEST and GROUP for the training of Korean vowels.
Table 25. Post-hoc analysis of interaction between GROUP and TEST for Korean vowels.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>Value</th>
<th>Df</th>
<th>Chisq</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowel-trained group: pre-post</td>
<td>0.21703</td>
<td>1</td>
<td>108.1564</td>
<td>&lt; 2e-16 ***</td>
</tr>
<tr>
<td>Coda-trained group: pre-post</td>
<td>0.42939</td>
<td>1</td>
<td>5.9271</td>
<td>0.01491 *</td>
</tr>
<tr>
<td>Control group: pre-post</td>
<td>0.48669</td>
<td>1</td>
<td>0.2185</td>
<td>0.64020</td>
</tr>
</tbody>
</table>

*Note. Significance codes: <0.001 ‘***’, <0.01 ‘**’, <0.05 ‘*’. Significant results are in bold.

As shown in Table 25, the vowel-trained group showed significant improvement in identifying Korean vowels between pre-test and post-test, confirming that explicit training is beneficial for the perception of L2 vowels. Interestingly, the coda-trained group which was passively exposed to Korean vowels during training also showed significant perceptual improvement for Korean vowels. This finding indicates that implicit training also has an impact on the perception of L2 vowels, although the explicitly-trained group improved more than the implicitly-trained group. In contrast, the control group did not show significant improvement for Korean vowels between pre-test and post-test, demonstrating that learners who did not receive training did not improve.

Turning to the effects of training for the perception of Korean codas, the results of the logistic mixed-effects model for Korean codas are summarized in Table 26, which shows significant main effects and interactions.

Table 26. Results of the mixed-effects model of perception accuracy of Korean codas.

<table>
<thead>
<tr>
<th></th>
<th>β</th>
<th>SE</th>
<th>Z-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.79306</td>
<td>0.18366</td>
<td>9.763</td>
<td>&lt; 2e-16 ***</td>
</tr>
<tr>
<td>Test (Pre-test vs. Post-test)</td>
<td>0.55359</td>
<td>0.09131</td>
<td>6.063</td>
<td>1.34e-09 ***</td>
</tr>
<tr>
<td>Control group vs. Two trained groups</td>
<td>-0.14416</td>
<td>0.23259</td>
<td>-0.620</td>
<td>0.5354</td>
</tr>
<tr>
<td>Coda-trained group vs. Vowel-trained group</td>
<td>-0.20631</td>
<td>0.26846</td>
<td>-0.768</td>
<td>0.4422</td>
</tr>
</tbody>
</table>
Table 26 shows that at pre-test, there was no significant difference in terms of perception accuracy of Korean codas between the control group and the two experimental groups, nor between the coda-trained group and the vowel-trained group. These findings indicate that the learners had a similar identification ability of Korean codas before training. However, there were significant main effects of TEST and interactions between TEST and GROUP. Taken together, the coda-trained group and the vowel-trained group experienced greater improvement in perceiving Korean codas from pre-test to post-test than the control group, which did not receive training. Comparing the degree of improvement for Korean codas between the two experimental groups, the results show that the explicitly-trained learners of Korean codas made significantly greater gains than implicitly-trained learners for the same target sounds after training. These results suggest that the web-based perceptual training tool produced a general learning effect, but perception of L2 codas can be further enhanced by attending to the target sounds with explicit attention, rather than merely listening to the target sounds. A post-hoc analysis for the perception of Korean codas is provided in Table 27.

Table 27. Post-hoc analysis of interaction between TEST and GROUP for Korean codas

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>Df</th>
<th>Chisq</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Coda-trained group: pre-post</strong></td>
<td>0.23709</td>
<td>1</td>
<td>52.1557</td>
<td>5.127e-13  ***</td>
</tr>
<tr>
<td>Vowel-trained group: pre-post</td>
<td>0.43264</td>
<td>1</td>
<td>3.0237</td>
<td>0.08206</td>
</tr>
<tr>
<td>Control group: pre-post</td>
<td>0.44497</td>
<td>1</td>
<td>2.0870</td>
<td>0.14856</td>
</tr>
</tbody>
</table>

*Note. Significance codes: <0.001 ‘***’, <0.01 ‘**’, <0.05 ‘*’. Reference levels for categorical variable are underlined. Significant results are in bold.*
As presented in Table 27, the post-hoc analysis for Korean coda identification demonstrates that the coda-trained group, which directly focused on codas during training, significantly improved in identification accuracy between pre-test and post-test. This finding confirms that perceptual training with explicit attention is effective in improving the perception of L2 consonants in coda position. However, the vowel-trained group, which did not focus on Korean codas, did not significantly improve in Korean coda perception. This indicates that passive exposure to L2 sounds during training does not significantly affect perception accuracy of L2 codas. These findings are in contrast to the results of identification improvement for Korean vowels, suggesting that the effects of attention type in L2 perception differ by L2 segments. That is, both implicit and explicit attention are effective for the perception of L2 vowels, whereas only explicit attention is effective for the perception of L2 codas. These findings indicate that enhancing perception ability for L2 codas is more difficult than for L2 vowels.

In summary, the results of Experiment 3 revealed that the training groups that received explicit training for their target sounds showed substantial gains in identification accuracy of their target sounds, providing evidence to support the effectiveness of web-based auditory training. Moreover, learners are able to improve their perception accuracy for L2 vowels through either deliberate listening with feedback or passive exposure to L2 vowels. On the other hand, learning L2 codas through deliberate listening leads to increased perception accuracy but passive exposure does not cause any significant improvement. That is, these results demonstrated that the type of attention makes a difference in L2 learners’ improvement in phoneme identification for their target language: consciously directed attention to target sounds plays a significant role in improving learners’ accuracy for their target sounds.

### 5.4.3 Effects of implicit and explicit training on individual L2 sounds

So far, I have investigated how and to what extent web-based HVPT training with two types of attention affects the perception of Korean vowels and codas. Overall, the results indicate that web-based training has a positive effect on the perception of Korean vowels and codas. Furthermore, training with explicit attention results in greater gains in the perception of target sounds. However, these findings do not show whether individual sounds show patterns in the same way. Carlet and Cebrian (2015) tested the effectiveness of HVPT methods on the perception of English vowels and consonants by Spanish/Catalan learners of English and found
that individual vowels and consonants patterned differently. It is therefore essential to explore how the perception of individual L2 segments improves before and after training with different types of attention.

Figure 21. All groups’ identification accuracy of individual Korean vowels and codas at pre-test and post-test

*Note.* Error bars represent the standard error.

Figure 21 illustrates all groups’ average correct identification scores for individual Korean vowels and codas at both tests. As is clear from the figure, not all vowels and codas are equally difficult to identify; some segments are more difficult for Mandarin speakers to identify than others. Overall, the hierarchy of accuracy of Korean vowel perception is: $i > i > a > e > A > o > u$. The vowel $/i/$ is most accurately perceived, while the lowest scores are obtained for $/o/$. As found in Experiment 2, which examined the perception of Korean vowels and codas by L2 learners, Korean $/o, A, u/$ are the most difficult vowels for Mandarin-speaking learners to accurately perceive.

In decreasing level of perception accuracy, the codas are ordered as follows: $l > m > p > n > t > \eta > k$. The coda $/l/$ is the most accurately perceived, followed by bilabial sounds $/m, p/$, alveolar sounds $/n, t/$, and velar sounds $/\eta, k/$. The ease of perceiving $/l/$ may be due to the fact that Korean has only one liquid that appears in coda position, so it is not easily confused with other phonemes. Perception accuracy of the bilabials $/m, p/$ in coda position is above 87%, suggesting that they are also relatively easy to identify. The alveolar sounds $/n, t/$ are in the middle of the
perception accuracy scale. The most difficult to perceive correctly are the velars /ŋ, k/, which was observed in Experiment 2 as well. As discussed previously in section 4.3.2, these findings suggest that markedness in terms of place of articulation may play a role in the acquisition of L2 codas.

Figures 22 and 23 display the identification accuracy of individual Korean vowels across the three groups and improvement between pre-test and post-test, respectively.

![Figure 22](image1.png)  
**Figure 22.** Identification accuracy of individual Korean vowels at pre-test and post-test by group  
*Note.* Error bars represent the standard error.

![Figure 23](image2.png)  
**Figure 23.** Perception improvement for individual Korean vowels between pre-test and post-test

As shown in Figures 22 and 23, the vowel-trained group experienced significant improvement for all Korean vowels after training (/ɑ/: p < 0.001, /e/: p < 0.001, /i/: p < 0.001, /o/: p < 0.001, /u/: p < 0.001, /ɨ/: p < 0.001, /i/: p < 0.001), suggesting that training with explicit attention is effective in improving the identification of not only difficult L2 vowels but also other L2 vowels.
In addition, /i/ and /ɨ/ reached 100% accuracy at post-test. In the coda-trained group, which was passively exposed to vowel targets, on the other hand, only two vowels /ʌ, ɨ/ showed significant improvement in perception (/ʌ/: p < 0.05, /ɨ/: p < 0.05). There was no significant improvement for any vowels in the control group, as expected. These results confirm that training with both explicit and implicit attention for L2 vowels can improve perception performance, but the degree of improvement for individual vowels differs depending on the type of attention. Explicit attention is more effective at improving all L2 vowels than implicit attention.

Figure 24. Identification accuracy of individual Korean codas at pre-test and post-test by group
Note. Error bars represent the standard error.

Figure 25. Perception improvement of individual Korean codas between pre-test and post-test
As both Figure 24 and 25 illustrate, the coda-trained group showed significant improvement for four codas /k, t, ŋ, m/ (/k/: p < 0.001, /t/: p < 0.001, /ŋ/: p < 0.001, /m/: p < 0.05), which indicates
that explicit training helps L2 learners enhance their perception of the most difficult consonants, as well as /m/, which was one of the easier consonants in coda position. In contrast, the vowel-trained group showed some improvement only for the coda /n/ (/n/: p < 0.01). Moreover, only the coda /ŋ/ was significantly improved in the control group (/ŋ/: p < 0.05). These findings provide evidence that training with explicit attention for L2 codas is more effective than training with implicit attention.

### 5.4.4 Individual learners’ development of L2 perception

I have so far examined how and to what extent implicit and explicit attention affect the three groups’ perception of Korean vowels and codas through perceptual training. For further investigation, I explored the improvement of individual learners within each group, at pre-test, during online training, and at post-test.

![Figure 26. Individual learners’ perception development for Korean vowels (left side) and codas (right side) at pre-test and post-test](image)

*Note.* The dotted lines show decrease in perception, whereas the solid lines show increase in perception.

Figure 26 shows individual learners’ identification accuracy for Korean vowels and codas across the three groups at pre-test and post-test. For Korean vowel perception, all 15 Mandarin L2 learners (V1- V15) in the vowel-trained group improved their perception accuracy after web-based HVPT training in which they were explicitly instructed to concentrate on Korean vowels. In the coda-trained group, on the other hand, 13 out of 15 learners increased their perception of Korean vowels, and two learners, C03 and C04, decreased their perception after training in
which learners were passively exposed to Korean vowels. The figure also shows that individual learners’ amount of improvement in perception in the vowel-trained group was greater than those in the coda-trained group in terms of performance on Korean vowels. In the control group, there was substantial variation in the level of development of Korean vowel perception: nine learners increased, but six learners decreased in their perception accuracy of Korean vowels between pre-test and post-test. This indicates that many individual learners who did not receive any training for L2 sounds slightly improved.

Similar to the individual differences across the three groups for Korean vowel perception, the coda-trained group with explicit attention showed substantial gains in Korean coda perception for all learners. However, four out of 15 learners in the vowel-trained group which were implicitly exposed to Korean codas during training and four out of 15 in the control group which did not receive any training decreased their perception of Korean codas. Given the results from individual Mandarin speakers’ improvement in identification of target sounds, we can conclude that explicit attention helps individual learners perceive L2 sounds.

5.4.5 Identification performance during online training

In the previous section, I found that the explicitly-trained groups performed the best at the post-test and generalization test for their respective target sounds. Based on this result, it would be worthwhile to examine how learners develop their perception abilities for their target sounds throughout the online training. Figure 27 displays the explicitly-trained Mandarin-speaking learners’ increase in identification accuracy over the course of the eight training sessions.
As seen in Figure 27, both groups demonstrated improvement in the identification of Korean vowels and codas across the eight training sessions. Mandarin speakers started at 83.7% accuracy for Korean vowels at session 1 and reached 91.2% accuracy at the end of the training. There was a significant increase in the perception accuracy of Korean vowels between the first and second sessions (P < 0.001), followed by a stable upward trend across the remaining sessions.

With regard to the development in coda perception during online sessions, Mandarin speakers identified 79.2% of Korean codas correctly at session 1 and reached 89.1% accuracy by session 8. The improvement slope was steeper for the first three sessions than the last six sessions. This might be due to the fact that immediate feedback in the early sessions was more helpful for learners in noticing their errors and shifting attention to the relevant acoustic cues to improve the accuracy of identification of their target sounds during online training. Both the explicit training groups for vowels and codas exhibited cumulative improvement from session 1 to session 8, with increases of 7.5% in identifying vowels and 9.9% for codas.

5.4.6 Response patterns for L2 sounds

Mandarin-speaking learners’ response patterns at pre-test and post-test for each Korean vowel and coda are presented in Figures 28 and 29, respectively.
Figure 28 shows how Mandarin-speaking learners responded to Korean vowels at pre-test and post-test. Mandarin-speaking learners were more likely to accurately identify the Korean vowels /a, e, i, i/ than other vowels in the identification tasks. However, Mandarin-speaking learners have difficulty with Korean vowels /o, u, ʌ/. For the target vowel /o/, they often incorrectly selected /u/ or /ʌ/. Similarly, the target vowel /u/ was often confused with /o/, as was /ʌ/. These response patterns for difficult Korean vowels by Mandarin-speaking learners are consistent with Ryu’s previous study (2018a), which investigated how English and Mandarin-speaking learners of Korean perceive the difficult vowels /i, o, u, ʌ/ using an identification task. When looking at development in perception accuracy between pre-test and post-test, overall Mandarin-speaking learners’ perception of each vowel improved, but the response patterns for vowels at the two tests are similar in terms of which vowels they are typically confused with.
As seen in Figure 29, the consonant /l/ in coda position is the most accurately identified among the seven consonants, and Mandarin-speaking learners did not typically confuse it with other consonants in either test. In the responses for the velar nasal /ŋ/, it was most likely to be misidentified as the alveolar nasal /n/, and vice versa, in both tests. In the responses for /k/, Mandarin-speaking learners misidentified /k/ as /t/ approximately 16.8% of the time and /p/ 24.6% of the time at pre-test, while they misidentified /k/ as /t/ approximately 15.6% of the time and /p/ approximately 15.1% of the time at post-test. Overall Mandarin-speaking learners’ perception of each coda improved between pre-test and post-test. However, the response patterns for Korean codas at the two tests are similar with respect to which codas they are typically confused with, as with the response patterns for Korean vowels.

5.4.7 Effects of generalization on L2 perception

To evaluate generalizability from the training, I examined whether Mandarin speakers who received vowel and coda training with explicit attention could generalize the knowledge from web-based HVPT training in which the context /hVC/ was used, to a new context, /kVC/. This was done by examining accuracy rates for the identification of Korean vowels and codas at pre-test, post-test and generalization test. The results of these tests indicate whether learning from the /hVC/ context during perceptual training can be successfully transferred to the /kVC/ context in the generalization test. Some previous studies that have examined generalization effects have
compared generalization results to pre-test results (Nobre-Oliveira, 2007), while other studies have compared generalization results to post-test results (Carlet & Cebrian, 2015; Hazan et al., 2005; Rato, 2013). In order to provide a complete picture of the generalization effects, this study follows Carlet (2017) in comparing the generalization results to both the pre-test and post-test scores. The scores by group for correct identification of Korean vowels and codas at pre-test, post-test, and generalization test are shown in Figure 30 and Table 28.

![Figure 30](image)

**Figure 30.** Identification accuracy for Korean vowels (left side) and codas (right side) across two groups at pre-test, post-test and generalization test

*Note.* Error bars indicate standard error.

**Table 28.** Mean % scores for Korean vowels and codas at pre-test, post-test and generalization test

<table>
<thead>
<tr>
<th></th>
<th>Korean vowels</th>
<th></th>
<th>Korean codas</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
<td>Post-test</td>
<td>Generalization test</td>
<td>Pre-test</td>
</tr>
<tr>
<td>Vowel-trained group</td>
<td>75.03 (43.30)</td>
<td>88.37 (32.07)</td>
<td>88.64 (31.74)</td>
<td>75.71 (42.90)</td>
</tr>
<tr>
<td>Control group</td>
<td>79.86 (40.12)</td>
<td>80.61 (39.55)</td>
<td>85.58 (35.14)</td>
<td>79.52 (40.37)</td>
</tr>
</tbody>
</table>

*Note.* Standard deviation is in parentheses.
As presented in Table 28, the explicitly vowel-trained group showed improved vowel identification accuracy from pre-test (75.03%) to post-test (88.37%) and this increase in performance was maintained in the generalization test (88.64%). Similarly, the explicitly coda-trained group improved coda identification accuracy from pre-test (75.71%) to post-test (88.30%) and this increase in performance was maintained in the generalization test (87.28%). As shown earlier, the control group did not show greater gains from pre-test to post-test for either Korean vowels or codas. However, the control group showed different results for vowels and codas in the generalization test. Korean vowel perception scores (85.58%) were higher than Korean coda perception scores (81.02%) in the generalization test. For analyzing the effect of generalization of Korean vowels, a logistic mixed-effects model was fit, and the results are shown in Table 29.

Table 29. Results of the mixed-effects model of generalization effects of Korean vowels

<table>
<thead>
<tr>
<th></th>
<th>β</th>
<th>SE</th>
<th>Z-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.7056</td>
<td>0.2483</td>
<td>6.869</td>
<td>6.45e-12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>Test (Pre-test vs. Post-test)</td>
<td>1.2734</td>
<td>0.1382</td>
<td>9.215</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>Test (Pre-test vs. Generalization)</td>
<td>1.4381</td>
<td>0.2997</td>
<td>4.798</td>
<td>1.60e-06</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>Vowel-trained group vs. Control group</td>
<td>0.3974</td>
<td>0.2440</td>
<td>1.629</td>
<td>0.103394</td>
</tr>
<tr>
<td>Test (Pre-test vs. Post-test) * Group (Vowel-trained group vs. Control group)</td>
<td>-1.2372</td>
<td>0.1883</td>
<td>-6.571</td>
<td>4.98e-11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>Test (Pre-test vs. Generalization) * Group (Vowel-trained group vs. Control group)</td>
<td>-0.7868</td>
<td>0.2023</td>
<td>-3.889</td>
<td>0.000101</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>***</td>
</tr>
</tbody>
</table>

Note. Significance codes: <0.001 ‘***’, <0.01 ‘**’, <0.05 ‘*’. Reference levels for categorical variable are underlined. Significant results are in bold.

The statistical results in Table 29 further indicate that Mandarin listeners’ perception accuracy for Korean vowels significantly differed between pre-test and post-test as well as pre-test to generalization test, as confirmed by significant effects of TEST. These findings again show a
positive effect of web-based HVPT. Furthermore, there were significant interactions between Test and Group. The vowel-trained group made significant improvements in the identification of Korean vowels from pre-test to post-test, as well as pre-test to generalization test, compared to the control group. As expected, this shows that vowel training with explicit attention is effective, and learners who take part in the training are able to generalize the newly acquired knowledge to a new phonetic context. The results of a post-hoc test for effects of generalization of Korean vowels is shown in Table 30.

Table 30. Post-hoc analysis of interaction between Group and Test for generalization effects of Korean vowels

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>Df</th>
<th>Chisq</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowel-trained group: pre-post</td>
<td>0.21867</td>
<td>1</td>
<td>84.9232</td>
<td>&lt;2.2e-16 ***</td>
</tr>
<tr>
<td>Vowel-trained group: pre-gen</td>
<td>0.19184</td>
<td>1</td>
<td>23.0218</td>
<td>1.602e-06 ***</td>
</tr>
<tr>
<td>Vowel-trained group: post-gen</td>
<td>0.45892</td>
<td>1</td>
<td>0.3044</td>
<td>0.58116</td>
</tr>
<tr>
<td>Control group: pre-post</td>
<td>0.49094</td>
<td>1</td>
<td>0.0789</td>
<td>0.77884</td>
</tr>
<tr>
<td>Control group: pre-gen</td>
<td>0.34270</td>
<td>1</td>
<td>4.8013</td>
<td>0.02844 *</td>
</tr>
<tr>
<td>Control group: post-gen</td>
<td>0.35091</td>
<td>1</td>
<td>4.5235</td>
<td>0.03343 *</td>
</tr>
</tbody>
</table>

Note. Significance codes: <0.001 ‘***’, <0.01 ‘**’, <0.05 ‘*’. Reference levels for categorical variable are underlined. Significant results are in bold.

As presented in Table 30, the post-hoc test confirms that the vowel-trained group’s performance on Korean vowels improved significantly between the pre-test and generalization test, but not between the post-test and generalization test. Interestingly, although the control group did not make significant gains for Korean vowels from pre-test to post-test, a significant improvement was found from pre-test to generalization and post-test to generalization. More specifically, the control group had significantly higher identification scores for vowels in the generalization test (85.58%) than in the post-test (80.61%). This result might be due to the fact that the initial consonant /k/, which occurs before a target vowel, affects vowel perception. That is, perceiving Korean vowels before /k/ might be easier than before /h/, which was used in the pre-test. To account for this finding, further study is necessary to investigate whether neighboring segments influence L2 vowel perception (see Gottfried, 1984; Levy, 2009)
I turn now to the output of the logistic mixed effects model of generalization effects and the post-hoc test for Korean codas, which are provided in Table 31 and Table 32, respectively.

Table 31. Results of the mixed-effects model of generalization effects of Korean codas

<table>
<thead>
<tr>
<th></th>
<th>β</th>
<th>SE</th>
<th>Z-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.6375</td>
<td>0.2288</td>
<td>7.157</td>
<td>8.24e-13 ***</td>
</tr>
<tr>
<td>Test (Pre-test vs. Post-test)</td>
<td>1.1385</td>
<td>0.1711</td>
<td>6.653</td>
<td>2.86e-11 ***</td>
</tr>
<tr>
<td>Test (Pre-test vs. Generalization)</td>
<td>0.9484</td>
<td>0.2606</td>
<td>3.639</td>
<td>0.000274 ***</td>
</tr>
<tr>
<td>Coda-trained group vs. control group</td>
<td>0.2494</td>
<td>0.2519</td>
<td>0.990</td>
<td>0.322130</td>
</tr>
<tr>
<td>Test (Pre-test vs. Post-test) * Group (Coda-trained group vs. Control group)</td>
<td>-0.9286</td>
<td>0.2349</td>
<td>-3.953</td>
<td>7.71e-05 ***</td>
</tr>
<tr>
<td>Test (Pre-test vs. Generalization) * Group (Coda-trained group vs. Control group)</td>
<td>-0.8599</td>
<td>0.2214</td>
<td>-3.884</td>
<td>0.000103 ***</td>
</tr>
</tbody>
</table>

Note. Significance codes: <0.001 ‘***’, <0.01 ‘**’, <0.05 ‘*’. Reference levels for categorical variable are underlined. Significant results are in bold.

Table 31 shows that there were significant improvements in Korean coda perception between the pre-test and post-test and between the pre-test and generalization test, indicating that web-based HVPT training is effective in helping L2 learners improve their perception accuracy and generalize their learning to new contexts. Furthermore, the explicitly coda-trained group showed significant improvement in identification accuracy from pre-test to post-test, and from pre-test to generalization test, compared to the control group.

Table 32. Post-hoc analysis of interaction between TEST and GROUP for generalization effects of Korean codas

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>df</th>
<th>Chisq</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coda-trained group: pre-post</td>
<td>0.24260</td>
<td>1</td>
<td>44.2681</td>
<td>2.86e-11 ***</td>
</tr>
<tr>
<td>Coda-trained group: pre-gen</td>
<td>0.27922</td>
<td>1</td>
<td>13.2422</td>
<td>0.0002737 ***</td>
</tr>
</tbody>
</table>
Table 32 illustrates that the coda-trained group showed a significant increase in accuracy in the perception of Korean codas from pre-test to post-test and pre-test to generalization test, whereas the control group did not show a significant increase from pre-test to generalization test. These results indicate that explicitly instructed training is helpful for L2 learners to generalize their learning to new stimuli.

To summarize, the results of the effect of training on the generalization test showed that perceptual accuracy for the vowel-trained group was significantly different from pre-test to post-test and from pre-test to generalization for Korean vowel perception. This suggests that L2 learners can generalize their learning to new words. The control group, on the other hand, did not yield significant improvement in Korean vowel perception from pre-test to post-test. Interestingly, there was a significant improvement from pre-test to generalization and post-test to generalization, suggesting that Mandarin-speaking learners had better identification abilities for Korean vowels which appear after the consonant /k/, rather than /h/. It would be interesting to further examine how initial consonant affects perception of the following vowel. For coda perception, the result of a post-hoc comparison revealed that the coda-trained group’s improvement of Korean coda perception between pre-test to post-test and pre-test to generalization was significant. For the control group, however, no differences were found between pre-test and post-test or pre-test and generalization test. Overall, the results for generalization revealed that explicit training for both vowels and codas was able to promote strong generalization to new phonetic contexts.

5.5 Summary and discussion

This study examined the role of web-based HVPT training as well as the effectiveness of implicit and explicit attention on the perception of Korean vowels and codas by Mandarin
learners of Korean. The current study aimed to address the following research questions: (1) How and to what extent do Mandarin-speaking learners of Korean improve their perception accuracy of Korean vowels and codas through web-based training? (2) Does explicit training result in greater enhancement in identifying Korean vowels and codas compared to implicit training, even when identical auditory input is used? (3) Can Mandarin-speaking learners of Korean who receive explicit training generalize learning to new words?

In relation to the first research question, the results revealed that the perceptual identification performance of both training groups as a whole increased significantly from pre-test to post-test, in comparison with the control group, suggesting a beneficial effect of online perceptual training of L2 sounds. This positive learning effect is largely consistent with previous work showing that L2 learners are able to improve their perception ability through perceptual training (Bradlow et al., 1999; Inceoglu, 2016; Rato, 2013; many others). This finding also suggests that, as Motohashi-Saigo and Hardison (2009) recommended, online training can encourage participation and can be an additional tool for L2 learners to supplement classroom instruction and interactions with native speakers of their target languages.

The important finding regarding the second research question addresses the impact of implicit vs. explicit attention on L2 speech perception. Although the effectiveness of implicit and explicit attention is a question of long-standing interest in the study of L2 acquisition, little research has been dedicated to directly measuring the influence of the two types of attention conditions on the identification of L2 sounds. Thus, this study was primarily concerned with the relative effects of implicit and explicit attention on L2 speech perception. Specifically, I examined how implicit and explicit attention in training affect the perception of Korean vowels and codas. Results comparing the effects of explicit vs implicit attention on L2 vowel and coda perception revealed different patterns of improvement. The vowel-trained group made strongly significant improvement and the coda-trained group yielded modest but significant improvement in identifying Korean vowels. In contrast, the control group did not show any significant gains. These results suggest that learners are able to improve their perception of L2 vowels through both deliberate listening and passive exposure to the target vowels, but it is important to note that focused attention for individual L2 vowels is more effective than simple exposure to stimuli containing vowels.
The pattern for Korean codas was different from the pattern for Korean vowels. Only the coda-trained group, which explicitly focused on codas during training, achieved a significant increase in coda identification accuracy. This suggests that the more learners are attentive to their L2 codas while learning, the more their ability to identify those sounds improves. On the other hand, the vowel-trained group, which did not explicitly attend to codas, did not show significant improvement in the identification of Korean codas. This suggests that implicit attention in training is not effective for learners to improve their perception of L2 codas. Taken together, these results suggest that attention type differentially affects perception depending on the target structure and highlights the benefits of explicit over implicit training for at least some L2 sound structures.

Based on these different improvement patterns, Korean coda perception learned through unconscious attention is arguably less productive than vowel perception under the same conditions. There are two possible explanations for these asymmetrical perceptual improvements in training. First of all, I speculate that these differential results for vowels and codas might be related to their relative positions in the stimuli presented: perhaps listeners can improve their perception of segments that precede the target sounds with both explicit and implicit attention, because when they focus on their target sounds, they also listen carefully to the segments right before the target sounds. Thus, the coda-trained learners were able to enhance their perception of Korean vowels which occur before their target codas, even though they were not instructed to focus on vowels. However, if listeners tend to not focus on segments that follow their target sounds, this can account for why significant perception improvement of Korean codas did not occur in the vowel-trained group. That is, when the vowel-trained group was asked to pay attention to their target vowels during training, they may have focused only on Korean vowels and ignored the following codas, and thus their perception of codas did not significantly improve.

Another possible explanation might be acoustic salience of Korean codas which are obligatorily unreleased. Flege & Wang (1989) observed that Mandarin learners of English had great difficulty identifying English word-final stops /t, d/ in coda position, especially when the codas did not have a release burst, while they had relatively little difficulty identifying tokens of the final stops that had audible release bursts. As they found in their previous work, Mandarin learners of Korean also experience difficulty identifying Korean codas accurately and it is
difficult to improve their perception accuracy by simple exposure to codas. More attention to codas is needed to enhance the ability to identify codas correctly.

These asymmetrical attention effects on the perception of Korean vowels and codas are puzzling, so it would be worthwhile to examine them further to determine whether these patterns appear in other L2 learners of Korean, or in learners of other languages.

These findings are partially consistent with those of Carlet (2017), which compared two types of perceptual tasks, identification and discrimination, and looked at the effects of training on trained and untrained segments at pre-test, post-test and a delayed post-test. Even though she found that Spanish/Catalan speakers who were trained on English vowels improved in final stop identification, and vice versa, in the discrimination task, she did not find any improvement on untrained segments in the identification task. That is, only the discrimination trainees, not the identification trainees, improved in the untrained structures. This is in contrast to my current result wherein identification training induced implicit learning of L2 vowels. This differing result with regard to the effect of implicit attention on L2 perception might be due to the fact that my study contained the full set of L2 segments, rather than only difficult L2 segments, as Carlet (2017) used in her experiment.

In terms of the perception of individual segments, I found that perception accuracy of all Korean vowels /α, e, i, o, u, ɨ, ʌ/ and the most difficult codas /k, t, ɳ, m/ significantly improved after training with explicit attention, and the perception of vowels /ɻ, ɨ/ and coda /n/ significantly increased after training with implicit attention. These results show the advantages of using the full sets of L2 phonemes in training, as suggested by Nishi and Kewley-Port (2007, 2008), who found that training larger sets of phonemes led to greater gains than training only the most difficult subsets of phonemes.

My last research question explored whether Mandarin-speaking learners who receive perceptual training are able to generalize learning to new words. Results of the generalization tests showed that groups that were explicitly trained on target sounds were able to generalize their learning to the new items. This indicates that the perceptual online training was successful in enabling Mandarin-speaking learners to establish more robust representations of Korean vowels and codas. From these results, I conclude that web-based HVPT can lead to significant gains in learners’ perception of L2 sounds, with this improvement extending to a new phonetic context.
This is in line with previous research demonstrating that perceptual training is effective for establishing new segmental categories (Bradlow et al., 1997; Lively et al., 1993; many others).
Chapter 6
Conclusion

6.1 Summary of the present study

The purposes of this dissertation were to investigate how and to what extent L2 learners of Korean perceive Korean vowels and consonants with regard to their L1 transfer and L2 experience and to provide an effective learning tool for them to improve their perception abilities of Korean segments.

For these goals, this dissertation consisted of three experiments: Experiment 1 showed the effects of L2 experience on the perception of the Korean three-way laryngeal contrast in stops and affricates by Mandarin learners of Korean; Experiment 2 presented the effects of L1 inventory sizes on the perception of Korean vowels and codas by both English and Mandarin learners of Korean; and Experiment 3 investigated the effects of web-based auditory training on the perception of Korean vowels and codas by Mandarin learners of Korean and compared the effects of attention conditions through training.

Experiment 1 was designed to test whether the extent of Korean language experience affects Mandarin listeners’ ability to discriminate and identify Korean three-way contrasts in stops and affricates. In both discrimination and identification tasks, it was found that most Mandarin-speaking learners did not reach native levels of perception accuracy for the Korean contrasts. However, in terms of the effects of L2 experience, the results demonstrated that more experienced learners, who learned Korean for more than 3 years at the university level, performed better than inexperienced learners, who learned Korean for less than 2 years at the university level. The overall finding suggests that learners’ L2 perception is positively influenced by experience with L2 sounds.

The aim of Experiment 2 was to examine how and to what extent adult native English and Mandarin listeners, who differ in L1 vowel and coda inventory sizes, identify Korean vowels and codas. With respect to the influence of L1 vowel inventory, English listeners were better at identifying Korean vowels than Mandarin listeners, suggesting that individuals with a larger L1 vowel inventory have an advantage in L2 vowel perception. In addition, with respect to perception accuracy of Korean codas, English listeners had higher identification accuracy with
Korean codas than Mandarin listeners. These overall results suggest that L1 inventory size and phonotactics has a significant effect on the performance in L2 segment perception.

From the findings of Experiment 1 and 2, it was found that beginner Mandarin-speaking learners have notable difficulty perceiving certain Korean vowels and codas. Thus, in Experiment 3, a training program was developed with the aim to enhance Mandarin learners’ perception ability of Korean sounds. Furthermore, this experiment demonstrated the role of web-based auditory training, as well as the effectiveness of implicit and explicit attention on the perception of Korean vowels and codas by Mandarin learners of Korean. The results showed that online training is effective in improving learners’ L2 perception accuracy and can be an additional educational tool for L2 learners to supplement classroom instruction and interactions with native speakers of their target languages. Moreover, with regard to the effects of implicit vs explicit attention on L2 perception, outcomes showed different patterns of improvement. The vowel-trained group made strong significant gains, while the coda-trained group yielded modest but significant gains in their ability to identify Korean vowels. In contrast, the control group did not show any significant gains. These results suggest that learners are able to improve their perception of L2 vowels through explicit attention and implicit exposure to the target vowels. On the other hand, for the perception of Korean codas, only the coda-trained group, which explicitly focused on codas during training, made significant gains in coda identification accuracy. Based on these different improvement patterns, Korean coda perception learned through unconscious attention is arguably less productive than vowel perception under the same conditions. These outcomes presented in Experiment 3 demonstrated that, in learning, learners’ perceptual gains on target sounds are associated with their attention during training.

### 6.2 Theoretical implications

Over the last several decades, many researchers have systemically examined the acquisition of L2 sounds in a variety of languages. From the linguistic perspective on L2 sound acquisition, this study has focused on the effects of linguistic experience on the perceptual accuracy of L2 segments by examining how and to what extent Mandarin listeners, who have different levels of Korean-language experience, identify and discriminate Korean three-way contrasts of stops and affricates which do not exist in their L1. This study also explored the role of L1 inventory size on
the perception of Korean vowels and codas by comparing Korean identification accuracy by both native English listeners and Mandarin listeners, having different L1 phoneme inventory sizes.

The theoretical contribution of this study to the field of L2 speech perception is threefold. First, this research provides empirical evidence for the role of learners’ linguistic experience in L2 perception. Most previous research that examines the effects of experience in L2 sound learning are based on the length of residency in a target language environment. However, the length of residency is not always a good index of L2 performance, since some adult immigrant learners may not interact frequently with native speakers of their target language even though they reside in a target language country. In contrast, some motivated learners who acquire a foreign language outside of a target language community may be exposed to intensive L2 input in a classroom setting through instruction and outside of the classroom by working on their assignments and projects using the target language and through exposure to L2 media such as music, film and drama. Thus, in Experiment 1, I investigated the role of L2 experience on a basis of learners’ amount of formal instruction in a Korean classroom-based setting at the University level in Canada and found that learners’ Korean language experience in an academic setting has a strong influence on the perception of Korean vowels and consonants. In other words, the result showed that the amount of formal instruction at the university level can be a useful index of L2 input for adult learners who do not live in an L2-speaking country.

Additionally, in Experiment 3, I examined how Mandarin learners of Korean improve their perception accuracy of Korean vowels and codas through web-based high variability phonetic training which provides intensive Korean input. Results demonstrated that learners enhanced their ability to accurately perceive Korean sounds by approximately 13% as they gain experience in their Korean sounds intensively in a training-based setting. More specifically, during training, their perception accuracy was at 83.7% in the first session and improved to 91.2% in the last session for Korean vowels and 79.2% in the first session and improved to 89.1% in the last session for Korean codas. This perceptual improvement for Korean vowels and codas suggests that the perception of L2 segments can be improved through auditory training that provides intensive L2 input and immediate feedback. That is, the more they increase experience with L2 sounds and the more attention they pay to them in training, the higher their identification accuracy of L2 sounds becomes.
Taken together, results from Experiment 1 and Experiment 3 showed that learners are able to improve their perceptual accuracy of L2 segments as learners with a massive amount of L2 input and an increasing amount of L2 experience in a classroom-based setting or a training-based setting.

Second, this research provides empirical evidence for the role of learners’ L1 background in L2 perception. The effects of L1 transfer continues to be of interest in the field of L2 acquisition, because transfer, in particular with regards to the differences between L1 and L2 phoneme inventories, is useful in predicting with which sounds L2 learners generally experience difficulty. The objective of Experiment 2, therefore, was to offer a more comprehensive investigation of L2 perceptual acquisition with a focus on the role of L1 inventory size using both L2 vowels and consonants. Experiment 2 provides an interesting test case because of the difference between English and Mandarin vowel and coda inventory sizes and phonotactic restrictions compared to the target language, Korean. This study examined how English and Mandarin listeners identify both Korean vowels and codas differently due to their different L1 phoneme sizes. The results for vowel perception showed that English listeners, whose L1 has a rich vowel inventory, are better at identifying Korean vowels than Mandarin listeners, who have a small L1 vowel inventory, suggesting that large L1 vowel inventories make learning large L2 vowel inventories relatively easier than smaller L1 inventories. With respect to Korean coda perception, English listeners performed on par with native speakers of Korean. However, Mandarin learners, who lack voiceless obstruents and /m/ codas in their L1, had more difficulty accurately identifying Korean codas than the English L2 learners, whose L1 is much less restricted in the segments that can appear syllable finally. These results indicate that L1 coda inventory size has a significant effect on performance in L2 coda performance. Overall, these findings provide promising evidence on the effects of L1 inventory size on L2 segmental learning.

However, it should be mentioned that L1 inventory size cannot account for the order of L2 perceptual difficulty. The awareness of the limitations of L1 transfer on L2 perception has fostered an interest in the role of universal markedness in L2 acquisition. In Experiment 2, the results from the nasal and stop coda identification tasks showed that both English and Mandarin learners of Korean achieved higher accuracy with nasal codas than stop codas indicating that Korean nasal codas are relatively easier to identify than stop codas. This finding might be explained by the Markedness Differential Hypothesis proposed by Eckman (MDH, 1977), which
states that marked structures are more difficult to acquire than unmarked ones. Moreover, according the Sonority Hierarchy, more sonorant consonants are the least marked for codas: (more sonorant, less marked codas) glides > liquids > nasals > fricatives > stops (less sonorant, more marked codas). Considered together, we can conclude that more marked Korean stop codas are more difficult for listeners to perceive than less marked Korean nasal codas. In addition, the results from the identification tasks showed that Korean coronal codas /n, t/ are likely to be perceived more accurately than Korean dorsal codas /ŋ, k/ by both Mandarin and English learners of Korean. This finding might be also accounted for by the markedness hierarchy for place of articulation. For example, Hume (2003) looked at phonological patterning in Korean and suggested the place of markedness hierarchy: (more marked) dorsal > labial > coronal (less marked). Based on the place of markedness hierarchy, I speculate that Korean dorsal consonants /ŋ, k/ are the most difficult to identify compared to other segments in coda position. This acquisition pattern shows an effect of markedness in place of articulation on L2 coda perception, which is also consistent with previous findings that dorsal codas are more difficult to perceive accurately than other codas (Cardoso, 2011; Hume, 2003; Woods, 2005). As seen, the results about the order of perceptual difficulty for Korean codas suggest that markedness should be considered as a language internal factor which also contributes to L2 perception.

As with the perceptual difficulty of Korean codas, the perceptual difficulty from the least to most difficult Korean vowels was independent of their L1 background: ɨ > ʌ > u > o. It is conjectured that L2 listeners were most likely to accurately identify Korean /ɨ/, because /ɨ/ is a new L2 phone to both English and Mandarin listeners. It suggests that with L2 experience, listeners can establish a new L2 phonological or phonetic category and identify it correctly. This result is consistent with the SLM which claims that new L2 sounds would be easier to learn than similar L2 sounds, given sufficient exposure to and experience with the target language. Another interesting finding from Experiment 2 is that between English and Mandarin listeners, English listeners were more successful at identifying /ʌ, u, o/ than Mandarin listeners, while they were less accurate at identifying /ɨ/ than Mandarin listeners. This might be due to the fact that Mandarin listeners, with a smaller vowel inventory, have more unoccupied vowel space for the new vowel /ɨ/ and have higher perception accuracy than English listeners, with a richer vowel inventory. Overall, these results in Experiment 2 are not only in line with the SLM but also extends such findings to previously understudied L1-L2 language combinations.
Finally, this research provides empirical evidence for the role of attention in L2 perception. In Experiment 3, I compared the effectiveness of implicit and explicit attention through training. Results of L2 vowel perception demonstrated that there was no difference between the vowel-trained and coda-trained groups, suggesting that both directed attention and implicit exposure to L2 vowels play an important role in improving perception. However, for L2 codas, the two types of training were not equally effective: The coda-trained group outperformed the vowel-trained group on Korean coda perception, indicating that training with explicit attention, rather than implicit attention, leads to greater improvement in the perception of L2 codas. These asymmetrical findings have two possible explanations: positional effects and acoustic salience. With respect to positional effects, listeners may improve their perception of segments that precede the target sounds with both explicit and implicit attention. That is, when trained on /hVC/ syllables, coda-trained listeners are also attending to the vowel directly preceding the coda. However, if listeners tend not to focus on segments that follow their target sounds, this can account for why significant perception improvement of Korean codas did not occur in the vowel-trained group. Another possible explanation is that acoustic salience affects L2 perception. Flege and Wang (1989) showed that English final stops prove difficult for Mandarin listeners due to the unreleased burst of stops. Similarly, Mandarin learners of Korean might experience more perceptual difficulty with Korean codas than vowels, because Korean stops are not released in coda position, causing Mandarin listeners to face difficulty in accurately identifying Korean codas, especially when their attention is drawn to the target sounds. It would be interesting to investigate whether other L2 learners of Korean show similar perceptual patterns for Korean sounds.

Altogether, the theoretical contribution of this dissertation to the field of L2 speech perception is threefold. This research contributes (1) to the role of linguistic experience in the perception L2 contrasts, which do not exist in L1 (2) to the impact of L1 inventory sizes in the perception of L2 vowels and codas, and (3) to the effects of phonetic attention type on the perceptual acquisition of L2 vowels and codas through web-based auditory training.

6.3 Pedagogical implications

The main findings reported in this study have several implications for L2 pedagogy. Previous perceptual training studies have provided evidence of the effectiveness of perceptual training;
however, these studies have been mostly laboratory-based, where learners participate in training in a controlled condition while being monitored by the experimenter. Due to this limitation, although perceptual training is effective for learners to increase their L2 perception abilities, language instructors have not actively used this training program in and outside of the classroom. In contrast, the web-based auditory training program I developed shows positive learning effects while providing a more learner-friendly learning environment.

In other words, adopting web-based auditory programs is useful for L2 learners to facilitate the language learning process because it provides intensive L2 input during training and is student-centered, which enables them to learn at their own pace and at any location. Implementing such a program would allow language instructors to provide students with useful tools to get sufficient practice for their L2 perception and pronunciation under conditions of autonomy. Moreover, L2 learners who do not reside in their target language community would be able to reap the benefits of being exposed to multiple talkers in diverse phonetic contexts by accessing this perceptual training online when and where they can provide their full attention.

In addition to the exposure to target language input, learners’ attention to input is necessary to facilitate their language development. This study supported the view that L2 learners are able to improve their perception abilities when they are intentionally attending to their target sounds during training. That is, use of explicit instruction contributes to L2 learning and teaching; thus, language instructors can benefit from explicit instruction on L2 speech perception in class. When instructors explicitly explain their target goals with L2 segment perception and/or pronunciation to their students, students are able to be consciously aware of their target sounds which can lead to perception improvement. Interestingly, this study also found that when learners are passively exposed to L2 vowels, they can acquire some knowledge unconsciously. Language instructors therefore need to provide enough L2 sound input for learners in class, and also encourage students to expose themselves to their target language outside of the classroom.

6.4 Contribution of this study

This study is intended to constitute an empirical and pedagogical contribution to the field of second language phonology. Most training studies have analyzed English as a first or target language (Lee et al., 2014), so it is important to increase the range of languages to determine how training works amongst the diversities of language-learning conditions. To my knowledge,
this is the first study to examine the effects of web-based perceptual training on the perception of Korean vowels and codas by L2 learners. By investigating a language other than English, we can provide a more accurate picture of how beneficial training is for the acquisition of non-native sounds.

With respect to the empirical contribution, this study showed the role of L1 transfer and L2 experience as well as effectiveness of web-based auditory training in the perceptual acquisition of L2 sounds. In particular, a strong benefit of this training is its ability to provide learners with more practice and promote greater learner autonomy so that training can help learners achieve more native-like perceptual accuracy in their L2. In addition, L2 researchers or instructors, who want to use a web-based training program for their research or teaching purposes, are able to create their own program using tutorial materials entitled “online linguistics experiments using jsPsych” (Ryu, 2018b). In this way, online training programs will help L2 learners enhance their abilities, paving the way for more effective methods of teaching L2 learners, as well as L2 researchers for their studies on L2 speech acquisition.

This study provides evidence suggesting an advantage for explicit attention in perceptual training. Specifically, it determines what training condition leads to the most improvement in L2 perceptual acquisition by looking at whether L2 learning can be attained through conscious attention to target sounds. This study can contribute to teaching in L2 speech as well as open new avenues to the development of more training research studies in the field of second language phonological acquisition.

6.5 Limitations and future directions

The aims of this dissertation have largely been attained, but there are some limitations. First, I only included inexperienced Mandarin-speaking learners in Experiment 3 because training would be more successful than including experienced L2 learners who were already exposed to intensive natural variability from multiple speakers in various phonetic contexts in a wide range of situations. However, Iverson, Pinet and Evans (2012) found that both inexperienced and experienced French listeners who received eight sessions of HVPT on English vowels showed similar degrees of improvement in perception and production. In future studies, I will examine whether the training method is effective for experienced L2 learners such as intermediate and advanced learners who already have ample exposure to the target language. Furthermore, this
study did not determine whether there is long-term retention in L2 perception through web-based training. Thus, it would be meaningful to investigate whether and to what extent there exists a long-term positive effect of training on L2 perception by adding a delayed post-test in three months.

Finally, as far as perceptual training and L2 production are concerned, results on whether perceptual training effects can transfer to production have been varied. Several studies have reported that gains in perception from perceptual training can be transferred to gains in production (Bradlow et al. 1997, 1999; Rochet, 1995; Lambacher et al., 2005; Lopez-Soto & Keywley-Port 2009, Wang, Jongman, & Sereno, 2003). For example, Bradlow et al. (1997) investigated the effects of perceptual training on the perception and production of English /r/-/l/ by adult Japanese speakers. The results showed that Japanese speakers who received intensive perceptual training improved in both perception and production of the English contrast. In particular, the Japanese speakers’ post-test productions were more accurately identified by native English listeners compared to pre-test productions. Yet, other studies have found that perceptual training does not lead to improvement of the production of L2 sounds. Hwang and Lee (2015) examined whether the learning effects gained from perceptual training can be successfully transferred to speech production by looking at Korean early learners’ perception and production of English vowels and consonants. They observed that learners showed greater improvement in perceiving consonants than vowels. However, native English listener judgements revealed no significant improvement of the Korean learners’ production of English vowels and consonants. This result suggests that perceptual training does not have a positive influence on the production of L2 segments. In future work, therefore, it will be imperative to determine whether the knowledge learned through perceptual training transfers successfully to production tasks, which may in turn show how speech perception and production modalities are connected in L2 learning. The effectiveness of the training will be assessed by both objective acoustic analysis and subjective ratings by native listeners. By measuring the effects of training on the perception and production of the L2, it will shed further light on the link between perception and production in L2 acquisition. In addition, the focus on pedagogically feasible perceptual training will contribute to identifying better practice in and outside of the language classroom.
References


Kuznetsova, A., Christensen, R. H. B., & Brockhoff, P. B. (2013). Different tests on lmer objects (of the lme4 package): introducing the lmerTest package.


## Appendices

### Appendix A. Korean stops and affricates in word-initial position (Real words)

<table>
<thead>
<tr>
<th></th>
<th>Aspirated</th>
<th>Lenis</th>
<th>Fortis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bilabial</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pul</td>
<td>pʰul ‘glass’</td>
<td>pul ‘fire’</td>
<td>p’ul ‘horn’</td>
</tr>
<tr>
<td>Pang</td>
<td>pʰang ‘to pop’</td>
<td>pang ‘room’</td>
<td>p’ang ‘bread’</td>
</tr>
<tr>
<td>Pe.ta</td>
<td>pʰe.ta ‘to beat’</td>
<td>pe.ta ‘to penetrate’</td>
<td>p’e.ta ‘to bleach’</td>
</tr>
<tr>
<td>Pi.ta</td>
<td>pʰi.ta ‘to bloom’</td>
<td>pi.ta ‘to be empty’</td>
<td>p’i.ta ‘to sprin’</td>
</tr>
<tr>
<td><strong>Alveolar</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thal</td>
<td>tʰal ‘mask’</td>
<td>tal ‘moon’</td>
<td>t’al ‘daught’</td>
</tr>
<tr>
<td>Tam</td>
<td>tʰam ‘greed’</td>
<td>tam ‘wall’</td>
<td>t’am ‘sweat’</td>
</tr>
<tr>
<td>Thang</td>
<td>tʰang ‘soup’</td>
<td>tang ‘sugar’</td>
<td>t’ng ‘land’</td>
</tr>
<tr>
<td>Te.la</td>
<td>tʰɛ.l.ta ‘to brush’</td>
<td>tɛ.l.ta ‘to lighten’</td>
<td>t’ɛ.l.ta ‘to shake’</td>
</tr>
<tr>
<td><strong>Velar</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Khi</td>
<td>kʰi ‘heigh’</td>
<td>ki ‘energy’</td>
<td>k’i ‘meal’</td>
</tr>
<tr>
<td>Khe</td>
<td>kʰe ‘to dig’</td>
<td>ke ‘dog’</td>
<td>k’e ‘sesame’</td>
</tr>
<tr>
<td>Kha.t</td>
<td>kʰa.ta ‘to be large’</td>
<td>ki.ta ‘to draw’</td>
<td>k’a.ta ‘to turn off’</td>
</tr>
<tr>
<td>Khi.t</td>
<td>kʰi.t.a ‘to turn on’</td>
<td>ki.ta ‘to crawl’</td>
<td>k’i.t.a ‘to be caught’</td>
</tr>
<tr>
<td><strong>Post-alveolar</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cha</td>
<td>cʰa ‘tea/car’</td>
<td>ca ‘ruler’</td>
<td>c’a ‘to be salty’</td>
</tr>
<tr>
<td>Cok</td>
<td>cʰok ‘point’</td>
<td>cok ‘a foot/ a tribe’</td>
<td>c’ok ‘page’</td>
</tr>
<tr>
<td>Cha.t</td>
<td>cʰa.ta ‘to kick’</td>
<td>ca.ta ‘to sleep’</td>
<td>c’a.ta ‘to be squeeze’</td>
</tr>
<tr>
<td>Chi.t</td>
<td>cʰi.ta ‘to hit’</td>
<td>ci.ta ‘to lose’</td>
<td>c’i.ta ‘to gain weight’</td>
</tr>
</tbody>
</table>
Appendix B. Korean stops and affricates in word-initial position (Nonce words)

<table>
<thead>
<tr>
<th></th>
<th>Aspirated</th>
<th>Lenis</th>
<th>Fortis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilabial</td>
<td>파아 pʰa.a</td>
<td>바아 pa.a</td>
<td>빼아 p’a.a</td>
</tr>
<tr>
<td></td>
<td>파미 pʰa.mi</td>
<td>바미 pa.mi</td>
<td>빼미 p’a.mi</td>
</tr>
<tr>
<td></td>
<td>파나 pʰa.na</td>
<td>바나 pa.na</td>
<td>빼나 p’a.na</td>
</tr>
<tr>
<td></td>
<td>파사 pʰa.sa</td>
<td>바사 pa.sa</td>
<td>빼사 p’a.sa</td>
</tr>
<tr>
<td>Alveolar</td>
<td>타아 tʰa.a</td>
<td>다아 ta.a</td>
<td>따아 t’a.a</td>
</tr>
<tr>
<td></td>
<td>타미 tʰami</td>
<td>다미 ta.mi</td>
<td>따미 t’a.mi</td>
</tr>
<tr>
<td></td>
<td>타나 tʰa.na</td>
<td>다나 ta.na</td>
<td>따나 t’a.na</td>
</tr>
<tr>
<td></td>
<td>타사 tʰa.sa</td>
<td>다사 ta.sa</td>
<td>따사 t’a.sa</td>
</tr>
<tr>
<td>Velar</td>
<td>카아 kʰa.a</td>
<td>가아 ka.a</td>
<td>까아 k’a.a</td>
</tr>
<tr>
<td></td>
<td>카미 kʰa.mi</td>
<td>가미 ka.mi</td>
<td>까미 c’a.mi</td>
</tr>
<tr>
<td></td>
<td>카하 kʰa.ha</td>
<td>가하 ka.ha</td>
<td>까하 c’a.ha</td>
</tr>
<tr>
<td>Post-alveolar</td>
<td>차나 cʰa.na</td>
<td>자나 ca.na</td>
<td>짜나 c’a.na</td>
</tr>
<tr>
<td></td>
<td>차미 cʰa.mi</td>
<td>자미 ca.mi</td>
<td>짜미 c’a.mi</td>
</tr>
<tr>
<td></td>
<td>차하 cʰa.hi</td>
<td>자하 ca.hi</td>
<td>짜하 c’a.hi</td>
</tr>
<tr>
<td></td>
<td>차하 cʰa.ha</td>
<td>자하 ca.ha</td>
<td>짜하 c’a.ha</td>
</tr>
</tbody>
</table>

Appendix C. Korean monosyllabic words for the vowel perception task

<table>
<thead>
<tr>
<th>Onset</th>
<th>Coda Vowel</th>
<th>No coda</th>
<th>ㄱ</th>
<th>ㅅ</th>
<th>No coda</th>
<th>ㄱ</th>
<th>ㅅ</th>
<th>No coda</th>
<th>ㄱ</th>
<th>ㅅ</th>
<th>No coda</th>
</tr>
</thead>
<tbody>
<tr>
<td>ㄱ</td>
<td>/ki/</td>
<td>/kik/</td>
<td>/kit/</td>
<td>/kip/</td>
<td>/kin/</td>
<td>/kom/</td>
<td>/kii/</td>
<td>/s/</td>
<td>/s/</td>
<td>/sip/</td>
<td>/sin/</td>
</tr>
<tr>
<td>끼</td>
<td>/ko/</td>
<td>/kok/</td>
<td>/kont/</td>
<td>/kop/</td>
<td>/kon/</td>
<td>/kom/</td>
<td>/ko/</td>
<td>/so/</td>
<td>/so/</td>
<td>/sop/</td>
<td>/son/</td>
</tr>
<tr>
<td>ㅏ</td>
<td>/kuk/</td>
<td>/kuk/</td>
<td>/kup/</td>
<td>/kuy/</td>
<td>/kun/</td>
<td>/kum/</td>
<td>/kul/</td>
<td>/su/</td>
<td>/su/</td>
<td>/sup/</td>
<td>/sun/</td>
</tr>
</tbody>
</table>
Appendix D. Korean monosyllabic words including stop codas for the coda perception tasks

<table>
<thead>
<tr>
<th>V1</th>
<th>ㅏ</th>
<th>ㅐ</th>
<th>ㅣ</th>
<th>ㅗ</th>
<th>ㅜ</th>
<th>ㅓ</th>
<th>ㅡ</th>
<th>ㅏ</th>
<th>ㅣ</th>
<th>ㅗ</th>
<th>ㅜ</th>
<th>ㅓ</th>
<th>ㅡ</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2</td>
<td>Reply</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
</tr>
<tr>
<td>C1</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
</tr>
</tbody>
</table>

Appendix E. Korean monosyllabic words including nasal codas for the coda perception tasks

<table>
<thead>
<tr>
<th>V1</th>
<th>ㅏ</th>
<th>ㅣ</th>
<th>ㅗ</th>
<th>ㅜ</th>
<th>ㅓ</th>
<th>ㅡ</th>
<th>ㅏ</th>
<th>ㅣ</th>
<th>ㅗ</th>
<th>ㅜ</th>
<th>ㅓ</th>
<th>ㅡ</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2</td>
<td>Reply</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
</tr>
<tr>
<td>C1</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>V1</th>
<th>ㅏ</th>
<th>ㅣ</th>
<th>ㅗ</th>
<th>ㅜ</th>
<th>ㅓ</th>
<th>ㅡ</th>
<th>ㅏ</th>
<th>ㅣ</th>
<th>ㅗ</th>
<th>ㅜ</th>
<th>ㅓ</th>
<th>ㅡ</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2</td>
<td>Reply</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
</tr>
<tr>
<td>C1</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>V1</th>
<th>ㅏ</th>
<th>ㅣ</th>
<th>ㅗ</th>
<th>ㅜ</th>
<th>ㅓ</th>
<th>ㅡ</th>
<th>ㅏ</th>
<th>ㅣ</th>
<th>ㅗ</th>
<th>ㅜ</th>
<th>ㅓ</th>
<th>ㅡ</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2</td>
<td>Reply</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
</tr>
<tr>
<td>C1</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>V1</th>
<th>ㅏ</th>
<th>ㅣ</th>
<th>ㅗ</th>
<th>ㅜ</th>
<th>ㅓ</th>
<th>ㅡ</th>
<th>ㅏ</th>
<th>ㅣ</th>
<th>ㅗ</th>
<th>ㅜ</th>
<th>ㅓ</th>
<th>ㅡ</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2</td>
<td>Reply</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
</tr>
<tr>
<td>C1</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>V1</th>
<th>ㅏ</th>
<th>ㅣ</th>
<th>ㅗ</th>
<th>ㅜ</th>
<th>ㅓ</th>
<th>ㅡ</th>
<th>ㅏ</th>
<th>ㅣ</th>
<th>ㅗ</th>
<th>ㅜ</th>
<th>ㅓ</th>
<th>ㅡ</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2</td>
<td>Reply</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
</tr>
<tr>
<td>C1</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
<td>Click</td>
</tr>
</tbody>
</table>
Appendix F. Korean monosyllabic words including all vowels and codas used at pre-test, training and post-test (Onset: /h/)

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Coda</th>
<th>k</th>
<th>p</th>
<th>t</th>
<th>n</th>
<th>m</th>
<th>n̄</th>
<th>l</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>학</td>
<td>/hak/</td>
<td>합</td>
<td>/hap/</td>
<td>한</td>
<td>/ha/</td>
<td>함</td>
<td>/ham/</td>
</tr>
<tr>
<td>e</td>
<td>핵</td>
<td>/hek/</td>
<td>합</td>
<td>/hep/</td>
<td>한</td>
<td>/hen/</td>
<td>함</td>
<td>/hem/</td>
</tr>
<tr>
<td>i</td>
<td>핍</td>
<td>/hik/</td>
<td>핍</td>
<td>/hip/</td>
<td>핍</td>
<td>/hit/</td>
<td>핍</td>
<td>/him/</td>
</tr>
<tr>
<td>o</td>
<td>혹</td>
<td>/hok/</td>
<td>홉</td>
<td>/hop/</td>
<td>홉</td>
<td>/hot/</td>
<td>홉</td>
<td>/hom/</td>
</tr>
<tr>
<td>u</td>
<td>혹</td>
<td>/huk/</td>
<td>홉</td>
<td>/hup/</td>
<td>홉</td>
<td>/hut/</td>
<td>홉</td>
<td>/hum/</td>
</tr>
<tr>
<td>A</td>
<td>핍</td>
<td>/hāk/</td>
<td>핍</td>
<td>/hap/</td>
<td>핍</td>
<td>/hat/</td>
<td>핍</td>
<td>/ham/</td>
</tr>
<tr>
<td>i</td>
<td>혹</td>
<td>/hik/</td>
<td>홉</td>
<td>/hip/</td>
<td>홉</td>
<td>/hit/</td>
<td>홉</td>
<td>/him/</td>
</tr>
</tbody>
</table>

Appendix G. Korean monosyllabic words including all vowels and codas used at generalization test (Onset: /k/)

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Coda</th>
<th>k</th>
<th>p</th>
<th>t</th>
<th>n</th>
<th>m</th>
<th>n̄</th>
<th>l</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>각</td>
<td>/kak/</td>
<td>갑</td>
<td>/kap/</td>
<td>갑</td>
<td>/kat/</td>
<td>갑</td>
<td>/kan/</td>
</tr>
<tr>
<td>e</td>
<td>객</td>
<td>/kek/</td>
<td>객</td>
<td>/kep/</td>
<td>객</td>
<td>/ket/</td>
<td>객</td>
<td>/ken/</td>
</tr>
<tr>
<td>i</td>
<td>급</td>
<td>/kik/</td>
<td>급</td>
<td>/kip/</td>
<td>급</td>
<td>/kit/</td>
<td>급</td>
<td>/kin/</td>
</tr>
<tr>
<td>o</td>
<td>곡</td>
<td>/kok/</td>
<td>곡</td>
<td>/kop/</td>
<td>곡</td>
<td>/kot/</td>
<td>곡</td>
<td>/kom/</td>
</tr>
<tr>
<td>u</td>
<td>국</td>
<td>/kul/</td>
<td>국</td>
<td>/kup/</td>
<td>국</td>
<td>/kut/</td>
<td>국</td>
<td>/kum/</td>
</tr>
<tr>
<td>A</td>
<td>급</td>
<td>/kāk/</td>
<td>급</td>
<td>/kap/</td>
<td>급</td>
<td>/kat/</td>
<td>급</td>
<td>/kan/</td>
</tr>
<tr>
<td>i</td>
<td>급</td>
<td>/kik/</td>
<td>급</td>
<td>/kip/</td>
<td>급</td>
<td>/kit/</td>
<td>급</td>
<td>/kin/</td>
</tr>
</tbody>
</table>
Appendix H. Instruction form for online perceptual training

Training Instructions (Group 1)

1. This training consists of 8 sessions and each session takes about 20 minutes.
2. Listen to the recordings of Korean sounds with either your headphones or earphones. Identify what you heard by pressing the corresponding numbers on the keyboard.
3. You will receive immediate feedback on your choice of answer. If you give a wrong response, you will see “incorrect” on the screen and you will need to try again until you respond correctly.
4. You should complete all sessions in a quiet room, preferably at home, within 16 days (8 days-16 days recommended).
5. You should take no more than one session per day.
6. After finishing the last session, you will need to decide on a day to return to the phonetics lab and complete the last experiment.
   Book the last experiment: http://kor1.youcanbook.me (March 21 –April 7)
7. You should submit this completed form to the research team on the day you complete the last experiment.
8. If you have any questions, please email the research team at michellemj.heo@mail.utoronto.ca

Online training
http://142.1.98.7/~phonlab/Exp/G1/
Recommended browsers: Chrome, Safari, Firefox

Participant Information

ID: 
Password: 

Session 1
Date: ________________ Time: __________~ __________
Where (Please circle one): Home Library Other______________
Hearing device used for training (Check one): Earphones ( ) Headphones ( )

Please write down the numbers of words you answered incorrectly per session.
Appendix I. [Experiment 2] A post-hoc analysis of interaction between Korean individual vowels and L1 for the perception accuracy of Korean vowels

<table>
<thead>
<tr>
<th>Vowels</th>
<th>L1</th>
<th>Value</th>
<th>Df</th>
<th>Chisq</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Korean vs. Mandarin</td>
<td>0.90982</td>
<td>1</td>
<td>8.5562</td>
<td>0.003444 **</td>
</tr>
<tr>
<td>i</td>
<td>Korean vs. English</td>
<td>0.92444</td>
<td>1</td>
<td>10.1440</td>
<td>0.001448 **</td>
</tr>
<tr>
<td></td>
<td>Mandarin vs. English</td>
<td>0.54805</td>
<td>1</td>
<td>0.3396</td>
<td>0.560066</td>
</tr>
<tr>
<td>o</td>
<td>Korean vs. Mandarin</td>
<td>0.98648</td>
<td>1</td>
<td>29.7115</td>
<td>5.014e-08 ***</td>
</tr>
<tr>
<td></td>
<td>Korean vs. English</td>
<td>0.96959</td>
<td>1</td>
<td>19.2497</td>
<td>1.147e-05 ***</td>
</tr>
<tr>
<td></td>
<td>Mandarin vs. English</td>
<td>0.30406</td>
<td>1</td>
<td>8.2984</td>
<td>0.003968 **</td>
</tr>
<tr>
<td>u</td>
<td>Korean vs. Mandarin</td>
<td>0.77118</td>
<td>1</td>
<td>8.0942</td>
<td>0.004441 **</td>
</tr>
<tr>
<td></td>
<td>Korean vs. English</td>
<td>0.64460</td>
<td>1</td>
<td>1.9112</td>
<td>0.166829</td>
</tr>
<tr>
<td></td>
<td>Mandarin vs. English</td>
<td>0.34988</td>
<td>1</td>
<td>4.5146</td>
<td>0.033606 *</td>
</tr>
<tr>
<td>Λ</td>
<td>Korean vs. Mandarin</td>
<td>0.96862</td>
<td>1</td>
<td>25.3582</td>
<td>4.761e-07 ***</td>
</tr>
<tr>
<td></td>
<td>Korean vs. English</td>
<td>0.94057</td>
<td>1</td>
<td>16.3202</td>
<td>5.349e-05 ***</td>
</tr>
<tr>
<td></td>
<td>Mandarin vs. English</td>
<td>0.33896</td>
<td>1</td>
<td>5.1636</td>
<td>0.023065 *</td>
</tr>
</tbody>
</table>

Note. Significance codes: <0.001 ‘***’, <0.01 ‘**’, <0.05 ‘*’. Reference levels for categorical variable are underlined. Significant results are in bold.

Appendix J. Online language background questionnaire for Experiment 3

1. Participant ID (Ex: P001):
2. Name (Firstname Lastname) (Ex: Terry Lee):
3. E-mail address:
4. Age:
5. Are you a standard Mandarin speaker? (Yes or No)
6. Do you speak any other Chinese dialects? (Yes or No)
7. Where were you born in China? (Ex: Hangzhou Zhejiang)
8. How many years have you lived in Canada? (Ex: 3 years)
9. Have you ever lived/visited in South Korea? (Yes or No)
9-1. (If yes) How long and where? (Ex: 2 weeks Seoul)
10. Have you ever studied in South Korea? (Yes or No)
10-1. (If yes) How long and where? (Ex: 6 months Seoul)
11. Thinking about your day-to-day life.
Which language do you use the most on a daily basis and what percent of the time you use it? (Ex: Mandarin 70%)
12. Which is your second and third most used language and what percent of the time you use it? (Ex: English 20% Korean 10%)
13. What is your gender?
   ○ Female        ○ Male

14. Which Korean course are you currently taking at Universities?
   ○ Modern Standard Korean I (EAS 110)
   ○ Modern Standard Korean II (EAS 210)
   ○ Accelerated Modern Standard Korean I & II (EAS 211)
   ○ other courses

15. What do you think about your Korean language ability?
   ○ Very poor   ○ Poor   ○ Fair   ○ Good   ○ Native-like
15-1. Korean speaking fluency?
   ○ Very poor   ○ Poor   ○ Fair   ○ Good   ○ Native-like
15-2. Korean listening comprehension?
   ○ Very poor   ○ Poor   ○ Fair   ○ Good   ○ Native-like
15-3. Korean reading proficiency?
   ○ Very poor   ○ Poor   ○ Fair   ○ Good   ○ Native-like
15-4. Korean writing proficiency?
   ○ Very poor   ○ Poor   ○ Fair   ○ Good   ○ Native-like

16. How many hours are you exposed to Korean per week (including lectures/tutorials/music/movie/drama/book)?
   ○ Less than 5 hours
   ○ 6 hours-10 hours
   ○ 11 hours-15 hours
   ○ 16 hours-20 hours
   ○ more than 20 hours

17. What do you think about your English language ability?
   ○ Very poor   ○ Poor   ○ Fair   ○ Good   ○ Native-like
17-1. English speaking fluency?
   ○ Very poor   ○ Poor   ○ Fair   ○ Good   ○ Native-like
17-2. English listening comprehension?
   ○ Very poor   ○ Poor   ○ Fair   ○ Good   ○ Native-like
17-3. English reading proficiency?
   ○ Very poor   ○ Poor   ○ Fair   ○ Good   ○ Native-like
17-4. English writing proficiency?
   ○ Very poor   ○ Poor   ○ Fair   ○ Good   ○ Native-like

18. Do you have a hearing impairment or disability that may affect your participation in this experiment?
   ○ Yes        ○ No