THE PERCEPTUAL PROCESSING
OF
SECOND LANGUAGE CONSONANT CLUSTERS

by

Baris Kabak

A dissertation submitted to the Faculty of the University of Delaware in partial requirements for the degree of Doctor of Philosophy with a major in Linguistics.

Committee Members:

William J. Idsardi, Ph.D. (Dissertation Supervisor)

Irene Vogel, Ph.D.

Benjamin Bruening, Ph.D.

John Kingston, Ph.D.

Spring 2003

ACKNOWLEDGMENTS

I would like to take this opportunity to express my gratitude to several people who made this dissertation possible and my graduate studies a meaningful experience. First, I am indebted to my advisor Bill Idsardi who has guided me with great enthusiasm and support in the making of this dissertation and throughout my graduate career. I have learnt and benefited tremendously from Bill's vision, expertise and knowledge, without which this dissertation would not have been possible. I would also like to thank my committee members, Irene Vogel, Benjamin Bruening, and John Kingston, for the time and effort they have put in reading this dissertation and the thorough comments and suggestions they have provided. I am especially thankful to Benjamin Bruening for his keen eye and challenging my thinking, which has greatly improved this dissertation.

Special thanks are in order for several people who have contributed to my education at Delaware. I have had the privilege to be a student of Irene Vogel and to collaborate with her. I am thankful to Irene for giving me the opportunity to work with her and sharing her expertise with me, which has given me the tools to grow as a phonologist. It has been a great pleasure to be around and a student of Colin Phillips, Satoshi Tomioka, Irene Kraemer, Roberta Golinkoff, Gabriella Hermon, Peter Cole, and Cindy Brown. My contemporaries at the Linguistics Department, Chang-Yong Sim, Woohyeok Chang, and Chonghyuck Kim, have taught me a lot about Korean. Special thanks to Chang-Yong Sim, who has spent many enjoyable hours with me patiently answering my questions about Korean. Many thanks to Chun-Chieh Hsu, Min-Jeong Son, Sun-Ah Son, Eun-kyung Sung, Perng Wang, Ngee Thai Yap, Heejong Yi, the English Language Institute, and the Filiz Family and Evi Pamer in Germany, who have generously helped me find participants for this and various pilot studies. I am indebted to the Korean community in Delaware for so kindly participating in my study. Many thanks to Stephanie Baker, who has been a wonderful friend and an important part of my life in America, for her support and help in formatting this dissertation.

I would like to thank my dear parents, Mustafa and Binnaz Kabak, who have sacrificed greatly for my education. Their unconditional love, support and unending faith in me have always given me the strength to move on and pursue my dreams despite being continents away from them. I am also grateful to Fritz and Frieda Altmann, who have made my revisions in peaceful Groβschwarzenlohe a pleasant one with their hospitality and support.

Finally, I give my warmest thanks to my wonderful wife, Heidi Altmann, for her endless love, support, patience and knowledge in helping me accomplish this sometimes seemingly unreachable goal. Without Heidi, I would have been lost in this journey.

TABLE OF CONTENTS

LIST OF TABLES ix
LIST OF FIGURES xii
ABSTRACT xiv

Chapter

1. INTRODUCTION 1

    1.1. Introduction 1
    1.2. Organization of the Dissertation 4

2. NATIVE LANGUAGE PHONOTACTIC KNOWLEDGE AND THE PERCEPTION OF CONSONANT CLUSTERS 6

    2.1. Introduction 6
    2.2. The Role of Language Specific Phonotactic Restrictions in Speech Perception 6
    2.3. Studies on the Perception of Consonant Clusters by Japanese listeners 12
    2.4. Shortcomings of the Previous Studies on Japanese 22
    2.5. Summary 28

3. MOTIVATING THE PRESENT STUDY: KOREAN PHONOTACTICS 30

    3.1. Introduction 30
    3.2. Phonotactics of Korean and English 32
    3.3. Restrictions on Consonantal Contact in Korean 36
    3.4. Loan Word Adaptations in Korean 43

4. RESEARCH QUESTIONS AND HYPOTHESES 53

4.1. Research Objectives and Questions 53
4.2. Hypotheses 56
4.2.1. The Consonantal Contact Hypothesis 56
4.2.2. The Coda/Onset Identity Hypothesis 58
4.2.3. Perceptual Level Representations and Sub-hypotheses 59

4.2.3.1. Phonetic Processing Hypothesis 60
4.2.3.2. Phonological Processing Hypothesis 62

5. EMPIRICAL TEST: METHOD AND PREDICTIONS 68

5.1. Method 68

5.1.1. Choice of Consonantal Contact Contexts 68
5.1.2. Choice of the Epenthetic Vowels 72
5.1.3. Materials 73
5.1.4. Design 79

5.2. Experimental Predictions 81
5.3. Subject Population 85
5.4. Procedure 86
5.5. Analysis Methods 87

5.5.1. Signal Detection Theory 87

6. RESULTS 97

6.1. Introduction 97
6.2. Descriptive Statistics 99
6.3. Analysis of Variance 103
6.4. Post Hoc Analyses on the Korean Data 109
6.5. Cluster Analysis 112
6.6. The Analysis of the Clusters [ln] and [lt^h] 117
6.7. Effect Size: Statistical Tests on Obstruent Clusters 120
6.8. Effect Size: Cost Analysis 123
6.9. The Perception of the Likely Output Forms of Illicit Clusters 130
6.10. Summary and Conclusions 132

7. CONSEQUENCES AND IMPLICATIONS ... 135

7.1. Introduction ... 135
7.2. Overview ... 136
7.3. Speech Perception is not Isomorphic to Speech Production ... 141
7.4. Frequency Information and Perceptual Illusion ... 150
7.5. Onset/Coda Detection Theory ... 156

8. CONCLUSIONS AND REMAINING QUESTIONS ... 168

APPENDIX A: CERTIFICATE OF INFORMED CONSENT ... 172

APPENDIX B: QUESTIONNAIRE ... 175

APPENDIX C: ENGLISH AND KOREAN MEDIAN AND MEAN D-PRIME SCORES AND STANDARD DEVIATIONS FOR EXPERIMENTAL DOUBLETS ... 176

APPENDIX D: INDIVIDUAL D-PRIME VALUES AND COST OF FEATURES ... 177

APPENDIX E: POST-HOC TEST TABLES FOR THE ENGLISH GROUP ... 179

APPENDIX F: POST-HOC TEST TABLES FOR THE KOREAN GROUP ... 182

REFERENCES ... 185

LIST OF TABLES

3.1. Heterosyllabic $C_1.C_2$ sequences necessary to separate contact violations from coda violations 31

3.2 Consonantal inventory of Korean 36

3.3 Vowel inventory of Korean 37

3.4 The layout of the subsequent tables of the surface forms of heterosyllabic consonantal contacts in Korean 48

3.5 Surface forms of [Stop]+[Stop] clusters in Korean 49

3.6 Surface forms of [Stop] + ([Sibilant] or [Sonorant] or /h/) in Korean 50

3.7 Surface forms of ([Sibilant] or [Sonorant] or /h/) + [Stop] in Korean 51

3.8 Surface forms of ([Sibilant] or [Sonorant] or /h/) + ([Sibilant] or [Sonorant] or /h/) in Korean 52

5.1 Consonantal contexts used in the experiment and their legitimacy in Korean 71

5.2. Test words used in the experiment 74

5.3 Evaluation table for the hypotheses 82

5.4 Possible response outcomes on a trial of a yes-no experiment 89

5.5 An SDT analysis scheme of an example doublet 90

5.6 Frequency matrix for participant K17's d' score for [p^hákt^{h}a] vs. [p^hákut^{h}a] 93

5.7 An example corrected d' calculation matrix 94

5.8 An example corrected d' calculation matrix 95

5.9 An example corrected d' calculation matrix 95

5.10 An example frequency matrix where p(F)=p(H)... 96

6.1 Consonantal contexts used in the experiment. 97

6.2. ANOVA Multivariate Tests 104

6.3 Mean d' scores, standard errors, lower and upper bounds within 95% confidence rate for CVCCV vs. CVCVCV discrimination............................ 106

6.4 Evaluation table for the hypotheses... 116

6.5 Mean d' scores, standard errors, lower and upper bounds within 95% confidence rate for [ln] and [lt^h]. 117

6.6 Sphericity Test... 121

6.7 Sequence of hierarchically included models from the General Linear Model statistical analysis.. 121

6.8 The effect size (R^2) for each model 122

6.9 The mean cost of features and one-sample t-test statistics............................. 125

6.10 Mean cost of features in units of d', standard errors, lower and upper bounds within 95% confidence rate. 126

6.11 Korean mean d' scores for the discrimination of test clusters with their likely output forms. 131

7.1 Summary of Korean Results in the CC vs. CVC condition 138

7.2 An example SDT analysis scheme ([p^hacma] vs. [p^hac^hima])........................ 146

7.3 Frequency of occurrence and structural well-formedness of stop-consonant clusters (adapted from Moreton (2002)) 153

7.4 The Onset/Coda specification of Korean consonants..................................... 161

English and Korean median and mean d' scores 176

D.1 Individual English d' scores and cost of features 177

D.2 Individual Korean d' scores and cost of features 178

E.1 The Scheffe Test for English d' scores 179

E.2 The Tukey/Kramer test for English d' scores 180

F.1 The Scheffe Test for Korean d' scores 182

F.2 The Tukey/Kramer test for Korean d' scores 183

LIST OF FIGURES

2.1 Percept [u] vowel judgments as a function of vowel duration (after Dupoux et al., 1999: Experiment 1) 15

2.2 Japanese coda filter........ 24

2.3 Syllabification of [tombo] in Japanese (adapted from Itô(1989))........ 24

5.1 AX discrimination protocol per trial 81

6.1 Korean and English *mean* d' scores for the discrimination of [CVCCV] vs. [CVCVCV] words........ 99

6.2 Korean and English *median* d' scores for the discrimination of [CVCCV] vs. [CVCVCV] words 102

6.3 The upper and lower bounds for d' scores for CVCCV vs. CVCVCV discrimination with 95% confidence interval........ 107

6.4 Consonant cluster groupings in the English group based on a Scheffe post-hoc test........ 108

6.5 Consonant cluster groupings in the Korean group based on post-hoc tests 110

6.6 Cluster analysis dendogram........ 113

6.7 Three-way clustering of the consonant clusters using Cluster Analysis........ 114

6.8 Two-way clustering of the consonant clusters using Cluster Analysis........ 115

6.9 Three-way clusterings using Cluster Analysis and ANOVA........ 118

6.10 Mean cost of features and error bars with the 95% Confidence Interval........ 127

7.1 Korean perception vs. production grammar 142

7.2 Korean coda constraint (perceptual)... 158

ABSTRACT

Listeners perceive epenthetic vowels within consonant clusters that violate the phonotactics of their first language (L1) (Dupoux et al., 1999). The present study tests two different approaches towards phonotactics for their predictions about perceptual epenthesis effects. While the string-based approach predicts perceptual epenthesis effects due to consonantal contact restrictions in the L1, the syllable-based approach predicts the same effects as a result of L1 syllable structure conditions.

These two approaches are tested with Korean second language speakers of English whose L1 bans certain consonants in coda position (e.g., *[c], *[g]) while allowing others (e.g., [k], [l]). Korean also disallows certain heterosyllabic contacts such as *[km] and *[ln], which are realized as [ŋm] and [ll] due to nasalization and lateralization, respectively. In a perceptual experiment, clusters that contain (1) coda violations (e.g., *[cm], *[gm]), and (2) contact violations (e.g., *[km], *[ln]) in Korean are used in nonce pairs of words comparing these clusters with their vowel-present counterparts (e.g., *[pʰákma] vs. [pʰákuma]), and also to their likely output forms (e.g., *[pʰákma] vs. [pʰáŋma]).

The results from an AX discrimination experiment using English and Korean listeners indicate that the English group successfully discriminates all clusters. The

Korean group is also successful in most clusters except for those where the cluster incurs a coda violation (e.g., *[c.m], *[j.t]). Finally, the Korean group does not confuse illicit clusters with their likely output forms.

These results show that perceptual epenthesis arises when there is a syllable structure violation, rather than a contact violation, confirming the syllable-based approach. The successful discrimination of clusters with voiced codas (e.g., *[gt]) suggests that voicing is perceptually suppressed leading to a licit coda ([g.t] → /k.t/). That illicit clusters are not misperceived as their likely output forms indicates that assimilation rules are not relevant for perception. Therefore, theories that explain phonological alterations based on perceived similarity between input and output forms (e.g., Steriade, 2001b) cannot account for Korean listeners' perception. Finally, frequency cannot explain present findings because, despite their zero frequency in Korean, only certain illicit clusters induce perceptual epenthesis. Instead, a perceptual model using onset and coda detectors is proposed, and is shown to account for perceptual epenthesis effects in a straightforward way.

Chapter 1

INTRODUCTION

1.1. Introduction

This dissertation is concerned with the influence of first language phonotactic knowledge on the perception of second language consonant clusters. Phonotactics broadly refers to the knowledge about which sound combinations are legitimate and which are not; that is, the "syntax" of sound combinations that make up the internal structure of words. In this dissertation, I propose that the syllable is a crucial component of listeners' knowledge of phonotactics based on empirical evidence from Korean. To achieve a better understanding of its influence on speech perception, I show that the role of syllable structure conditions in the perception of consonant clusters needs to be distinguished from restrictions on sound combinations.

A number of studies have indicated that the way non-native speech is processed is heavily influenced by listeners' knowledge of the phonotactic restrictions in their L1 (Dupoux, Kakehi, Hirose, Pallier and Mehler, 1999; Dahaene-Lambertz, Dupoux and Gout, 2000). For instance, Dupoux and colleagues (Dupoux et al., 1999) show that Japanese listeners report perceiving epenthetic vowels within consonant clusters in nonsense words such as *[e**bz**o] or *[e**gd**o], which contain consonantal contact situations that never occur in Japanese. Dupoux and colleagues have employed Japanese, a language that is limited with respect to consonant clusters due to its conservative syllable structure, in their studies. From a string-based approach towards phonotactics, these nonsense words violate the phonology of Japanese because the consonantal strings *[bz] and *[gd] in these words do not occur in Japanese. To a great extent, however, the same phonotactic violations can be stated with reference to syllable structure. Accordingly, *[e**b**.**z**o] and *[e**g**.**d**o] violate the phonology of Japanese because [b] and [g] cannot occur in the coda position in Japanese. Since Japanese is predominantly a CV language with severe restrictions on coda consonants, syllable structure restrictions and consonantal contact restrictions are confounded in these studies. The exact nature of L1 phonotactic influence on L2[1] speech perception thus remains to be unknown.

My primary goal in this dissertation is to examine and compare the consequences of syllable structure violations and consonantal contact violations for perceptual epenthesis. The present study employs an experiment using Korean, which provides a unique case to dissociate syllable structure restrictions and consonantal contact restrictions. Korean has syllable structure restrictions on the

[1] The term "second language" (L2) generally refers to a language acquired usually in naturalistic settings after the first language has been mastered. In this dissertation, I use the term L2 to refer to a non-native language and to those situations where individuals encounter non-native stimuli. First language (L1) refers to a person's native language.

type of consonants that can surface in coda position (e.g., *[c.] *[s.], etc., where the dot indicates a syllable boundary)[2]. Furthermore, Korean does not allow certain heterosyllabic consonant contact situations (e.g., *[k.m], *[l.n], etc.) although the first consonant in these contacts can surface in coda position (e.g., [k.t], [l.t]). Regarding syllable structure conditions, the experiment only focuses on coda violations. It also compares these to contact violations.

Based on the results of the present study, I show that Korean listeners' perception of consonant clusters can only be explained with reference to the restrictions on coda position in Korean. String-based phonotactic restrictions, in contrast, fail to account for the Korean findings in the present study.

There are several other consequences of the present study for existing phonological theories and speech perception models. First, I confirm that speech perception is not isomorphic to speech production (see Boersma, 1998 for similar conclusions). I show that Korean listeners do not misperceive illicit sequences (e.g., *[k.m]) as their likely output forms in Korean phonology (e.g., *[k.m]→[ŋ.m]). That is, Korean listeners' perceptual behavior does not parallel the specific phonological

---

[2] In this dissertation, Korean voiceless alveo-palatal affricate consonants are represented with [c] in order to be consistent with the way these sounds are traditionally transcribed in the Korean linguistics literature. Accordingly, /c/ with no diacritic marker indicates a plain consonant, as opposed to tense (e.g., /c'/) and aspirated (e.g., /c^h/) in Korean. The symbol [j], on the other hand, is used to represent the voiced allophone of the plain alveo-palatal affricate consonant (i.e., [c]). For English examples, however, widely accepted IPA symbols are adhered to unless otherwise noted.

rules of Korean (in this case, nasalization) that apply to such illicit consonantal contact situations in production. Consequently, I argue that the P-map Hypothesis (Steriade 2001 a, b), which derives phonological phenomena from listeners' knowledge of the perceptibility of contrasts, makes false predictions regarding Korean phonology.

Last, I discuss the present findings in the context of theories that make frequency-based predictions. I show that certain illicit consonantal sequences in Korean (e.g., *[l.n], *[k.m]), which presumably have zero probability of occurrence, do not cause perceptual epenthesis effects for Korean listeners. Frequency, therefore, does not predict the present results. Instead, I propose a portion of a speech perception algorithm that incorporates the knowledge of syllable structure. Using syllable-based concepts such as onset and coda, the algorithm predicts the perceptual epenthesis facts in a simple and straightforward manner.

1.2. Organization of the Dissertation

Chapters 2 and 3 of this dissertation lay the grounds for the present study. Specifically, I discuss relevant studies on the perception of consonant clusters and their shortcomings in Chapter 2. I motivate the present study on Korean listeners' perception of consonant clusters in Chapter 3. The chapter discusses various phonotactic restrictions and phonological processes of Korean that are relevant to the issue at hand. Chapter 4 lays out the research questions and hypotheses that are tested in a speech perception experiment. The design of the experiment, the nature of stimulus items, procedures, and subject populations are discussed in Chapter 5. Chapter 6 presents the results of the empirical test and summarizes the findings. I also consider alternative interpretations of the data from the present study in Chapter 6. Finally, the implications of the study are discussed in Chapter 7. The first part of Chapter 7 discusses the findings of the study in the context of the P-map hypothesis (e.g., Steriade, 2001a, b) and theories that make predictions based on frequency of speech strings in the language. I show that these two views cannot adequately explain Korean listeners' perception. In the second part of the chapter, I propose a speech perception algorithm that employs syllabically conditioned feature detectors. Chapter 8 concludes the dissertation with remaining questions that are pending future investigation and suggestions for further research on the perceptual processing of consonant clusters.

Chapter 2

NATIVE LANGUAGE PHONOTACTIC KNOWLEDGE AND THE PERCEPTION OF CONSONANT CLUSTERS

2.1. Introduction

The primary aim of this chapter is to provide a basic understanding of some of the studies that were conducted specifically to investigate the influence of phonotactic restrictions. The majority of the discussion in this chapter is devoted to recent perceptual studies on Japanese since they directly investigate the issue cross-linguistically and bear assumptions that also underlie the present study.

2.2. The Role of Language Specific Phonotactic Restrictions in Speech Perception

Speech perception is a prerequisite for phonological development. It plays a crucial role in acquiring the sound patterns of our first language. Our speech perception abilities, however, seem to become more specialized after having achieved the mastery of our first language (see Jusczyk, Houston and Goodman, 1998 for a review of studies on infants' early speech perception abilities). It has long been observed that the internalized knowledge about the sounds and how they pattern in the

L1 may strongly influence our perception of non-native speech (Polivanov, 1932; Sapir, 1933; Trubetzkoy, 1939/1969).

Knowledge about the phonotactic regularities of our native language seems to emerge at a very early point in language development. For instance, Jusczyk, Luce, and Charles-Luce (1994) investigated infants' sensitivity to sound sequences that occur with different frequencies within English words. Jusczyk and colleagues calculated phonotactic probability based on positional phoneme frequency and bi-phone frequency (i.e., the phoneme-to-phoneme co-occurrence probability) using an online version of a dictionary. They found that 9-month olds, but not 6-month-olds, listened longer to CVC nonwords with high-probability phonotactic sequences (e.g., [ʃæn]) than to ones containing low-probability phonotactic sequences (e.g., [jɔb]). Their findings indicate that between 6 and 9 months of age, infants attuned to the sound combinations that typically occur in their native language. Similarly, Friederici and Wessels (1993) presented 4.5 - 6-, and 9-month-old Dutch infants with pairs of monosyllables with identical phonetic segments but with different orderings of these segments. Specifically, each pair contained a given cluster of consonants (e.g., [br]), which occurred at the onset of one set of stimulus items (e.g., [bref]) and at stimulus offset in another set (e.g., [febr]). The clusters used in the experiments occur in Dutch either at word onset and never at word offset (e.g., [br]), or at word offset and never at word onset (e.g., [rt]). Only 9-month-olds showed a listening preference for stimuli containing the cluster in a permissible position (e.g., [murt] or [bref]) over the stimuli containing the same cluster in an impermissible position (e.g., [febr] or [rtum]). Although a few languages permit such clusters (e.g, Russian permits [rt] and [br] both syllable initially and finally; French has [br] in coda position)[3], the illicit clusters in these stimulus items also violate general syllable structure patterns, specifically the Sonority Sequencing Generalization, not just those of Dutch. Therefore, it is not clear whether the infants' preference to listen to the possible clusters in this experiment is due to their sensitivity to their native-language patterns.

These findings have been extended in a crosslinguistic study done by Jusczyk and colleagues which compared Dutch and American infants (Jusczyk, Friederici, Wessels, Svenkerud, and A. Jusczyk (1993). Both Dutch and English are similar in prosody but they differ in certain phonotactic restrictions. For instance, both languages are stress-timed and display an opposition between strong and weak syllables, and the pitch, amplitude, and durational correlates of stress in both languages tend to pattern the same way (Crystal & House, 1998). As for phonotactic differences, Dutch permits sequences like [kn] at the onset of syllables while English does not. Jusczyk and colleagues used two- to three-syllable-long low-frequency words for each language. Most of the words in each list for each language violated the phonotactics constraints of the other language. For instance, Dutch allows words to end in [mst] as in 'opkomst' but English does not have any words that end in such a cluster. Furthermore, while English allows words to end in voiced stops such as [b],

---

[3] Thanks to John Kingston for bringing these examples to my attention.

Dutch does not. It was shown that 9-month olds, but not 6-month-olds, preferred listening to words that accord with the phonotactic constraints of their native language. The converging evidence from these infant speech perception studies shows that children by the first year of life become sensitive to the phonotactic regularities of the language that they are acquiring.

There is an abundance of literature on the role of phonotactics in adult speech perception. Most of the research in this field is on monolingual perception and studies typically focus on response biases in favor of perceiving ambiguous sound combinations as legal sequences, a pattern which is assumed to reflect the role of phonotactic knowledge in perception (e.g., Massaro & Cohen, 1983; Hallé, Seguí, Frauenfelder & Meunier, 1998; Pitt, 1998). These studies typically make use of steps along a continuum, such as /l/-/r/, and embed each step in a consonantal context in such a way that the consonant from the continuum (i.e., the liquid), is the second consonant of a two-member consonant cluster (e.g., [tri]). The clusters are formed in such a way that one endpoint of the continuum creates a phonotactically legal word-initial cluster in English (e.g., [tr.], [sl.]) and the other endpoint creates an illicit one (e.g., *[tl.], *[sr.]). Massaro and Cohen (1983) found that listeners' classification responses showed a bias in favor of legal sequences. For instance, the steps at the /l/ endpoint were identified frequently as /r/ when preceded by /t/ but less often as /r/ when preceded by /s/. Those at the /r/ endpoint, on the other hand, were heard as /l/ when preceded by /s/, but as /r/ when preceded by /t/. Massaro and Cohen interpreted their findings to suggest that knowledge of the phonotactic constraints of English affected perception of the liquid by English speakers.

More recently, studies on phonotactics and speech perception have been more concerned with the source of phonotactic information, such as frequency and lexical effects (e.g., Luce & Pissoni, 1998; McClelland & Elman, 1996; Norris, 1994; Pitt & McQueen, 1998; Vitevitch and Luce, 1998). Research in this area has typically investigated probabilistic phonotactic information, which refers to the frequencies of segments and sequences of segments in syllables and words. For instance, Vitevich and colleagues (Vitevitch, Luce, Charles-Luce & Kemmerer, 1997) looked at the effects of probabilistic information on processing times for spoken stimuli using bisyllabic nonce words with legal phonetic sequences in English. These words varied in their segmental and sequential probabilities. Vitevitch and colleagues used a speeded auditory naming task where the subjects heard a spoken stimulus and had to repeat it as quickly and accurately as possible. They found that words that were composed of common segments and sequences of segments were repeated faster than those with less common segments and sequences. The same study also involved subjects' preference ratings for spoken nonce words. Vitevitch and colleagues found that the subjects' ratings mirrored the pattern of results from their reaction time experiment: nonce words composed of common segments and sequences of segments were rated as being "better" possible words than those composed of less common segments and sequences.

A recent study by Moreton (2002) held frequency constant while manipulating structural constraints on English onset clusters. The English stop-sonorant clusters [dl] and [bw] , which are unattested in English, were used as syllable onsets. Moreton claims that while [dl] is impossible (e.g., English bans [coronal] [coronal] sequences), [bw] is "marginal" in English (see Moreton (2002: 56-57) for a detailed discussion on this). Models that distinguish attested from unattested consonant cluster configurations predict equal biases for these two onsets because both are unattested in English. On the other hand, structural models predict a larger perceptual bias against [dl] than against [bw]. Moreton's experimental findings indicated a perceptual bias against [dlæ], but not against [bwæ], disconfirming those models which attribute bias to the frequency of clusters. The experiment's results favor models that emphasize the use abstract featural generalizations over frequency information. Further discussion of this study will be presented in Chapter 7 in the context of the findings from the present study, which also indicates that frequency cannot explain Korean listeners' successful perception of certain consonant clusters.

In short, the results of the studies cited above demonstrate that listeners' information about the legality as well as the probability of phonotactic patterns seem to influence the processing of spoken stimuli. Given the possibility that frequency may influence listeners' judgments, the present study will consider these effects in interpreting the results of the perceptual experiment.

As language-specific phonotactic phenomena seem to shape infants' perceptual abilities as well as adult's perceptual decisions, would we expect it to also play a role in L2 speech perception? From the production side, there are several pieces of anecdotal evidence reflecting the ways in which L2 speakers accommodate non-native sound sequences that are not legal in their first language. For instance, Spanish speakers are known to produce the vowel [e] before English words that begin with an /s/+ Consonant sequence. This is perhaps no surprise given that Spanish does not allow such clusters in syllable onsets. Similarly, Mandarin L2 speakers' tendencies to drop coda consonants in English words (e.g., [dai ko:] 'diet coke', [gu:] 'good') are frequently noticed as being markedly non-native. Coda dropping in these pronunciations can be seen as a result of the fact that Mandarin has only nasals and glides in the coda position. The question remains as to whether what these speakers produce could be a reflection of what the way they perceive L2 forms.

2.3. Studies on the Perception of Consonant Clusters by Japanese listeners

Japanese is perhaps the most frequently cited language that disallows complex consonant clusters. This restriction appears to be found not only in word initial positions but at every position in a word due to the simple syllable structure of the language. The syllable inventory of Japanese is V, VV, CV, CVV, CVN and CVG_{1}, where V is a vowel, C is a consonant, N is a nasal and G_1 is the first half of a geminate consonant. When consonants come into contact with other consonants, such as in loan words, they create impermissible consonant clusters in Japanese. These illicit clusters are then adapted into the native language phonology, typically, by the insertion of an epenthetic vowel, which breaks up the illegal cluster. As can be seen in (1), the epenthetic vowel is [o] after a dental stop, and elsewhere it is [ɯ].[4]

(1) a. MacDonald → mak**u**donar**u**d**o**

b. brother → b**u**raza:

c. volcano → bor**u**kano

d. fight → fait**o**

e. festival → fes**u**tibar**u**

Epenthesis in loan words shows that speakers obey phonotactic constraints in their production. Could it also be possible that Japanese speakers apply vowel epenthesis perceptually? Dupoux, Kakhei, Hirose, Pallier & Mehler (1999) tested this question in a cross-linguistic design in which they compared Japanese speakers with French speakers in their perception of consonant clusters. In their first two experiments, Dupoux and colleagues presented their participants with an off-line phoneme detection task and instructed them to respond whether each item they heard

[4] As an orthographic convention, the high back unrounded vowel /ɯ/ is usually represented in the Roman alphabet as <u>. For convenience, in the examples given in (1) the symbol <u> is used. It should be noted that these examples reflect a broad transcription and do not necessarily show the actual pronunciations of these words. The crucial point, however, is to illustrate vowel epenthesis in these loan words.

contained the sound [u]. Ten naturally produced nonsense items containing the sequence VC_1 u C_2 V were recorded (e.g., *ebuzo, egudo, abuno, akumo*). These words were manipulated to produce a series of six items from each one. Each set of six items included the original item (that is, the item that was naturally produced with the vowel [u]) and five other items modified from the original item such that the duration of the vowel [u] was gradually shortened to lengths of 72 ms, 54 ms, 36 ms, 18 ms, and 0 ms, respectively. The results from the phoneme detection task revealed that, unlike French participants, Japanese participants predominantly judged that the vowel was present at all levels of vowel length. This was even true for the condition where the vowel had been completely removed. In this case, the Japanese listeners reported that the vowel was present in more than 70% of the presentations. In contrast, the French participants, as expected, judged that the vowel was absent in the no-vowel condition about 90% of the time and that a vowel was present only in 50% of the intermediate cases. Figure 2.1 shows the steady decrease as a function of decreasing vowel length for both French and Japanese subjects.

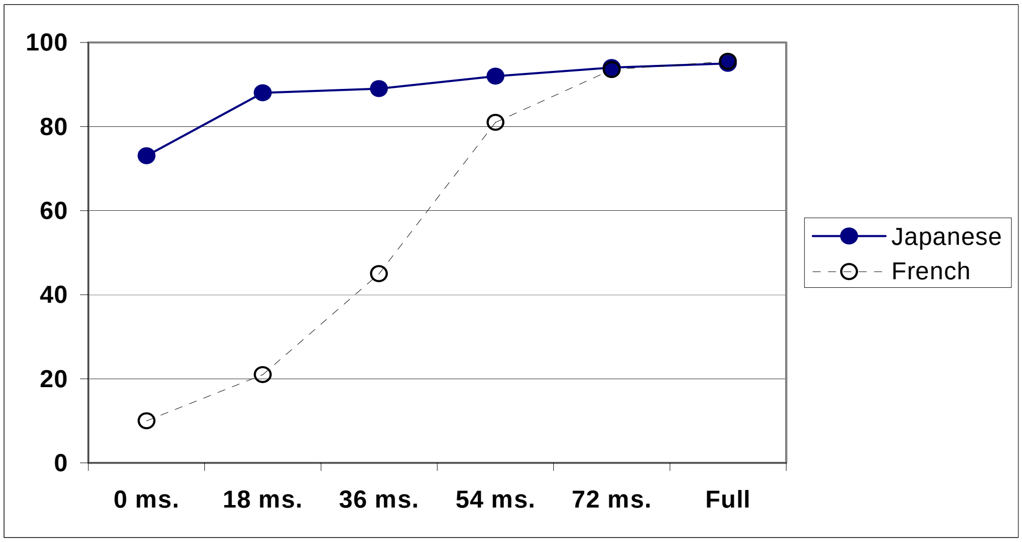

Figure 2.1: Percept [u] vowel judgments as a function of vowel duration (after Dupoux et al., 1999: Experiment 1)

The second experiment in the study had the same items as well as the same paradigm as in the first experiment except that the experimenters added two extra conditions: one with no vowel (e.g., a naturally produced *ebzo*) and one with a vowel other than [u] (e.g., *ebizo*). While the former was introduced to control for any co-articulation cues for the rounded vowel that are likely to be present on the surrounding consonants, the latter was introduced to measure the baseline performance. The results of the second experiment showed that the Japanese participants did not perceive more [u] vowels in the digitalized than in the natural cluster condition. The baseline condition (i.e., ebizo) yielded almost no [u] judgments from both groups of listeners.

Dupoux and colleagues also ran experiments in which they presented the subjects with speeded ABX paradigms where they heard three stimuli in a row and were told to decide whether the third stimulus was the same as the first or the second. In addition, they created an extra condition, in which they used a long [u] instead of the short one (e.g. *ebuuzo*). Unlike in French, vowel length is contrastive in Japanese. By adding vowel length, the researchers constructed a complete crossover design in which both language groups would impose the phonology of their native language on unfamiliar linguistic stimuli. Sixteen triplets were constructed conforming to the model V_1 C_1 C_2 V_2 (e.g., *ebzo*), V_1 C_1 U C_2 V_2 (e.g., *ebuzo*), and V_1C_1 UU C_2V_2 (e.g., *ebuuzo*). One hundred and twenty eight experimental trials were constructed using these 16 experimental triplets. These triplets yielded experimental trials consisting of the presentation of the three stimuli A, B, and X with an inter-stimulus interval of 500 ms. The stimuli corresponding to A and B in the design were taken from the same triplet but differed depending on whether the vowel was absent, present, or long. The experimental items were recorded by two native speakers of Japanese, who had some training in phonetics and spoke French fluently. The first two words in the presentation were spoken in a female voice, and the last one in a male voice. In the production of some items, however, these speakers could not prevent themselves from inserting an epenthetic vowel into clusters. The vocalic part of these words, thus, had to be digitally removed. Unlike in the first experiment, where the language of the test stimuli was not told to the subjects, the subjects were told that the words were from a foreign language and the purpose of the experiment was to test their intuitions about the sounds contained in the words. The analysis of the error data showed that the French participants had more difficulty with the vowel length contrast than the epenthesis contrast ($MinF'(1,16)=18.11$, $p<.001$), while the Japanese participants encountered relatively more difficulty with the epenthesis contrast, although this difficulty was only significant in the items analysis ($F1(1,18)=3.71$, $p=.07$; $F2(1,15)=20.17$, $p<.001$). Taken together, Dupoux and colleagues' experiments demonstrate that not only do Japanese listeners report more vowels than are really present in the signal, but their ability to distinguish two stimuli, one that has a vowel and one that does not, is also strongly affected.

The Dupoux et al. 1999 study established yet another piece of evidence for the notion that the way in which a continuous speech stream is processed in the mind can be heavily influenced by the typical patterning of sounds in the first language of the listener. What is more crucial is that the study has shown that listeners tend to make adjustments in processing in such a way that they may "invent" illusory vowels to accommodate an unusual speech string in their native language phonology.

An alternative explanation for the perception of illusory vowels in the Dupoux et al. (1999) study is one that refers to the listener's lexical knowledge. Of particular relevance to this explanation are the existing approaches towards phonotactics in the speech processing and spoken word recognition literature. In the 1980s, the central objective of research that particularly focused on spoken word recognition was to tease apart the types of linguistic information that are potentially utilized in perceptual decisions. Researchers typically questioned whether acoustic featural information and lexical context make independent contributions to perceptual recognition (Ganong, 1980; Massaro & Oden 1980 a, 1980 b; Massaro & Cohen, 1983). From that perspective, a number of researchers showed that a listener's knowledge of phonotactic patterns can be helpful in providing a rich source of information that could be used to facilitate auditory word recognition in the first language (e.g., Church, 1987; Frauenfelder & Lahiri, 1985). For instance, English speakers may encounter illicit consonant clusters due to a well-known schwa-deletion phenomenon. Thus, *tomato* can turn into [tmato], yielding an illegal [tm] cluster in casual speech. However, knowledge of the English lexicon could be used to recover the intended vowel, leading to successful recognition of the word.

Likewise, it could be that the Japanese failure to perceive clusters is due to the frequency with which these clusters occur in the language. Accordingly, phonological processing per se may not be necessary to produce phonotactic effects. Instead, as in the TRACE model of spoken word recognition (McClelland & Elman, 1986), potential biases towards perceiving legal sequences can arise through a top-down activation of phonemes due to a general frequency effect. In this model, all lexical entries are activated when they match the speech input. We can assume that there are more lexical entries with legal than illegal sequences (presumably zero of the latter). The effect of frequency in activation of lexical entries makes strong predictions regarding the perception of legal clusters with low probability. Thus, the perceptual illusions of Japanese listeners could have resulted from top-down lexical influences. After all, there are many Japanese words that contain the sequence C_1uC_2. In fact, this alternative explanation is supported by previous studies showing that lexical knowledge can yield the perception of a phoneme that is not present in the signal (e.g., Ganong, 1980; Samuel, 1987). A follow up study by the same group of researchers (Dupoux, Pallier, Kakehi, and Mehler, 2001) tested this alternative explanation in experiments that involved a transcription task. The participants were asked to transcribe the words into the Roman alphabet and, furthermore, to complete a lexical decision task in which they had to decide whether the items were words in Japanese or not. Dupoux and colleagues created non-words containing illegal consonant clusters in Japanese, yielding only one lexical neighbor when a vowel is inserted between the consonants. For some items, the lexicon biased the insertion of the /u/ to generate an actual word (e.g. *sokdo* → *sokudo* 'speed', **sokado, *sokedo, *sokido, *sokodo*); in others, the lexicon biased the insertion of a vowel other than /u/ (e.g. *mikdo* → *mikado* 'emperor', **mikudo, *mikodo, *mikido, *mikedo*). It was expected that if Japanese listeners' perception is influenced by the "nearest" real Japanese word, then, for instance, a nonsense word with an illicit cluster such as *sokdo* should be perceived as *sokudo*, while *mikdo* should be perceived as *mikado* since the insertion of /a/ would be the only way to obtain an already existing Japanese word. They found that Japanese speakers reported hearing a vowel /u/ between the consonant clusters irrespective of the lexical status of the perceptual outcome (i.e., they reported hearing *mikudo* even though this is not a word of Japanese). The results obtained from the lexical decision task showed that participants classified stimuli such as *sokdo* as real words, and stimuli like *mikdo* as non-words. These results indicate that the illusory vowel effect in Japanese is not due to "top-down" lexical influences, but rather it arises from language-specific pre-lexical processes. In other words, phonological context is sufficient to trigger the illusory perception of /u/.

That the perception of the epenthetic vowel /u/ is a pre-lexical process is further supported by an electrophysiological study conducted by Dehaene-Lambertz and colleagues (Dehaene-Lambertz, Dupoux & Gout, 2000). The researchers used electroencephalagraphic (EEG) event-related potentials (ERPs) in a cross-linguistic design involving once again Japanese and French participants. An oddball paradigm was employed in which each trial consisted of four similar precursor stimuli, followed by a test stimulus that was either identical to or different from the previous stimuli, depending on the presence or absence of the vowel /u/ between the two consonants (e.g., *ebuzo-ebuzo-ebuzo-ebuzo-ebzo*). The Mismatch Negativity (MMN) is a response that is typically elicited whenever there is a mismatch between the features of the perceived stimulus and the representation in the sensory memory left by the stimuli immediately preceding it. For French subjects, the deviant versus control (i.e., nondeviant) comparisons produced three mismatch responses in ERPs: (1) 139 to 283 ms, (2) 291 to 419 ms, and (3) 523 to 651 ms post-deviance onset. These responses, however, were either absent or had a shorter duration and weaker amplitude in Japanese subjects. Specifically, the Japanese subjects showed no early mismatch negativity response (MMN) to the deviant stimulus whereas French subjects did show an early MMN response to the deviant stimulus 140-280 ms after the offset of the first consonant in the cluster. Dehaene-Lambertz and colleagues interpreted the early MMN response and its absence in Japanese listeners to suggest that phonotactic effects emerge very early in the course of speech perception, which is potentially related to the coding of phonetic properties. The second and third responses had somewhat similar voltage cartographies in both the French and the Japanese subjects (see Dehaene-Lambertz, Dupoux & Gout, 2000 for different interpretations of the difference between the first response and subsequent responses). The major finding of this study is that Japanese subjects show no evidence of a deviance effect at the time of the first mismatch response shown by French subjects. Therefore, Dehaene-Lambertz and colleagues concluded that a fast and automatic coding of the speech input takes place early in speech processing. That is, phonotactic constraints influence the brain response to speech strings that do not conform to the native language sound patterns at a very early stage in speech perception.

2.4. Shortcomings of the Previous Studies on Japanese

The studies on Japanese listeners' perception of consonant clusters by Dupoux and his colleagues have contributed a new perspective to our understanding of how the perceptual system is vulnerable to and constrained by our native language grammar. These findings now show that epenthesis, which has traditionally been viewed as a response to a difficulty in production, also arises as a perceptual problem.

Dupoux and colleagues used the term *phonotactics* in the sense common in the spoken word recognition literature, where it refers to the sequential arrangement of phonetic segments in morphemes, syllables, and words. Phonotactics, by this definition, is a component of the grammar that describes the ordering of the basic phonetic units, such as segments, in utterances. When sequences do not conform to this grammar, a phonotactically illegal sequence is obtained. That is, the illegality of *[bz] in Japanese is due to a co-occurrence restriction on these two Japanese consonants.

There is another possible explanation for why this sequence is illicit: the first consonant in the cluster (i.e., [b]) can never surface in the coda position in Japanese. That is, knowing that [b] is an illicit coda, Japanese listeners might repair *[eb.zo] as [e.bu.zo] by positing an epenthetic vowel to put this consonant in onset position, thereby satisfying the coda restriction on [b].

Indeed, very few coda-onset clusters arise in Japanese since it is predominantly a CV language and the Japanese syllable licenses only a few coda consonants under a very restricted set of circumstances. Specifically, a consonant can only occur in coda position if it is the first member of a geminate (2a, b), or a homorganic nasal (2c, d) (Itô, 1986; 1989). Thus, while all the words in (2) are phonologically well-formed, those in (3) are not (examples are from Itô, 1989).

(2) a. kap.pa 'a legendary being'

b. gak.koo 'school'

c. tom.bo 'dragonfly'

d. kaŋ.gae 'thought'

(3) a. * kap.ta

b. * tog.ba

c. * pa.kap

According to Itô, (1986), the restrictions on coda consonants in Japanese can be expressed as a filter referring to the syllable final position and the melody, as can be seen in Figure 2.2 below.

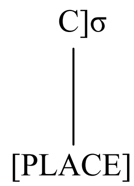

Figure 2.2: Japanese Coda Filter (Itô, 1986; 1989).

Accordingly, the Coda Filter rules out syllables with final consonants that have their own independent place specification. However, since geminates and homorganic clusters are doubly place-linked, they are immune to the Coda Filter. That is, the Coda Filter applies only to a single association line, as the one in Figure 2.2 above. Consequently, permissible syllable-final consonants in Japanese are always place-linked to a following consonant, yielding forms such as [tombo], which is properly syllabified in Figure 2.3.

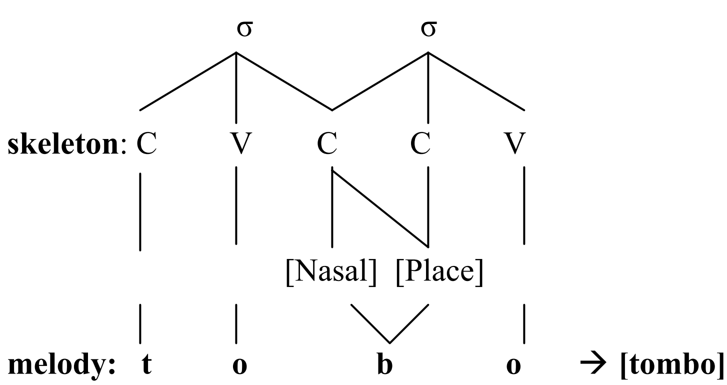

Figure 2.3: Syllabification of [tombo] in Japanese (adapted from Itô (1989)).

Closer examination of the stimulus items used in Dupoux and colleagues' study reveals that all the test words contained impermissible codas in Japanese (e.g. [ab.ge], [iʃ.to], [og.za], [eb.zo], [ek.to][5]). For instance, voiced stops such as [b] and [g], as well as fricatives such as [ʃ] can never occur as coda consonants. Furthermore, voiced stops cannot geminate. Voiceless stops such as [k], on the other hand, can surface as a coda, but only when they are part of a geminate construction (e.g., [gak.koo] 'school'). However, the researchers did not present an analysis of their data for any effects of specific consonants. Therefore, it is not known if the Japanese listeners initially accepted /k/ as a coda consonant in items such as [ik.ma] or [ek.to]. Due to the fact that /k/ (as well as other voiceless stops) can never surface word-finally but can occur in coda position only as the first half of a geminate, it can be assumed that Japanese listeners never encounter singleton /k/ in coda position. Therefore, they must interpret /k/ as an onset consonant. Since Japanese onsets cannot branch, the combination of /k/ with a following /m/ would induce a syllable structure violation, causing perceptual epenthesis to arise to repair the illicit syllable structure.

In short, it is not known whether perceptual epenthesis in Japanese in words such as [egdo] or [ekto] arises due to contact violations (i.e., because *[gd] and *[kt]

[5] John Kingston has suggested that there are cases where heterosyllabic clusters such as [ʃ.t] and [k.t] seem to arise as a result of high vowel devoicing (HVD) in Japanese (e.g., Beckman,1982; Keating and Hoffman, 1984). In the case of [g.z] and [b.z], HVD does not apply because of the voiced consonants. Therefore, HVD would not account for all observed cases of perceptual epenthesis in Japanese.

are illicit in Japanese) or coda violations (i.e., because *[g.] and *[k.] cannot be independent coda consonants). The alternative explanation, which is solely based on the knowledge of syllable structure, relies specifically on listeners' knowledge of coda restrictions in the L1. Coda restrictions simply dictate what features or feature bundles are not permitted as coda consonants. From this perspective, the perception of an illusory vowel is caused by the fact that none of the coda consonants used in the Dupoux et al.'s experiments obeyed the coda restrictions of Japanese. Specifically, the knowledge about the featural content of coda consonants, that is the coda identity of the segments, could have also evoked perceptual epenthesis effects in Japanese.

Last, there is also a contradiction in the literature with regard to a model that Dupoux et al. (1999) refers to for potential explanations for Japanese perceptual epenthesis results. Dupoux and colleagues suggest that their findings can be accounted straightforwardly using SARAH (Mehler, Dupoux and Segui, 1990), an acquisition model that postulates a correspondence between the processes used by infants and those underlying lexical access in adults. SARAH (Syllable Acquisition, Representation, and Access Hypothesis) assumes three levels of processing in adult speech perception: syllabic, lexical, and phonological. The syllabic level segments speech into elementary units (syllabic frames) that roughly correspond to the syllable. Syllabic frames are the minimal functional units relevant to speech production and are recognized by a bank of syllabic detectors. At the Lexical Level, the bank of syllable detectors makes up the code that is used to access the lexicon. Phonemes, which do not play a direct role in speech perception, are extracted from the pre-lexical syllabic frames at the Phonological Level (See Mehler, Dupoux, and Segui, 1990: 255-256 for the specific mechanisms that SARAH proposes for the way syllable detectors are used to construct potential lexical entries by infants). Using the basic premises of SARAH, Dupoux and colleagues explain perceptual epenthesis effects in the following way. Faced with a foreign language, the perceptual system tries to parse the signal using the available L1 syllabic categories. However, Japanese lacks syllable categories containing consonant clusters or coda consonants. Therefore, a stimulus like /ebzo/ activates syllable categories for /e/ (instead of /eb/) and /zo/. It also activates all syllables that start with /b/ (i.e., /bu/, /ba/, /bi/, and /bo/, etc.). The syllable /bu/ emerges as the best match since /u/ is frequently shortened or devoiced and it shows considerable allophonic variation in Japanese. This is, however, an incomplete explanation because /i/ also frequently devoices in Japanese. Furthermore, by resorting to a model that relies on syllable repertoires, Dupoux and colleagues are moving away from their original goal. That is, if the knowledge of the repertoire of syllables, which includes the totality of all possible syllables in the L1, caused listeners to find the best match in syllable type for the incoming acoustic signal, Japanese listeners would then never have to make use of their (in one sense) phonotactics knowledge about what consonant sequences are licit or illicit in Japanese.

Since top-down models crucially rely on the high frequency of words with legal clusters (as in TRACE) or knowledge of all syllables that occur in the entirety of the L1 (as in SARAH), they predict that those cases where a given syllable occurs in the language with a low or zero probability will be problematic. In fact, there are consonant clusters with zero probability, which are still permissible across languages. For instance, there are no English words that begin with /stw/ and, therefore, no syllable frames that correspond to /stwa/, /stwo/, and /stwi/, etc can be construed for English. Yet, the sequence /stw/ constitutes a potential onset cluster since it fits the sonority profile of possible English onset clusters (e.g., /str/ occurs in English as onset cluster) and English speakers do not have any trouble with this cluster. Analysis based on the frequency of syllable types in the lexicon or a list of syllable frames, therefore, cannot predict English speakers' behavior on this possible but virtually absent clusters. I will show that frequency-based analyses also do not account for the present findings about Korean listeners' successful perception of illicit consonant clusters. This will be discussed extensively in Chapter 7.

2.5. Summary

The accounts proposed so far for the observed differences between the Japanese and French listeners in their perception of consonant clusters attribute epenthesis effects broadly to the listeners' knowledge of the permissibility of sound sequences in the L1, a notion typically referred to as phonotactics. However, all the evidence for the influence of phonotactic knowledge also converges on listeners' knowledge of the syllable structure restrictions in their L1.

I have shown above that two different but not necessarily independent phonological phenomena can adequately explain the Japanese results: (1) L1 restrictions on the co-occurrence of consonants (i.e., consonantal contact violations), and (2) the well-formedness of a consonant occupying a particular position of a syllable (i.e., syllable structure violations such as coda or onset violations). It appears that both of these factors could be independently contributing to the perception of epenthetic vowels in the previous studies on Japanese.

The present study aims to tease apart consonantal contact restrictions from coda restrictions by employing Korean. I will show that Korean listeners hear epenthetic vowels when there is a syllable structure violation in the form of an illicit coda consonant. Consequently, I will argue that perceptual epenthesis emerges to create well-formed syllables rather than to create licit sequences of consonants.

Chapter 3

MOTIVATING THE PRESENT STUDY: KOREAN PHONOTACTICS

3.1. Introduction

In this chapter, I present the necessary theoretical considerations and assumptions that will set the ground work for an empirical study that can tease apart consonantal contact restrictions from coda restrictions.

First, it is necessary to find a language that employs two kinds of consonants: those which are permissible codas and those which are not. Moreover, it is crucial to have the same legal coda consonants yield both permissible and impermissible clusters when they are followed by other consonants. Experimental items with these consonant clusters need to be prepared in the native language of a control group where none of these consonant clusters yields any phonological violations.

Table 3.1 outlines the necessary conditions for experimental clusters in the form of a heterosyllabic $C_1.C_2$ contact and the status of the contact in the ideal language to be tested. It is crucial to understand that while $C_{(x)}$ is a permissible coda consonant in the table, $C_{(y)}$ is not. Therefore, the combination of $C_{(y)}$ with any following consonant produces a bad cluster. $C_{(x)}$, on the other hand, produces a licit cluster only when it is followed by a certain set of consonants (i.e., $C_{(a)}$). When it is combined with $C_{(b)}$, the $C_1.C_2$ sequence becomes an illicit one.

Table 3.1: Heterosyllabic $C_1.C_2$ sequences necessary to separate contact violations from coda violations.

<table>
<tr><td></td><td></td><td colspan="2">C_2</td></tr>
<tr><td></td><td></td><td>C(a)</td><td>C(b)</td></tr>
<tr><td rowspan="2">C_1</td><td>C(x)</td><td>OK</td><td>BAD</td></tr>
<tr><td>C(y)</td><td colspan="2">BAD</td></tr>
</table>

The present study will employ Korean and English. Korean provides a number of consonant clusters that fit the prerequisites outlined in Table 3.1. It not only bans certain consonants from certain syllable positions, but it also puts phonological restrictions on what other segments the very same segment may or may not co-occur with. With respect to the positional identity of consonants, Korean has an absolute ban on strident consonants such as /c/ (voiceless palatal affricate) occupying the coda position. When /c/ occurs in morpheme-final positions underlyingly, it neutralizes to [t] (e.g., /nac/ → [nat] 'daytime'; /ic+ta/→ [it.t'a] 'forget'). However, Korean does not prohibit all consonants from surfacing in the coda position. For example, there is no such restriction on /k/ in this position (e.g., /pak/→[pak] 'gourd'; /sək.t^{h}an/ → [sək.t^{h}an] 'coal'). Interestingly, however, /k/ can form permissible heterosyllabic clusters only with certain segments. For example, a sequence of [k] and [m] can never surface in Korean. Instead, the sequence /k.m/ is only licensed if /k/ undergoes nasalization (i.e., /k.m/→[ŋ.m]). In contrast, English does not seem to impose any such substantial restrictions on the co-occurrence of consonantal segments across syllables. In the following section, I present the assumptions related to the phonotactics of Korean and English that are employed in the present study. Since the study focuses on a particular kind of coda-onset cluster found across syllables rather than those found within the syllable (e.g., rhyme or onset), the discussion is restricted only to coda-onset interactions in the two languages.

3.2. Phonotactics of Korean and English

Phonotactics, in its most commonly used sense, refers to the restrictions that govern sound sequences in a language. The specification of a phonological domain in which these restrictions hold, however, is not an essential part of this definition. Alternatively, starting from the smallest relevant phonological constituent, the syllable can be considered to define precisely at which location certain sequential restrictions on sounds are established. In that respect, the coda or onset positions, the subconstituents of a syllable, may precisely anchor where particular restrictions on the consonant are observed. For instance, while three-member consonant clusters in English must start with /s/ in the onset position (e.g., *spread, street, square, splash*, etc.), the same or reverse requirement does not hold for coda clusters (e.g., *burst*). Likewise, the defective distribution observed with regard to consonantal phonemes such as /ŋ/ in English and the alveo-palatal affricates (/c, c', c^h/) in Korean can be specified with reference to concepts such as coda and onset in these two languages. Simply, while /ŋ/ is not allowed to occur in the onset position in English, the onset is the only position that alveo-palatal affricates are allowed to surface in Korean.

Concerning possible heterosyllabic contacts, that is coda plus onset combinations, Korean and English exhibit different restrictions. At this point, it is useful to refer to Ewen and Van der Hulst's (2001) discussion of English medial clusters. The authors illustrate the sequence /-pkm-/ as an ill-formed English medial cluster. Ewen and Van der Hulst entertain the idea that part of the phonological knowledge of a native speaker involves the specification of which consonant clusters are ill-formed in English, giving the examples in (1) (p. 123).

(1)	*initial*	*medial*	*final*
	*km-	*-pkm-	*-pk
	*mr-	*-kmr-	*-km
	*mw-	*-tnw-	*-tn

A fairly obvious redundancy arises in the above clusters. Consequently, Ewen and van der Hulst note that the constraints on medial clusters in English are not independent of those on initial and final clusters. Since no syllable will begin with /km/, /mr/, or /mw/ in English, these illicit onset clusters can never form medial clusters when preceded by syllables that end with consonants (e.g., /-t/+*/mw-/ → */-tmw-/). Likewise, a sequence of /pk/, /km/, or /tn/ is impossible in the coda position in English. Thus, such clusters also will never be able to come together with consonant-initial syllables (e.g., /*-pk/+/-m/ → */-pkm-/). Ewen and van der Hulst make the important observation that medial clusters are not completely independent from the syllables that form them. Consequently, beyond the restrictions that English has at the level of a syllable, English seems to place no constraint on coda-onset contact. That is, given the condition that syllables that are put together to form a phonological word must contain well-formed codas and onsets, no illicit medial clusters will arise. It should be noted, however, that not all consonantal combinations are observed in English. For instance, while there are monomorphemic words with -kt- (e.g., *doctor*, *October*, *Victor*, etc.), the reverse sequence -tk- is very difficult to find in the language (except for *catkin*, or proper names such as *Atkin* or *Ratko*). Despite its rarity, English speakers do not find the sequence very difficult as evidenced by the pronunciation of loan words containing this cluster (e.g., *kamchatka*).

Korean, in contrast, integrates cross-morphemic phonotactic relations in the same phonological domain in which the language also handles morpheme-internal phonotactics. That is, phonological rules at the level of a syllable that dictate the shape of consonantal sequences seem to be duplicated for those that hold across-morphemes in Korean. To be more precise, inflected or derived words provide potential sites for heterosyllabic consonantal contacts, as consonants belonging to different morphemes come into contact at morpheme junctures. The rules that apply morpheme internally, such as nasalization, also apply across morphemes. Therefore, for instance, a sequence of /C+N/ never surfaces across words (e.g., /puək^h/+/mun/ → [pu. əŋ.mun] 'kitchen door'; /os/+/noŋ/→[ot.noŋ]→[on.noŋ] 'clothes chest') except where there is an Intonational Phrase (IP) boundary.

In English, however, although some word-internal consonant clusters such as /gt/ are nearly impossible to find, compounds with such clusters *do* exist in the English lexicon without undergoing any phonological alterations (e.g., *pigtail, ragtime*). That is, English does not impose any phonological rules on unusual combinations of consonants in order to make them fit into the regular patterns of the language. Considering the productive nature of the English compound formation, virtually any consonantal combinations are possible across morpheme boundaries. (e.g., [ʒ.b]: *bei**ge*** ***b**ag*; [f.ʃ]: *bee**f** **sh**op*; [θ.m]: *Be**th** **M**iller*, etc.)[6]. Moreover, many unusual heterosyllabic combinations of consonants in loan words, novel words, and acronyms can easily be adapted into English without modifications and pronounced without any difficulty by English speakers (e.g., [ŋ.j]: *Pyo**ngy**ang*; [g.d]: *Ma**gd**eburg*; [t.k]: *do**t***

---

[6] Especially in casual speech, postlexical assimilation processes such as place and manner of articulation assimilation and consonant elision may optionally apply to heterosyllabic consonantal combinations (e.g.; [d.j]/[dʒ.(j)]: would you; [s.ʃ]/[ʃ:]: this shoe, etc.).

com, etc.). As it stands, since English seems to posit virtually no restrictions on the type of *coda-onset* clusters, the discussion in the following section will focus on Korean, which reveals some crucial phonotactic restrictions for the purposes of this study.

3.3. Restrictions on Consonantal Contact in Korean

Korean has nineteen consonants, two semi-vowels and ten vowels (front rounded vowels are not found in all dialects of Korean). There is a three-way contrast for oral stops and affricates, namely plain (C), aspirated (C^h), and tensed (C'). Fricatives, however, only have a two-way contrast, namely /s/ and /s'/. The following tables show the consonantal inventory (Table 3.2) and vowel inventory (Table 3.3) of Korean.

Table 3.2: Consonantal Inventory of Korean

		Labial	***Alveolar***	***Alveo-palatal***	***Velar***	***Glottal***
	Plain (Lenis)	p	t	c	k	
Plosives	***Tense (Fortis)***	p'	t'	c'	k'	
	Aspirated	p^h	t^h	c^h	k^h	
Fricative	***Plain (Lenis)***		s			h
	Tense (Fortis)		s'			
Nasal		m	n		ŋ	
Liquid			l			

Table 3.3: Vowel Inventory of Korean

	Front		***Back***	
	Unround	Round	Unround	Round
High	i	y	ɯ	u
Mid	e	ø	ə	o
Low	ɛ		a	

Korean syllables are maximally CGVC. The onset and coda cannot branch in Korean; therefore, there are no complex onsets or codas in the language. The glide (G) in CGV is part of the nucleus. As mentioned before, the phonology of Korean places restrictions on the positional identity of its segments in relation to the syllable structure. For instance, only seven (i.e., [p, t, k, m, n, ŋ, l]) among nineteen consonantal phonemes can appear as a coda on the surface. When these coda consonants come in contact with other consonants at morpheme junctures, they undergo certain assimilatory processes. Even consonants that do not appear in coda position interact with following consonants (e.g., /h/+/t/→[t^h]: /coh/+/ta/→[cot^ha] 'is good'). These processes may alter the segmental realization of the consonant cluster on the surface, although the consonants that created the contact would independently be able to surface if they occurred elsewhere. Table 3.4 at the end of this section lists all surface forms of Korean consonantal phonemes that may come into contact across morphemes. In the following, the assimilatory processes that are relevant for the distribution of Korean consonants in consonant clusters will be explained and illustrated.[7] In the following discussion, the reader is recommended to simultaneously refer to Table 3.4 to see how each Korean consonantal phoneme in the C_1 position surfaces before other consonantal phonemes.

First, due to *coda neutralization*, the number of coda consonants that can possible surface in Korean is reduced. It should be noted that this process applies regardless of what follows the consonant in the next syllable. Coda neutralization, a weakening process, states that all consonants must be plain in coda position. In syllable-final position, therefore, the three-way distinction among plosives and the two-way distinction in alveolar fricatives are neutralized, yielding only plain versions of these consonants on the surface. For instance, the base (morphophonemic) forms /cip^h/ 'straw' and /cip'/ 'house' become homophonous when they occur independently as words (i.e., [cip] 'straw/ house').

Another restriction on consonantal contact in Korean arises due to the rule of *tensification*, a strengthening process. The plain series of the obstruent consonants is tensified when they are preceded by a plosive. While the coda plosives are neutralized and become unreleased as a result of a weakening process, any plain obstruent that follows is tensified as a result of the strengthening process. Consequently, the three-way distinction among obstruents is reduced to two ([C^h] and [C']) in onset position following coda plosives. This process is illustrated in (2) below.

---

[7] Most of the examples used in this section were compiled and adapted from Sohn (1994; 1999) and Ahn (1998).

(2) a. /cap/+/**k**o/ → [cap.**k'**o] 'catches and'

b. /kip^h/+/**t**a/ → [kip.**t'**a] 'it is deep'

c. /k'ak'/+/**c**a/ → [k'ak.**c'**a] 'Let us shear it'

Furthermore, all fricatives (/s, s', s^h, h/) and affricates (/c, c^h/) neutralize when they occur in coda position, surfacing as [t]. For instance, all the underlying forms in (3) become homophonous when they occur in isolation.

(3)	a. /nac/	'daytime'	→	[nat]
	b. /nac^h/	'face'	→	[nat]
	c. /nas/	'sickle'	→	[nat]
	d. /nas'/	'be cured'	→	[nat]
	e. /nat^h/	'piece'	→	[nat]

They may optionally undergo further changes by receiving [place] features from the consonant that follows them. For instance, coronal obstruent phonemes (e.g., /t, t', t^h/) are highly vulnerable to segmental assimilatory processes such as assimilation in place of articulation. They typically become labial ([p]) or dorsal ([k]) before labial and dorsal consonants, respectively (e.g., /mit^h+pota/→[mipp'ota] ‘more than the bottom’; /mit+ko/→[mikk'o] ‘believe and’). Labial to dorsal assimilation, although variably, can also occur (/ip+ko/→[ikk'o] ‘wear and’). Naturally, when fricatives (/s, s', h/) and affricates (/c, c', c^h/) occur before labial and dorsal consonants, they can also undergo these optional changes. First, they neutralize to [t] syllable-finally and then receive the relevant [place] feature from the following consonant, as shown in (4).

(4) a. /ka**s**/+/kɯn/ → [ka**t**k'ɯn] or [ka**k**k'ɯn] 'hat string'

b. /k'o**c**^{h}/+/pat^{h}/ → [k'o**t**p'at] or [k'o**p**p'at] 'flower garden'

Alveolar fricatives (/s/ and /s'/) can only be realized as [s] in the coda as the first part of a geminate (5a). Furthermore, before one of /s/ or /s'/, the alveolar stop /t/ (and any obstruent that is neutralized to [t]) surfaces as [s] due to an assimilatory process of *sibilation* (5b and c).

(5) a. /mo**s**/+/s'ɯnta/ → [mo**s**.s'ɯnta] 'bad'

b. /ho**t**^{h}/+/sil/ → [ho**s**.s'il] 'single-ply thread'

c. /na**c**/+/sul/ → [na**t**.s'ul] → [na**s**.s'ul] 'liquor of daytime'

Second, processes such as *nasalization* and *lateralization* also affect surface forms of consonants in Korean. The process of nasalization turns all obstruents before nasal consonants into nasal consonants (6). That is, a sequence of /C+N/ must surface as [N+N].

(6) a. /cip/+/mun/ → [cimmun] 'house gate'

b. /hak/+/mun/ → [haŋmun] 'learning'

c. /k'oc^h/+/namu/ → [k'onnamu] 'flower tree'

Furthermore, in pre-nasal contexts, fricative and affricate sounds also undergo nasalization. This may yield surface forms identical to those of alveolar stops in coda positions. For instance, in (7) all the underlying forms become homophonous before a nasal consonant. First, the coda consonant is neutralized to [t]. Second, the neutralized consonant receives its [+nasal] feature in the pre-nasal context. Third, adjacent nasal stops must share their [place] features. Accordingly, the nasal stop becomes [labial] preceding a labial nasal.[8]

---

[8] Naturally, alternative analyses can be proposed here with respect to the way neutralization of fricatives and affricates to alveolar stops and other assimilatory processes are represented. For instance, in models that base their analyses on feature geometry, the de-linking of the node that contains the relevant manner features (e.g., [+continuant], [+delayed release], etc.) of these coronal consonants in syllable-final contexts can also produce a coronal consonant on the surface. In pre-nasal contexts, the same consonants also lose the relevant manner and place nodes as they are linked to the [+nasal] and [place] nodes of the following nasal. The discussion of what model of phonology or feature system would be more economical in terms of the representation of these assimilatory processes is, however, beyond the scope of this dissertation.

(7) a. /na**s**/+/man/ → /na**t**man/ → /na**n**man/ → [na**m**man] 'sickle only'

b. /na**t**/+/man/ →/pa**t**man/ → /na**n**man/ → [na**m**man] 'grain only'

c. /na**c**/+/man/ →/na**t**man/ → /na**n**man/ → [na**m**man] 'daytime only'

Last but of crucial importance to the present study, [ln] and [nl] contacts are impermissible in Korean and undergo a process usually referred to as *lateralization*. The lateralization process results in a geminate [l] on the surface (8).

(8) a. /tal/+/nara/ → [tallara] 'moon land'

b. /man/+/li/ → [malli] 'ten thousand miles'

Finally, voicing is predictable in Korean. Voicing takes place intersonorantly. However, it is only the plain stops and affricates that undergoes voicing (/s/ never voices). This is illustrated in (9).

(9) a. /han/+/kɯl/ → [haŋ.gɯl] 'the Korean alphabet'

b. /pa/+/po/ → [pa.bo] 'idiot'

It should be noted that nasalization, lateralization and voicing are very general processes of Korean *without* any exceptions. These rules not only robustly apply within words but also across word boundaries although not across IP boundaries.

These phonological processes also influence the adaptation of non-native forms borrowed from other languages into Korean, as I will show below. Patterns observed in loan word phonology may help us understand some of the dynamics of what Korean listeners might do with non-native strings. It must be noted, however, that the issues involved with loan words are incredibly complicated, particularly due to multiple possible sources.

In summary, Korean has coda restrictions as well as consonantal contact restrictions due to various phonological phenomena such as coda neutralization, nasalization and lateralization. In light of these processes, both types of restrictions will be employed to construct the stimuli for the present study.

3.4. Loan Word Adaptations in Korean

When loan words come into a language, not only the distribution of each sound, but also their segmental composition are adjusted to fit into the language's phonological patterns. Vowel epenthesis is an important strategy for accommodating illicit sound strings and unwanted coda segments in loan words in Korean. The epenthetic vowel is a high, back, unrounded vowel ([ɯ]) or, in the context of palatal consonants, the high front vowel [i] (10) (examples are from Sohn, 1999).

(10) a. [a.i.sɯ.kɯ.ɾim] 'ice cream'

b. [k^hɯ.ɾi.sɯ.ma.sɯ] 'Christmas'

c. [o.ɾen.ji] 'orange'

d. [sɯ.p^hən.ji] 'sponge'

Loan words like those listed above must be listed in Korean speaker's lexicon with the epenthetic vowels. For instance, the English word *bus* is pronounced roughly as [pasɯ] since /s/ cannot surface independently in the coda position. When this word is inflected with the nominative suffix, it triggers the allomorph that goes with vowel-final stems, -*ka* ([pasɯ-ka]), instead of the allomorph -*i* that suffixes to consonant-final stems (*[pas-i]). This indicates that these vowels are not epenthetic in production at all, but are part of the stored form of the borrowed form.

Based on the evidence from vowel epenthesis in the production of loan words in Korean as well as the empirical evidence for perceptual epenthesis from the previous studies on Japanese, vowel epenthesis should also arise in Korean listeners' perception of illicit consonant clusters. This assumption will shape the design of the perceptual experiment in the present study.

Since Korean employs a three-way laryngeal distinction among stop sounds and voicing is not phonemic, the adaptation of voiced and voiceless stops from English to Korean is complex but generally predictable. In prevocalic positions, voiced and voiceless stops are typically mapped to Korean plain (11) and aspirated (12) stops, respectively.

(11) a. data → [te.i.tʰa][9]

b. guitar → [ki.tʰa]

(12) a. tennis → [tʰe.ni.sɯ]

b. camera → [kʰa.me.ra] (Examples from Y. Kang (2001)).

Y. Kang (2001) points out unexpected cases where the substitution of English stops is not uniform. For instance, while in 'hip', the final consonant is realized as aspirated ([hi.pʰɯ]), the same consonant is plain in 'cap' ([kʰæp]). Also, in words such as 'cake', it seems that the final /k/ can either be aspirated (i.e., [kʰe.i.**kʰ**i]), or plain (i.e., [kʰe.i**k**]). That is, an epenthetic vowel is variably inserted after the post-vocalic word-final consonant, although this is not motivated by the syllable structure requirement.

Finally, there seems to be some variation, possibly sociolinguistically motivated, in the way Korean speakers accommodate English /km/ sequences, which

[9] The word 'data' can also be pronounced as [dæ.ɾa], reflecting more of a General American English pronunciation, where /t/ is realized as the flap [ɾ] intervocalically, following a stressed syllable.

can never surface in the language. For instance, H. Kang (1995) notes that, in loan words, an output with a vowel inserted after [k] surfaces in adult speakers (e.g., [p^{h}æk^{h}ɯmæn] 'Pac-man', [pik^{h}ɯmæk] 'Big Mac'). Younger generations, however, tend to spread the nasal feature to the pre-nasal consonant (e.g., [p^{h}æŋmæn]] and [piŋmæk], respectively). In the latter case, the modifications could be due to the way these words are spelled in the Hankul orthographical practice, which follows the morphophonemic spelling principle. A series of phonological rules are applied to Hankul spellings to obtain correct pronunciations. Therefore, nasal assimilation is unavoidable when, for instance, *Pac-man* is spelled as /p^{h}æk +mæn/ in Hankul, respecting the morphological integrity of the individual words forming these compounds in English. As these words become nativized in the language, they are pronounced just as any word with /k+m/ in Korean.

Having discussed the phonological processes that determine the surface forms of consonantal strings in Korean and how these processes may influence the production of loan words, a number of interesting questions arises as to whether the very same processes may also influence Korean listeners' perception. Specifically, do Korean listeners employ processes such as nasalization and lateralization in order to perceptually repair L2 strings that would undergo these processes in Korean production? Or, does speech perception has its own set of rules and restrictions that are not necessarily derived from speech production phenomena? Another interesting question pertains to how detailed Korean listeners' representation of L2 forms can be. Are allophonic features, such as voice, in L2 forms represented by Korean listeners? I will also investigate these questions in the present study along with the main research question in the dissertation, which asks whether it is coda restrictions or contact restrictions that cause perceptual epenthesis.

Table 3.4: The layout of the subsequent tables of the surface forms of heterosyllabic consonantal contacts in Korean**.**

		C_2	
		[STOP]	**[SIBILANT], [SONORANT], /h/**
C_1	**[STOP]**	see Table 3.5	see Table 3.6
	[SIBILANT], [SONORANT], /h/	see Table 3.7	see Table 3.8

Note: The tables reflect the ultimate surface form of each cluster with the assumption that all the optional place/manner of articulation assimilation processes have applied to the underlying forms. C_1 stands for a coda and C_2 for an onset consonant in a hypothetical heterosyllabic contact context ([VC_1.C_2V])

Table 3.5: Surface forms of [Stop] + [Stop] clusters in Korean (with place assimilation)

		C_2								
		p	**p'**	**p^h**	**t**	**t'**	**t^h**	**k**	**k'**	**k^h**
C_1	**p**	p.p'	p.p'	$p.p^h$	p.t'	p.t'	$p.t^h$	k.k'	k.k'	$k.k^h$
	p'	p.p'	p.p'	$p.p^h$	p.t'	p.t'	$p.t^h$	k.k'	k.k'	$k.k^h$
	p^h	p.p'	p.p'	$p.p^h$	p.t'	p.t'	$p.t^h$	k.k'	k.k'	$k.k^h$
	t	p.p'	p.p'	$p.p^h$	t.t'	t.t'	$t.t^h$	k.k'	k.k'	$k.k^h$
	t'	p.p'	p.p'	$p.p^h$	t.t'	t.t'	$t.t^h$	k.k'	k.k'	$k.k^h$
	t^h	p.p'	p.p'	$p.p^h$	t.t'	t.t'	$t.t^h$	k.k'	k.k'	$k.k^h$
	k	k.p	k.p'	$k.p^h$	k.t'	k.t'	$k.t^h$	k.k'	k.k'	$k.k^h$
	k'	k.p	k.p'	$k.p^h$	k.t'	k.t'	$k.t^h$	k.k'	k.k'	$k.k^h$
	k^h	k.p	k.p'	$k.p^h$	k.t'	k.t'	$k.t^h$	k.k'	k.k'	$k.k^h$

Note: There are no morphemes that underlyingly end in /p'/ or /t'/ in Korean

Table 3.6: Surface forms of [Stop] + ([Sibilant] or [Sonorant] or /h/) in Korean

		C_2									
C_1		**c**	**c'**	**c^h**	**s**	**s'**	**m**	**n**	**ŋ**	**l**	**h**
	p	p.c'	p.c'	p.c^h	p.s'	p.s'	m.m	m.n	__	m.n	p^h
	p'	p.c'	p.c'	p.c^h	p.s'	p.s'	m.m	m.n	__	m.n	p^h
	p^h	p.c'	p.c'	p.c^h	p.s'	p.s'	m.m	m.n	__	m.n	p^h
	t	t.c'	t.c'	t.c^h	s.s'	s.s'	m.m	n.n	__	n.n	t^h
	t'	t.c'	t.c'	t.c^h	s.s'	s.s'	m.m	n.n	__	n.n	t^h
	t^h	t.c'	t.c'	t.c^h	s.s'	s.s'	m.m	n.n	__	n.n	t^h
	k	k.c'	k.c'	k.c^h	k.s'	k.s'	ŋ.m	ŋ.n	__	ŋ.n	k^h
	k'	k.c'	k.c'	k.c^h	k.s'	k.s'	ŋ.m	ŋ.n	__	ŋ.n	k^h
	k^h	k.c'	k.c'	k.c^h	k.s'	k.s'	ŋ.m	ŋ.n	__	ŋ.n	k^h

Notes: (1) No morpheme underlyingly begins with /ŋ/.
(2) /k'/+/h/ may be unattested.

Table 3.7: Surface forms of ([Sibilant] or [Sonorant] or /h/) + [Stop] in Korean

		C_2								
		p	**p'**	**p^h**	**t**	**t'**	**t^h**	**k**	**k'**	**k^h**
C_1	**c**	p.p'	p.p'	$p.p^h$	t.t'	t.t'	$t.t^h$	k.k'	k.k'	$k.k^h$
	c'	p.p'	p.p'	$p.p^h$	t.t'	t.t'	$t.t^h$	k.k'	k.k'	$k.k^h$
	c^h	p.p'	p.p'	$p.p^h$	t.t'	t.t'	$t.t^h$	k.k'	k.k'	$k.k^h$
	s	p.p'	p.p'	$p.p^h$	t.t'	t.t'	$t.t^h$	k.k'	k.k'	$k.k^h$
	s'	p.p'	p.p'	$p.p^h$	t.t'	t.t'	$t.t^h$	k.k'	k.k'	$k.k^h$
	m	m.b	m.p'	$m.p^h$	m.d	m.t'	$m.t^h$	ŋ.g	ŋ.k'	$ŋ.k^h$
	n	m.b	m.p'	$m.p^h$	n.d	n.t'	$n.t^h$	ŋ.g	ŋ.k'	$ŋ.k^h$
	ŋ	ŋ.b	ŋ.p'	$ŋ.p^h$	ŋ.d	ŋ.t'	$ŋ.t^h$	ŋ.g	ŋ.k'	$ŋ.k^h$
	l	l.b	l.p'	$l.p^h$	l.d	l.t'	$l.t^h$	l.g	l.k'	$l.k^h$
	h	p^h	p'	p^h	t^h	t'	t^h	k^h	k'	k^h

Note: The so-called epenthetic-s (*Bindungs-s* or *sai-sios*), which is morphologically conditioned, obscures some sonorant combinations, especially those with /l/. When the epenthetic /s/ occurs following an /l/-final morpheme, it is deleted after it tensifies the following plain consonant (e.g., /mil/+/kalu/→/mil-**s**+ kalu/→[mil.k'a.ɾu] ‘wheat flour’. That is, the tensification of a plain C_2 after /l/ is due to the morphologically governed epenthetic /s/.

Table 3.8: Surface forms of ([Sibilant] or [Sonorant] or /h/) + ([Sibilant] or [Sonorant] or /h/) in Korean

		C_2									
		c	**c'**	**c^h**	**s**	**s'**	**m**	**n**	**ŋ**	**l**	**h**
C_1	**c**	t.c'	t.c'	t.c^h	s.s'	s.s'	m.m	n.n	__	n.n	t^h
	c'	t.c'	t.c'	t.c^h	s.s'	s.s'	m.m	n.n	__	n.n	t^h
	c^h	t.c'	t.c'	t.c^h	s.s'	s.s'	m.m	n.n	__	n.n	c^h
	s	t.c'	t.c'	t.c^h	s.s'	s.s'	m.m	n.n	__	n.n	s'/t^h
	s'	t.c'	t.c'	t.c^h	s.s'	s.s'	m.m	n.n	__	n.n	s'/t^h
	m	m.j	m.c'	m.c^h	m.n	m.n	m.m	m.n	__	m.n	m
	n	ɳ.j	ɳ.c'	ɳ.c^h	n.n	n.n	m.m	n.n	__	l.l	n
	ŋ	ŋ.j	ŋ.c'	ŋ.c'	ŋ.s'	ŋ.s'	ŋ.m	ŋ.n	__	ŋ.n	ŋ
	l	l.c'	l.c'	l.c^h	l.s'	l.s'	l.m	l.l	__	l.l	ɾ
	h	c^h	c'	c^h	s'	s'	m.m	n.n	__	?	t^h

Notes: (1) The combination /h+l/ is unattested.
(2) Sonorant+/h/ combinations are simplified via h-deletion.
(3) The combination /h+h/ first turns into /t+h/ and then merge into an aspirated stop[t^h].

Chapter 4

RESEARCH QUESTIONS AND HYPOTHESES

4.1. Research Objectives and Questions

The primary motivation of the present study is to find out what linguistic knowledge forces listeners to resort to perceptual epenthesis. Specifically, this study asks whether perceptual epenthesis arises due to coda violations or consonantal contact violations. L1 phonological systems may involve specific mechanisms to resolve both types of violations in production. For instance, in the case of Korean, coda violations caused by strident segments such as /c/ are resolved through neutralization by delinking the feature [strident] from the segment (/c/→[t.]). Contact violations such as *[k.m] are resolved by nasalization (i.e., /k.m/→[ŋm]). The question remains as to whether and how the same violations resolved in perception. The issue here leads us to a broader idea: that speech perception may have its own forces and restrictions and it is not necessarily governed by the rules that apply to speech production. One of the goals of the present study is to test the idea that speech perception may not be isomorphic to speech production. Furthermore, the perception of L2 forms also launches a secondary question that regards the detail of feature representations in listeners' perception of L2 forms. Specifically, do allophonic features that create illicit consonant clusters cause perceptual epenthesis?

In the present study, I explore these questions by employing English [$\ldots VC_{1.}C_2 V\ldots$] strings, where [$C_{1.}C_2$] forms a heterosyllabic consonantal contact relationship. The primary aim of the experiment is to tease apart the two factors discussed in Chapter 3 above that can potentially evoke perceptual epenthesis when such [$C_{1.}C_2$] sequences are illicit in the L1 of listeners. These factors are particularly related to the L1 knowledge of (1) [$C_{1.}C_2$] co-occurrence constraints (i.e., *[$C_{1.}C_2$]), and (2) coda violations that pertain to syllable structure constraints (i.e., *[C_1.]). These two types of constraints will constitute the focal point of the hypotheses that will be tested in this dissertation.

Specifically, as mentioned in Chapter 3, the present study varies C_1 and C_2 to give different sorts of consonant clusters. The clusters that violate consonantal contact restrictions (e.g., Obstruent + Nasal) either contain a permissible coda obstruent (e.g., /k/, /l/) or an impermissible one (e.g., /c/, /g/). The impermissible coda consonant has either a phonemic status (e.g., /c/, a [strident] consonant) or an allophonic status (e.g., /g/, a [voice] consonant) in Korean. This variable feature of the coda consonant is introduced into the experiment specifically to test the status of voicing in Korean listeners' representations of L2 forms. The crucial difference between the two features is that while allophonic features do not create a meaning difference, phonemic features are contrastive in Korean. The question is whether

Korean listeners suppress this allophonic feature in their percept of consonant clusters since it is redundant and predictable. If Korean listeners represent voicing in their percept then those clusters with [g] should cause perceptual epenthesis just as those with [c]. If they suppress this allophonic feature then the clusters with [g] should be comparable to those with [k] since the percept of [g] in the absence of voicing information must be /k/ for Korean listeners.

The experiment employs an AX discrimination paradigm, which comes from an adaptation of Dupoux et al.'s (1999) discrimination experiments. Dupoux and colleagues employed an ABX paradigm where their participants heard three stimuli in a row and have to decide whether the third stimulus is the same as the first or the second. In ABX experiments, it is not clear whether this intended decision strategy is always followed. That is, it is not necessarily the case that the only way to correctly respond is to consider A and B first and then compare them to X. Another strategy is possible, where subjects can simply compare the second stimulus (B) with the third (X) and decide if they are the same. If B and X are different then X must be the same as the first stimulus (A). In order to eliminate these different response strategies, the present study uses an AX paradigm, which requires subjects to compare only two stimuli in a row. The same stimuli are tested on two subject groups: native speakers of Korean and native speakers of English. The way it runs can roughly be described in the following way. Upon hearing doublets in the form of [VC_1.C_2.V] vs. [VC_1.V.C_2.V], English and Korean listeners are asked to respond whether they heard the same word or different words. Following Dupoux et al. (1999), the assumption behind this design is that if Korean listeners hear epenthetic vowels then they should perceive the two words to be the same. It should be noted that the experiment is not designed to directly test if listeners hear epenthetic vowels. Rather, it builds on the assumption that if they hear an epenthetic vowel then they should perceive the given words in the doublet to be the same[10]. More information about the design of the experiment is provided in Chapter 5.

4.2. Hypotheses

Present study tests two conflicting main hypotheses, namely the Consonantal Contact Hypothesis and the Coda/Onset Identity Hypothesis, and two conflicting sub-hypotheses, namely the Phonetic Processing Hypothesis and Phonological Processing Hypothesis.

4.2.1. The Consonantal Contact Hypothesis

The Consonantal Contact (CC) hypothesis asserts that listeners' precise knowledge that certain consonantal contact situations are not allowed in the language

---

[10] It could also be the case that, instead of epenthesis, one or both of the segments in the illicit cluster can be misperceived by Korean listeners. For instance, hypothetically, Korean subjects might perceive [p^hacma] as /p^hatma/ or /p^hamma/, etc. All of these misperceptions, however, should evoke a "different" response because the comparison in the test doublet containing this item still remains to be between a word with a consonant cluster and a word without it (i.e., /p^hatma/ vs. /p^hacima/ or /p^hamma/ vs. /p^hacima/).

(i.e., *[C_xC_y]) will force Korean listeners to reanalyze the consonantal string as having an epenthetic vowel between the two consonants. Regardless of what the unit of perception is assumed to be (be it feature bundles or segments), once two consonantal units are detected in the speech string that violate a contact restriction, *[C_xC_y] will cause a percept of a missing vowel. Since the hypothesis makes unit-independent predictions, it is compatible with any model that assumes the unit of perception to be segment-like units or features.

The CC hypothesis is conceptualized through a string-based approach towards explaining sound distributions. In string-based phonotactics, the structure of speech input in these views consists of linear strings of discrete abstract linguistic units (feature bundles or segments) that are ultimately bound by word or morpheme boundaries (cf. Steriade, 1999; Blevins, 2002). Steriade (1999) classifies the positions of segments not in syllabic terms (e.g., as onset vs. coda) or in linear terms (e.g., "before a vowel" vs. "after a consonant") but as positions where certain featural contrasts are more vs. less perceptible. Accordingly, certain consonant clusters are illegal because the featural contrast on one or both of the segments in the cluster is difficult to perceive in that context. Accordingly, a string-based approach, *a là* Steriade, motivates perceptual epenthetic vowel insertion as a way to create a position where the featural contrasts in question become more perceptible. In simple linear terms, the percept of an epenthetic vowel ensures a buffer sound that breaks up the unwanted consonantal contact.

In summary, the CC Hypothesis motivates perceptual epenthesis to avoid an impermissible consonantal contact situation without the necessity to refer to any sort of phonological constituency such as the syllable.

4.2.2. The Coda/Onset Identity Hypothesis

The Coda/Onset Identity (COI) hypothesis, on the other hand, is based on the view that phonological structure is hierarchical. In particular, the hypothesis crucially refers to syllable structure conditions in the L1 system. These conditions specifically state that certain consonants can never surface as coda or as onset. That is, a given consonant may crucially lack a particular positional identity (e.g., coda identity or onset identity) in the inventory. Therefore, following a vowel, when the parser encounters a consonantal element that does not carry a coda identity (e.g., /c/ in Korean only occurs in the onset position, and therefore it does not carry a coda identity), the parser will automatically interpret the consonant as an onset. Accordingly, in a [...$VC_{1.}C_2V$...] string where two consonants (i.e., $C_{1.}C_2$) are in contact, perceptual epenthesis will only arise if the C_1 slot contains a consonant that lacks a coda identity in the L1. Consequently, the COI hypothesis predicts that Korean listeners will analyze those consonants as an onset. Because onsets must be non-branching in Korean, a vocalic segment must follow to form the nucleus.

Just like the earlier hypothesis, the implications of the COI hypothesis are also unit-independent. That is, the unit of perception can be a segment or a bundle of features. Crucially different from the CC hypothesis, however, the COI hypothesis motivates epenthetic vowels out of a necessity to build well-formed syllables. The percept of a missing vowel will emerge as the material interpreted as an onset in the perceptual representation requires a vocalic element to satisfy the syllable structure conditions of the language: onsets must be followed by a vocalic element[11].

In short, while the CC hypothesis motivates the perceptual epenthetic vowels in order to avoid unwanted consonantal contact, the COI hypothesis motivates the percept of these vowels out of a necessity to create well-formed syllables.

4.2.3. Perceptual Level Representations and Sub-hypotheses

The perception of L2 forms also launches a secondary question, perhaps as intriguing as the first one from a speech processing perspective. How detailed are L2 input representations? The present study proposes two potential answers to this question, which will form the basis for the two sub-hypotheses that are tested in the experiment. First, L2 phonological processing could aim to use all phonetic features including allophonic and redundant information in the acoustic stream. Second, L2 phonological processing could be phonological, which suppresses the predictable and redundant information, thereby storing only contrastive information.

---

[11] The content of the vocalic element that will emerge in such contexts can be specified by a default rule in both perception and production. Notice that such a rule would not only be necessary for the epenthesis predicted by the COI hypothesis but also by the CC hypothesis. Therefore, the specifics of such a rule are not crucial for the purposes of this study.

An example for such predictable phonological information in Korean is voicing. Voicing in Korean is non-contrastive and results from a systematic process without exceptions. Plain obstruents, which are voiceless, surface as voiced between sonorant sounds. The question is whether the perceptual system will build input representations by coding all phonetic information including the allophonic voicing properties of sounds, or whether it will be phonological in nature thereby suppressing voicing information since it is redundant and not contrastive in Korean. Thus, the test clusters will also contain voiced allophones of the coda consonants (e.g., [g], [j] as allophones of /k/ and /c/, respectively). Just like with [k] and [c], these consonants will be followed by a stop (i.e., [g.t], [c.t]) and by a nasal (i.e., [gm] vs. [jm]).

4.2.3.1. Phonetic Processing Hypothesis

According to the Phonetic Processing hypothesis, L1 segmental representations reflect all possible *surface* forms of discrete phonological segments. That is, the mapping of extracted acoustic features from the speech stream to L1 segmental categories is by no means constrained by phonological processes that systematically derive these features on the surface in the language. The assumption in the hypothesis is that perceptual representations are fully equipped with all available surface phonetic information.

The view advocated by this hypothesis is comparable to Silverman (1992). Silverman hypothesizes that speakers parse non-native speech signals in such a way that they match the signal to native feature matrices that most closely approximate the phonetic properties of the signal. Similarly, the Phonetic Processing hypothesis developed here asserts that the material detected in the speech stream will be compared to the feature matrices of phonetic forms in the L1. Specifically, if the acoustic stream contains [voice], which closely approximates to the phonetic properties of the voiced segments in Korean, [voice] will be represented at the perceptual level although voicing is predictable in Korean. Voiced segments such as [g] then will be faithfully represented by Korean listeners; in a consonantal contact situation such as [g.t] or [g.m], the clusters will then violate both the Consonantal Contact (CC) and Coda/Onset Identity (COI) hypotheses. This is because (1) these consonantal combinations never surface in Korean violating contact restrictions, and (2) [g] can never surface in the coda position, violating positional constraints (see the following section for another possible explanation). Additionally, these combinations violate the distribution of voiced segments according to the intersonorant voicing rule. The obvious way to correct this is to hypothesize an unheard vowel after [g].

There are speech perception models, particularly in the spoken word recognition literature (e.g., Christie, 1974; Church, 1987), which employ such a phonetic approach towards speech processing. These models typically pack as much information into the representation of lexical items as possible, storing possible variations of a word. However, the similarity between the phonetic-processing hypothesis developed here and the assumptions of such models are only relevant to the richness of phonetic information that is accessed in the acoustic stream. Since only nonce test items will be employed, the focus of the present study will be on the perception of forms that do not have any lexical status in the L1 or L2. That is, the present study is not concerned with making any predictions or proposals regarding the way lexical items are stored.

4.2.3.2. Phonological Processing Hypothesis

The phonological processing hypothesis claims that speech perception functions in an abstract and minimal manner. Accordingly, listeners filter the acoustic stream by detecting only a minimal amount of articulatory and/or acoustic cues. These cues correspond to the concept of distinctive features that are used to set up oppositions in underlying representations of the segments in the inventory of each language[12]. The perceptual mechanism consistently detects only certain subsets of these features and not others, suppressing redundant and predictable information coded in the surface representations. Korean voicing, under this hypothesis, is not represented by Korean listeners in their phonological perception since it is predictable.

---

[12] The Phonological Processing hypothesis is compatible with Underspecification Theory (cf. Archangeli, 1984; 1988) as well as the fundamentals of the Featurally Underspecified Lexicon Model (FUL) (Lahiri and Jongman, 1990; Lahiri and Marslen-Wilson, 1991, 1992, Lahiri and Reetz, 1999, see also Fitzpatrick and Wheeldon, 2000 for a summary of the FUL model). It should be remembered that models such as the FUL make specific predictions about what information can be used to access the lexicon and presumably to construct new lexical entries. The similarity between the Phonological Processing hypothesis and the assumptions of the FUL model rests only on the abstract nature of underlying phonological representations, where non-distinctive information is not represented.

When [voice] is not represented, a segment such as [g] must be represented as the less specified plain velar stop, /k/.

Unlike the Phonetic Processing hypothesis, the Phonological Processing hypothesis makes different predictions under the two main hypotheses raised earlier. If the CC hypothesis holds, then only the combination of [g.t], which is perceived as [k.t] with the suppression of voicing, should be possible because [k.t] is a permissible combination in Korean. The cluster [g.m], which is perceived to be [k.m], on the other hand, should be impossible since it is an impermissible combination. In contrast, if the COI hypothesis holds up, both [g.t] and [g.m] should uniformly be possible since [k.], the phonologically representation of the surface segment [g], is a permissible coda consonant in Korean.

Successful discrimination of the [g.t] and [g.m] clusters by Korean listeners can alternatively be related to the fact that the non-codahood of voiced consonants is epiphenomenal in the language[13]. That is, there are no phonological rules or constraints that ban these consonants altogether from the coda position in Korean, meaning that the distribution of these phones happens to never overlap with the coda position. To reach a better understanding of this alternative explanation, voiced segments such as [g] should be compared to other consonants such as [c] that also never surface in coda position in Korean. While [g] is an illicit coda as an indirect

---

[13] I am thankful to Benjamin Bruening for bringing this point to my attention.

consequence of the statement of its allophonic distribution, as given in (1), [c] in the coda position violates a syllable structure condition as stated in (2).

(1) Plain→[voice] / [sonorant] __ [sonorant]

(2) *CODA
|
[strident]

As will be argued below, despite the absence of an explicit rule that specifically bans the feature [voice] from coda position, there is also no evidence for Korean speakers to know the lack of such a restriction in Korean. Although the rule in (1) does not make any reference to syllable structure, there are several reasons that a plain obstruent cannot surface as voiced in coda position. First, when a vowel follows a plain obstruent, that obstruent will surface in the onset rather than the coda position. In this case, voiced obstruents appear in onset position, rather than in coda position, as a result of onset maximization. Second, although Korean voicing is typically described to occur intersonorantly, and thus it may involve segments other than vowels and glides, voicing of plain ostruents through a sonorant consonant is possible only when the sonorant precedes the plain obstruent (e.g., /kalpi/→[kal.bi] 'spare ribs'; /han+kɯl/→[han.gɯl] 'the Korean alphabet'). A plain obstruent never undergoes voicing in coda position before a sonorant consonant. When the following sonorant is a glide, just like with vowels, the obstruent is syllabified as the onset of the following syllable (e.g., /cik+wən/→[ci.gwən] 'staff'). The appropriate context for voicing before liquids cannot be constructed due to some unusual historical restrictions on morpheme shapes in Korean. Therefore, a [sonorant] consonant, excluding glides, following a plain obstruent could only be a nasal. In the context of a nasal, however, an obstruent undergoes nasalization (e.g, /cip+mun/ → [cim.mun] 'house gate'). In short, the reason that voiced consonants only surface in the onset position is not that they are not allowed in the coda position. It is rather an indirect effect of the phonological context that makes the allophonic rule in (1) possible. This is different from strident segments: the language strictly bans them from coda position on the surface.

The alternative explanation introduced above would then attribute the successful performance on the voiced clusters [gt] and [gm] to the lack of a rule that specifically bans voiced segments from coda position in Korean. This explanation can indeed account for successful performance on [gt] and [gm] in a straightforward manner since it excludes phonetic and phonological processing from consideration. It simply relies on the COI hypothesis, according to which only strident segments are excluded from the coda position. Voiced velar segments, thus, have the same status as voiceless velar ones in that they are not banned from this position via any explicit rule of Korean. By doing so, in the context of [gt], this explanation actually ignores the intersonorant voicing rule, according to which voicing between V_ [t] cannot take place.

However, just like there is lack of evidence for a rule that bans voiced consonants from coda position, there is also no evidence in the language that would enable Korean speakers to know the *absence* of such a rule. In fact, the latter case creates a problem for the alternative approach under current phonological theories such as Optimality Theory (OT) (Prince and Smolensky, 1993). OT assumes the existence of a universal set of constraints, which are ranked according to how strongly they hold in a particular grammar. Assuming the Biased Constraint Demotion (BCD) algorithm (Prince and Tesar, 1999; Tesar, 2002), in the absence of evidence for voicing in coda position, a markedness constraint banning voiced segments from coda position (i.e., *CODA[voice]) is undominated in Korean. Due to the high ranking of this constraint, candidate structural descriptions containing a voiced segment in coda position are always rejected as optimal candidates. Accordingly, the fact that Korean speakers lack evidence for voiced segments surfacing in coda position does not necessarily imply the non-existence of a constraint banning such a phenomenon.

Conceptually, the alternative explanation overlooks the difference between the status of voiced and voiceless segments in Korean phonology. The Phonological Processing hypothesis, however, derives the same finding based on the very fact that voicing does not carry linguistically meaningful information for Korean listeners; therefore, it may be ignored. That is, the Phonological Processing hypothesis is conceptually reflective of an important generalization about the allophonic status of voicing in Korean. An explanation of the same facts based on a causal relationship between the knowledge they possess about the status of voicing in their L1 and the detail they encode with respect to voicing in their L2 representations is conceptually more advantageous. Therefore, I will entertain the different predictions of the Phonetic versus the Phonological Hypothesis and evaluate them in this dissertation.

In summary, the two main hypotheses postulated above deal with the specific phonotactic knowledge in the L1 of the listener, which may independently exert forces on the perception of illicit consonant clusters in the L2 and their perceptual adaptation to the L1 system. In previous studies on the same issue, these two forces were not separated. Furthermore, where we draw the fine line between phonetic and phonological processing may make different predictions during the course of speech perception, where there seems to be an active interference of L1 phonology. Therefore, it is important to understand what dynamics are involved in speech perception and its interface with what the L1 phonological system has to offer towards building up perceptual representations of L2 input forms in the mind. I introduce the empirical test developed for this enterprise in the following chapter.

Chapter 5

EMPIRICAL TEST: METHOD AND PREDICTIONS

5.1. Method

This section presents the methodology employed in the present study. It lays out the criteria with which the test materials are constructed, and presents the design of the perceptual experiment.

5.1.1. Choice of Consonantal Contact Contexts

The primary aim of the experiment conducted in this dissertation was to test for perceptual epenthesis for Korean listeners in English coda-onset clusters of different types as a function of two distinct sets of information: coda restrictions and consonantal contact restrictions in Korean. The templatic form of the English nonce words was [p^háC_1(V)C_2a], where the form varied with regard to the consonant cluster formed by C_1 ([c], [j], [k], [g], [l], [ŋ], [n]) and C_2 ([t^h], [m], [n]), and by V ([i], [ʊ]) that was either present or not between C_1 and C_2. Since the forms were intended to be possible English words, the sounds forming the words had to be chosen from the English sound inventory. Furthermore, to test the effect of coda identity, C_1 consonants were selected in such a way that they varied between a permissible coda consonant and an impermissible one in Korean. Moreover, in order to examine whether co-occurrence restrictions played a role in perceiving epenthetic vowels, it was necessary that the very same coda consonants yielded a permissible cluster and an impermissible one depending on the following consonantal context.

The Korean consonantal inventory provided a small number of consonants that met this stringent set of criteria (see Section 3.3 in Chapter 3 for more information). Specifically, the experimental materials made use of [k], [l], and [tʃ] for the coda (C_1) position. It should be remembered that, just like in Korean, English obstruents /k/ and the palato-alveolar affricate /tʃ/ are typically unaspirated in the coda position. Therefore, these English coda obstruents typically correspond phonetically to Korean plain [k] and [c], respectively. For ease of presentation, I will adhere to the symbol /c/ for both affricates in English and Korean.

The reason that [k] and [l] were specifically chosen as licit coda consonants (C_1) was that these consonants are least subject to assimilation (either optional or compulsory) in Korean when they precede other consonants. In fact, a closer analysis of Table 3.4 (and the subsequent ones) in Chapter 3 reveals that when these two consonants are the initial member of a cluster (i.e. when they occur as C_1), they remain unaltered except before certain nasal contexts. Using both [k] and [l] as a licit coda consonants ensures that any perceptual illusion effect is not due to the particular phonetic characteristics of an individual sound and to yield a broader generalization of the results to obstruents and sonorants. Furthermore, while [k] is the consonant that gets modified when combined with a following nasal consonant, it is [l] that affects the following nasal. Therefore, having [l] in the experiment in addition to [k] allows us to investigate a variety of consonantal contact violation phenomena: nasalization in the case of *[k.m] and lateralization in the case of *[l.n].

Illicit C_1 was one of [c], [g] or [j]. Regardless of whether these consonants precede an oral stop ([t^h]) or a nasal ([m]), they yield impossible phonetic clusters in Korean because they never occur in the coda position. Unlike [c], the voiced consonants [g] and [j] have an allophonic and not phonemic status in Korean. As mentioned before, voiced obstruents only surface intersonorantly as allophones of voiceless plain obstruents. The reason that illicit C_1 varied between a phoneme and an allophone was to measure different types of processing procedures that might be used by the listeners. Specifically, if voicing is represented, which would be predicted by the Phonetic Processing hypothesis, then both [g] and [j] should pattern with [c]. If, on the other hand, voicing is suppressed, [g] should exhibit a similar profile with [k], a licit coda consonant, and [j] with [c].

The choice of the onset consonant (C_2) was carried out in the following way. C_2 would have to yield either a possible or an impossible consonant cluster when combined with [k] or [l]. Therefore, C_2 was chosen to vary between an oral stop[14]

---

[14] Most prevocalic voiceless obstruents (e.g., /t/) in English are aspirated and thus phonetically correspond to aspirated consonants in Korean (i.e., [t^h]). This correspondence can be seen in loan words, in which English voiceless stops in prevocalic positions (e.g., *c̲amera, guit̲ar*) are mapped to Korean aspirated stops (e.g., [$\underline{k}^h$a.me.ra], [ki.$\underline{t}^h$a] (Kang, 2001).

(i.e., [t^h]) and a nasal (either [m] or [n]). Accordingly, both [k.t^h] and [l.t^h] yield permissible contacts in Korean while [k.m] and [l.n] do not.

Table 5.1 summarizes the sound combinations used in the experiment and their phonotactic status in Korean phonology.

Table 5.1: Consonantal contexts used in the experiment and their legitimacy in Korean.

		C_2	
		Oral Stop (i.e., [t^h])	***Nasal (i.e., [n]or[m])***
C_1	**[k]**	Licit	Illicit
	[g]	Illicit	Illicit
	[c]	Illicit	Illicit
	[j]	Illicit	Illicit
	[l]	Licit	Illicit

The loan word evidence presented in Chapter 3 suggests that words with [km] (e.g., *Pac-man, Big Mac*) are produced in two different ways: (1) by the insertion of the epenthetic vowel (i.e., as [kɯm]), or (2) by the nasalization of the preceding obstruent (i.e., as [ŋm]). Judgments from native speakers that I have consulted suggest that, there is also variation, even within the same speaker, in the pronunciation of English words with [ln] clusters such as *hazelnut*, or *Telnet*. This cluster is pronounced either with the insertion of an epenthetic vowel (i.e., as [lɯn]) or by lateralizing the nasal consonant in the cluster (i.e., as [ll]). Therefore, extra conditions which involved sequences such as [ŋ.m], [l.l] and [n.n] were created for comparison as likely output forms of *[k.m] and *[l.n] in Korean. Perhaps the most interesting question that these extra conditions will answer is whether Korean production rules, which resolve illicit sequences such as [km] and [ln], also apply to Korean listeners' perception of the very same sequences.

5.1.2 Choice of the Epenthetic Vowels

The experiment employed a same-different AX discrimination paradigm, where subjects hear two stimuli in a row and decide whether they are the same or different. The experiment compared the perception of nonce forms that contained consonant clusters (e.g., [p^hac.ma]) with those that did not due to the presence of a vowel (e.g., [pa.c^hi.ma]). The participants were asked to respond, upon hearing two stimuli in a row, whether the two words they heard were the same or different. The assumption here was that if the perceptual system inserted an epenthetic vowel to break up the unwanted consonant clusters, Korean participants would have trouble distinguishing the group of words that only differed with regard to the presence of the vowel. Therefore, the items that contained clusters (e.g., [p^hac.ma]) were contrasted with the other exemplars of the same cluster type but contained the vowel [i] (i.e., [p^ha.ci.ma]). The epenthetic vowel was chosen to be [i] in this context where the preceding consonant is a palatal. In other contexts, the English [ʊ], which is the closest approximation of the Korean [ɯ], was used. The vowel [ʊ] is a high, back, lax vowel, with a very subtle degree of rounding, and is typically used in words such as *put*, *foot*, *butcher* and *could*. Contrary to the basic assumptions regarding the articulatory properties of rounding, the lips are hardly ever substantially protruded in the pronunciation of these words. Therefore, the English [ʊ] sounds acoustically extremely close to the Korean [ɯ]. The fact that many words in languages such as Korean and Japanese that contain [ɯ] are typically represented with the letter <u> in the Roman alphabet can also be attributed to the fact that the closest approximation of this sound in English is [ʊ], which is generally spelled as <u> in English.

5.1.3. Materials

Nonce words that conform to the model [páC_1(V)C_2a] were constructed. C_1 was one of [k, g, c, j, l, n, ŋ] while C_2 was one of [t^h, m, n, l]. The vowel (V) was either [i] or [ʊ]. The combination of these variables was permuted to create forms that conformed to (1) [paC_1C_2a] ("cluster"-words), and (2) [paC_1VC_2a] ("epenthesis"-words). The word stress was placed on the first syllable in these forms in order to keep the prosody of the experimental items as constant as possible. None of the test words corresponded to actual words in Korean or English. The complete list of test words is given in Table 5.2.

Table 5.2: Test words used in the experiment

Cluster	***Cluster-words***	***Epenthesis-words***	***Example English word with the cluster***
[kt^h]	[p^{h}ákt^{h}a]	[p^{h}ák^{h}ʊt^{h}a]	*October, factory, doctor*
[km]	[p^{h}ákma]	[p^{h}ák^{h}ʊma], [p^{h}ák^{h}ima]	*acme*
[gt^h]	[p^{h}ágt^{h}a]	[p^{h}ágʊt^{h}a]	*pigtail, ragtime*
[gm]	[p^{h}ágma]	[p^{h}ágʊma]	*dogma, segment*
[ct^h]	[p^{h}áct^{h}a]	[p^{h}ác^{h}it^{h}a]	*pitch-tracker*
[cm]	[p^{h}ácma]	[p^{h}ác^{h}ima]	*Richmond, attachment*
[jt^h]	[p^{h}ájt^{h}a]	[p^{h}ájit^{h}a]	*vegetable (dialectal)*
[jm]	[p^{h}ájma]	[p^{h}ájima]	*arrangement*
[lt^h]	[p^{h}ált^{h}a]	[p^{h}álʊt^{h}a]	*saltines, shelter*
[ln]	[p^{h}álna]	[p^{h}álʊna]	*walnut, vulnerable*
[ŋm]	[p^{h}áŋma]	-	*Ingmar, kingmaker*
[nn]	[p^{h}ánna]	-	*pinenut*[15]
[ll]	[p^{h}álla]	-	*mail-list*

The test words were then placed in doublets for a relevant comparison. It should be noted that, since the analysis of the results was carried out within the framework of the Signal Detection Theory (Green and Swets, 1974), it was also necessary to examine listeners' behavior with doublets that compared different exemplars of same words (e.g., [p^{h}akma] vs. [p^{h}akma]; [p^{h}ak^{h}ʊta] vs. [p^{h}ak^{h}ʊta]) to see if there were any response biases (see Section 5.5 for more information). The word [p^{h}ak^{h}ima] was

[15] Variation in geminate reduction can be observed in the pronunciation of these words in English.

created to get a baseline measurement of listeners' behavior towards pairs that contained different vowels[16]. Thirty-nine experimental doublets were constructed based on the following templates:

(1) *Experimental Doublets*

i. "Different"-doublets:

(a) Pairs of words that contrasted "cluster"-words with "epenthesis"-words:

[p^{h}ákt^{h}a]	*vs.*	[p^{h}ák^{h}ʊt^{h}a]
[p^{h}ákma]	*vs.*	[p^{h}ák^{h}ʊma]
[p^{h}ákma]	*vs.*	[p^{h}ák^{h}ima]
[p^{h}ágt^{h}a]	*vs.*	[p^{h}ágʊt^{h}a]
[p^{h}ágma]	*vs.*	[p^{h}ágʊma]
[p^{h}áct^{h}a]	*vs.*	[p^{h}ác^{h}it^{h}a]
[p^{h}ácma]	*vs.*	[p^{h}ác^{h}ima]
[p^{h}ájt^{h}a]	*vs.*	[p^{h}ájit^{h}a]
[p^{h}ájma]	*vs.*	[p^{h}ájima]
[p^{h}ált^{h}a]	*vs.*	[p^{h}álʊt^{h}a]
[p^{h}álna]	*vs.*	[p^{h}álʊna]

[16] The epenthetic vowel is [ɯ] in the environment of velars. Therefore, the vowel [i] was used in the baseline pair with the assumption that if Korean listeners perceive *[km] as [kɯm] and not as with any other vowel, they should hear the difference between [i] and [ɯ]. However, the same reasoning could not be applied to [c]-clusters because my pilot experiments with Korean as well as Turkish subjects show that in the context where C_1 is a palatal, the vowel [ɯ] is often confused with [i]. This is perhaps not surpising given that palatals usually attract other segments in their environments. Since this issue is peripheral to the main arguments of this thesis, I will not speculate further on it.

(b) Pairs of words that contrasted cluster-words with their likely output forms from the production grammar:

[p^{h}ákma]	*vs.*	[p^{h}áŋma]
[p^{h}álna]	*vs.*	[p^{h}ánna]
[p^{h}álna]	*vs.*	[p^{h}álla]
[p^{h}ágma]	*vs.*	[p^{h}áŋma]

ii. "Same"-doublets:

(a) Pairs of words that contrasted the different exemplars of the same cluster-words:

[p^{h}ákt^{h}a]	*vs.*	[p^{h}ákt^{h}a]
[p^{h}ákma]	*vs.*	[p^{h}ákma]
[p^{h}ágt^{h}a]	*vs.*	[p^{h}ágt^{h}a]
[p^{h}ágma]	*vs.*	[p^{h}ágma]
[p^{h}áct^{h}a]	*vs.*	[p^{h}áct^{h}a]
[p^{h}ácma]	*vs.*	[p^{h}ácma]
[p^{h}ájt^{h}a]	*vs.*	[p^{h}ájt^{h}a]
[p^{h}ájma]	*vs.*	[p^{h}ájma]
[p^{h}ált^{h}a]	*vs.*	[p^{h}ált^{h}a]
[p^{h}álna]	*vs.*	[p^{h}álna]
[p^{h}áŋma]	*vs.*	[p^{h}áŋma]
[p^{h}ánna]	*vs.*	[p^{h}ánna]
[p^{h}álla]	*vs.*	[p^{h}álla]

(b) Pairs that compared the different exemplars of the same epenthesis-words:

[p^{h}ák^{h}ʊt^{h}a]	*vs.*	[p^{h}ák^{h}ʊt^{h}a]
[p^{h}ák^{h}ʊma]	*vs.*	[p^{h}ák^{h}ʊma]
[p^{h}ák^{h}ima]	*vs.*	[p^{h}ák^{h}ima]
[p^{h}ágʊt^{h}a]	*vs.*	[p^{h}ágʊt^{h}a]
[p^{h}ágʊma]	*vs.*	[p^{h}ágʊma]
[p^{h}ác^{h}it^{h}a]	*vs.*	[p^{h}ác^{h}it^{h}a]
[p^{h}ác^{h}ima]	*vs.*	[p^{h}ác^{h}ima]
[p^{h}ájit^{h}a]	*vs.*	[p^{h}ájit^{h}a]
[p^{h}ájima]	*vs.*	[p^{h}ájima]
[p^{h}álʊt^{h}a]	*vs.*	[p^{h}álʊt^{h}a]
[p^{h}álʊna]	*vs.*	[p^{h}álʊna]

Due to the nature of the experiment, Korean listeners were potentially at a disadvantage since most of the stimulus pairs were expected to yield a "same" response from Korean listeners. Thus, special attention was given to filler doublets in order to increase the number of "different"-doublets in the experiment so as to balance the responses better for both groups. Filler doublets were created in such a way that the different words contained in them should pose no potential difficulty for Korean listeners to discriminate on any grounds (e.g., [p^{h}ákt^{h}a] *vs.* [p^{h}ákma]; [p^{h}ált^{h}a] *vs.* [p^{h}á**nn**a], etc.). Furthermore, to control for any potential effect of order of presentation in doublets, each "different"-doublet presented the words in both possible orders (e.g., [p^{h}ákt^{h}a] vs. [p^{h}ák^{h}ʊt^{h}a]; [p^{h}ák^{h}ʊt^{h}a] vs. [p^{h}ákt^{h}a]). Different exemplars of the same words were randomly chosen from a pool of exemplars that consisted of ten recordings of each test word. A different exemplar of the same item was used for each different order (e.g., [pʰáktʰa]$_1$ vs. [pʰákʰʊtʰa]$_2$; [pʰákʰʊtʰa]$_3$ vs. [pʰáktʰa]$_4$).

In addition, another set of cluster- and epenthesis-words was created for a practice session (e.g. [pʰábma], [pʰálma], [pʰáltʰa], [pʰábʊma], [pʰábʊtʰa], etc.). The doublets developed for the practice session had 5 "same" and 5 "different" pairs and they were expected to be successfully discriminated. None of the practice items were used in the actual test.

Each nonce item was produced ten times by a male native speaker of American English, who is a trained phonetician. Using a high impedance microphone (Audio-Technica, Omnidirectional, ATR35S), these words were recorded on a Compact Disc (CD) using a CD recorder (TEAC RW-800) in a sound attenuated room. The recorded tracks on the CD were then transferred onto a Mac OS 9.1 computer, where they were converted to 22kHz (16-bit resolution; 1-channel) and stored as ".aiff" files. Using Praat 4.0.1. (Boersma & Weenink), each token word was cropped out and saved into a separate file. The best exemplars of each word were chosen based on how good and natural they sounded. For instance, since the words were read 10 times from a list, special attention was paid to select the ones that did not carry the intonational properties of list reading. The selected words were then transferred into Sound-Designer-II files, keeping the same sampling parameters, to be used in PsyScope 1.2.5 PPC (Cohen, MacWhinney, Flatt and Provost, 1993).

5.1.4. Design

An AX discrimination protocol was constructed using the PsyScope software. Following Werker and Logan (1985), all pairs had an inter-stimuli interval (ISI) of 1500 milliseconds (ms) to bias listeners towards forming phonological representations in their comparisons. It must be noted that biasing listeners to form phonological representations does not confound the distinction that was made between phonological processing and phonetic processing. The use of the term "phonological" is perhaps misleading in this context; the purpose was to avoid any sort of comparison on purely acoustic grounds. Such behavior would otherwise cause all "same"-doublets to sound different to the listeners since the doublets contained different exemplars of the same item. By using a long ISI, the aim was to make this a purely linguistic, rather than an auditory, task.

At the mid-point of the ISI, a 500 ms long tone was played to further ensure that the word just heard would be encoded in a linguistic representation. That is, immediately after the presentation of the first word, there was 500 ms of silence, followed by a tone of 500 ms length, plus another 500 ms of silence. Then followed the second word in the doublet. About 150 ms after the onset of the second word, the computer allowed the participants to respond by pressing either the key <A> for "same", or <L> for different on the keyboard in front of them. The 150-ms-time lag was placed specifically to prevent subjects from responding abruptly without hearing the second word. A beep after each response confirmed to the subject that the response had been entered. Each trial ended either immediately after a response was given, or after 6000 ms had elapsed if the subject gave no response. The next trial started 2000 ms later. Using this protocol, an experimental block with was created with 2 random repetitions of each of the experimental and filler doublets.[17] The total number of trials was 118 in the block (39 test doublets presented in both possible orders plus 20 filler doublets also presented in both possible orders). The block was presented 5 times in a row, yielding 10 total repetitions of each doublet.

Figure 5.1 illustrates the components of the AX discrimination protocol.

[17] The experiment also contained 10 doublets that were designed for a pilot study, which will not be reported in this dissertation. The stimuli in these pairs were very similar in form to the actual test words; however, they contained digitally altered variations of some of the sounds in them. These items will be considered extra filler items for the purposes of the present study.

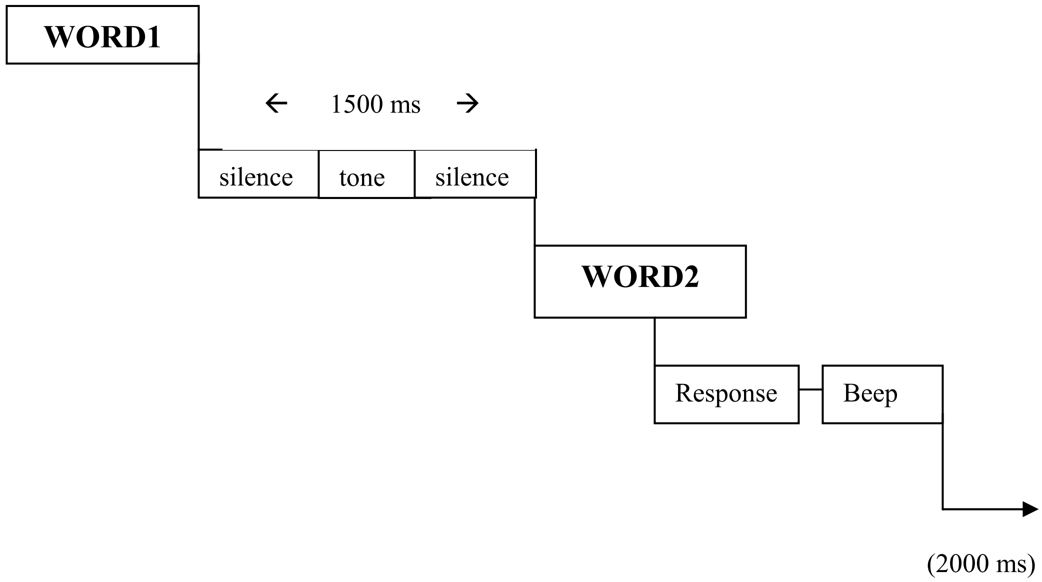

Figure 5.1: AX discrimination protocol per trial

5.2. Experimental Predictions

The two competing hypotheses tested in this dissertation, namely the Consonantal Contact and Coda/Onset Identity hypotheses generate a set of predictions in combination with the two sub-hypotheses proposed for the nature of L2 input representations, namely the Phonetic and Phonological Processing hypotheses. Some of the experimental predictions have already been pointed out in the previous sections when the motivation behind choosing specific types of experimental materials was discussed. The purpose of this section is to summarize and contrast the predictions of all the hypotheses raised so far in relation to one another. Accordingly, Table 5.3 combines each of the main hypotheses with the two sub-hypotheses. In doing so, it evaluates the experimental items that are specifically designed to address these hypotheses. It should be noted that the predictions made in the table only apply to the Korean listeners' performance. It is assumed that English listeners will perform better than or at least equal to Korean listeners in responding to the experimental trials. This is because English listeners are at an advantage due to the fact that they will be listening to words that are intended to be in their native language.

An "X" in the evaluation table stands for a predicted problem with discriminating pairs that contrast the cluster in question (e.g., [p^hakma] vs. [p^hak^{h}uma]). That is, these are the items for which we expect perceptual epenthesis. On the other hand, the absence of an "X" conveys that Korean listeners are predicted to have no problems with distinguishing the pair in question.

Table 5.3: Evaluation table for the hypotheses

	CC Hypothesis		**COI Hypothesis**	
	Phonetic	**Phonological**	**Phonetic**	**Phonological**
k.t^h				
g.t^h	**X**		**X**	
k.m	**X**	**X**		
g.m	**X**	**X**	**X**	
l.n	**X**	**X**		
l.t^h				
c.t^h	**X**	**X**	**X**	**X**
c.m	**X**	**X**	**X**	**X**
j.t^h	**X**	**X**	**X**	**X**
j.m	**X**	**X**	**X**	**X**
	A	**B**	**C**	**D**

The left and the right two columns in the table belong to the CC hypothesis and CIO hypothesis, respectively. These two hypotheses are individually evaluated in combination with the two sub-hypotheses, namely the Phonetic Processing hypothesis (in shaded columns) and the Phonological Processing hypothesis (in unshaded columns). Notice that all hypotheses predict that [kt^h] and [lt^h] will be non-problematic. That is, successful discrimination is expected on these clusters since they contain a licit contact and a licit coda consonant. Likewise, since both [k] and [l] have phonemic status in Korean, the Phonetic Processing hypothesis and the Phonological Processing hypothesis predict these consonants to be mapped to Korean phonetic [k] and [l] and to phonemic /k/ and/l/, respectively. Since these two sets of mappings correspond to the same sounds under both sub-hypotheses, we expect the same outcome. Likewise, all hypotheses predict that [ct^h],[cm],[jt^h] and [jm] will be problematic. That is, since strident codas, namely [c] and [j], will be mapped to the Korean [c] and [j], respectively under the Phonetic Processing hypothesis, and to /c/ under the Phonological Processing hypothesis, we expect unsuccessful performance regardless of which sub-hypothesis we consider.

The crucial differences between the two main hypotheses arise in the experimental conditions that involve [km], [ln]. Both of these clusters have a permissible coda consonant but its occurrence with the following nasal sound is impermissible. Both clusters are expected to be unproblematic under the COI hypothesis since they contain a permissible coda consonant in the cluster. The CC hypothesis, in contrast, predicts that the same clusters will be problematic since the consonantal contacts in question are impermissible.

The pairs containing [gt^h] and [gm] distinguish between the two sub-hypotheses. Namely, if the Phonetic Processing hypothesis were confirmed, both of these clusters would be interpreted as [gt^h] and [gm] at the perceptual level, keeping all the phonetic information available in the mapping process. The Phonological Processing hypothesis, on the other hand, will suspend the predictable information, namely the cues in regard to voicing since it is allophonic in Korean. All voiced obstruents are derived systematically from voiceless plain obstruents in Korean. Therefore, the two clusters in the input, namely [gt^h] and [gm], will be interpreted as [kt^h] and [km], respectively, at the perceptual level of representation. What would the Phonetic Processing hypothesis predict in combination with the main hypotheses in this case? Since there are no consonant clusters in Korean with voiced consonants, the CC hypothesis will automatically rule out the [gt^h] and [gm] combinations (Column A). Likewise, since [g] cannot surface in coda position, both clusters are also ruled out by the COI hypothesis (Column C).

In contrast, if the Phonological Processing hypothesis were confirmed, [gt^h] and [gm] would be postulated as [kt^h] and [km], respectively, and since [k] is a legitimate coda, Korean speakers are predicted by the COI hypothesis to have no trouble with these clusters, which only accepts legal consonantal sequences (Column D). However, [km] is predicted to be problematic by the CC hypothesis, which only accepts legal consonantal sequences. The [kt^h] cluster will pose no problem under the same hypothesis (Column B).

5.3. Subject Population

Twenty-seven native speakers of Korean were asked for their voluntary participation. At the time of the investigation, all of the Korean participants were residing in the USA for educational purposes and had an affiliation with the University of Delaware either as a student (25) or a spouse of a student (2). None of the Korean subjects had started learning English before the age of 12. The age of the Korean participants varied between 20 and 37. As for the English speakers, 26 undergraduate students were recruited from an introductory linguistics course at the University of Delaware. The age of the English participants ranged between 18 and 22. None of them knew Korean. All of the English participants received course credit for their participation in the experiment.

Two of the Korean participants reported that they experienced extreme fatigue during the experiment and were unable to respond to a large number of trials. Therefore, these two participants were excluded from the analysis, leaving only 25 Korean participants. One of the participants from the English group was excluded from the analysis since he/she was not a native speaker of English, leaving 25 English participants.

5.4. Procedure

Each participant performed the AX discrimination task in a single testing session. Participants were presented with pairs of stimulus items through headphones (Plantronics Digital DSP300) at a comfortable volume by a desktop computer (Mac OS 9.1). They were asked to indicate whether the words in the pair sounded the same or different. Instructions about the experiment were given to each subject on the computer screen and they were allowed to read them at their own pace and ask questions of the experimenter for clarification. In the instructions they were specifically told that they were going to hear an American man saying nonsense words of English in pairs, and their task was to determine, after hearing each pair, if the man repeated the same word the second time or said a different word. The participants were asked to press <A> on the keyboard if they had heard the man say the same word the second time or <L> if they heard the man say a different word. Throughout the experiment, the computer screen displayed these messages constantly, saying "<A> for same" and "<L> for different" on the left and the right side of the screen, respectively. It should be noted that these two keys are the farthest apart in the middle row of letters on conventional Q-keyboards.

Unlike in Dupoux and colleagues' experiments, there was no training session. However, a practice session was conducted with each subject, which demonstrated exactly the same experimental protocol as the one in the actual test session. The practice session contained different items. The whole purpose of practice was to familiarize the participants with the nature of the task. No feedback was given regarding their responses. Each participant was given an "Informed Consent Form" (Appendix A) and a questionnaire (Appendix B) before the experiment started.

The experiment lasted about 50 minutes. There were 4 self-terminated rest periods. After each block in the experiment, the computer displayed a message saying that the rest period had begun. The participants could determine when they wanted to continue. The total number of blocks was 5. All experiments were carried out in a sound treated room at the Phonetics Laboratory of the Linguistics Department at the University of Delaware.

5.5. Analysis Methods

5.5.1. Signal Detection Theory

How can accuracy be quantified in discrimination tasks so that a subject's decision making process is made explicit by eliminating errors as much as possible? According to Signal Detection Theory (Green & Swets, 1974), a person's sensory response to change in a task can vary as a function of the knowledge and motivation that the circumstances necessitate under which the task is carried out. That is, certain situations may allow room for errors depending on how fatal or trivial the consequence of making that error is for the particular situation. For instance, diagnosis of cancer by using various methods such as X-ray or MRI imaging poses a serious challenge to doctors and radiologists. If there is a tumor and it is not detected by the experts, this may lead to a serious consequence. The question, then, becomes whether it is better to erroneously detect a tumor when it is not present than to not detect one when it is actually present.

Likewise, in the case of speech perception, it is very important to understand the nature of the judgment processes that underlie listeners' accuracy levels. These processes may differ from one subject to another as well as from one task to another. Therefore, it is crucial to verify how listeners' sensitivity to differences may be contaminated by the nature of the criteria in the task in question. For instance, listeners may have different ways of interpreting the instructions for a simple word recognition task. While one listener may wait until he/she is very sure that he/she can recognize the word before responding "yes", another may decide to respond "yes" upon hearing any hint of a recognizable word. Such differences in threshold levels in decision making will eventually yield a lot of "no" responses in the case of the first listener, or "yes" responses in the case of the second.

Similarly, a given discrimination task may include a high probability of encountering a change in the stimuli. Accordingly, the participants may set their threshold at a low level for comparing two stimulus items. In that case, a response bias towards saying “yes, I detected a change” becomes highly likely whenever he/she detects the slightest change in the stimuli. Accordingly, the person’s threshold turns out to be lower than it would be if the task had an equal likelihood of “yes” or “no” responses.

Signal Detection Theory (SDT) (Green and Swets, 1974; MacMillan & Creelman, 1991) was developed from a need to find an objective method to separate such response bias effects from sensitivity measures, thus to objectively evaluate the judgment of the listener. As its basic premise, SDT assumes that a listener's performance is a result of two processes: (a) a coding process that transforms the stimulus into an internal percept or sensation; (b) a decision process that maps the internal percept onto a response (Sawusch, 1996). The decision process compares the internal percept with a criterion. When the percept exceeds the criterion, a *hit* occurs. When the percept is less than the criterion a *miss* occurs. Whenever a non-target stimulus produces an internal percept that exceeds the criterion, a *false alarm* occurs. The four response outcomes are summarized in Table 5.4 below.

Table 5.4: Possible response outcomes on a trial of a yes-no experiment (from MacMillan, 2002: 46)

	RESPONSE	
STIMULUS	**"YES"**	**"NO"**
Signal	***Hit***	***Miss***
Noise	***False Alarms***	***Correct Rejection***

SDT has also been adapted to situations where subjects are being asked to detect a difference between two signals (e.g., "same-different" tasks; MacMillan & Creelman, 1991), as in the discrimination task in the present study. Accordingly, if stimuli S_1 and S_2 are being discriminated, there are four possible sequences to be taken into consideration: $<S_1, S_1>$;$<S_2, S_2>$;$<S_1,S_2>$; $<S_2, S_1>$. A hit will occur if the $<S_1, S_2>$ and $<S_2, S_1>$ pairs are judged to be different, a miss if they are judged to be the same. A false-alarm will occur if $<S_1,S_1>$ and $<S_2, S_2>$ are judged to be different, a correct rejection if they are judged to be the same.

In the present study, the participants' performance on each "different" doublet yielded a total of 30 data points per subject, 10 for hits or misses (e.g., 10 instances of [p^{h}ákma] vs. [p^{h}ák^{h}ʊma]) and 20 for false alarms or correct rejections (e.g., 10 instances of [p^{h}ákma] vs [p^{h}ákma], and 10 instances of [p^{h}ák^{h}ʊma] vs. [p^{h}ák^{h}ʊma]). Table 5.5 provides the possible outcomes of the example pair [p^{h}ákma] vs. [p^{h}ák^{h}ʊma] according to SDT:

Table 5.5: An SDT analysis scheme of an example doublet

Trial: ([p^{h}ákma]-[p^{h}ákuma])	**RESPONSE**	
	DIFFERENT	**SAME**
(Target Present) [p^{h}ákma]-[p^{h}ák^{h}ʊma] [p^{h}ák^{h}ʊma]-[p^{h}ákma]	***Hits***	***Misses***
(No-Target) [p^{h}ák^{h}ʊma]-[p^{h}ák^{h}ʊma] [p^{h}ákma]-[p^{h}ákma]	***False Alarms***	***Correct Rejections***

Essentially, SDT reduces the total number of representations of each stimulus type to proportions, namely p(H) (the proportion of "different" responses when the words in the doublet are different, or hit rate) and p(F) (the proportion of "different" responses when the words in the doublet are same, or false alarm rate). Accordingly, p(H) and p(F) for each "different" doublet (e.g., [p^hákma]-[p^hák^{h}ʊma]) in the experiment were calculated using the formulae in (2):

(2) p(H)= #Hit / (#Hits+#Misses)

p(F)= #False Alarms / (#False Alarms+# Correct Rejections)

A d' score for each "different" doublet was calculated per participant in the following way. The p(H) and p(F) values were first converted to z-scores, yielding z(H) and z(F), respectively. Doing so standardized the scores obtained for the groups with a normal distribution for the population. Then, z(F) was subtracted from z(H), leaving only those cases where a participant correctly detected a difference. Using the Independent-Observation Model[18], the d' value was then calculated with the formula from Table A.5.3 in MacMillan and Creelman (1991; see also MacMillan, 2002).

[18] It should be noted that the AX paradigm involved in the present experiment is neither a classic roving design nor a classic fixed design (MacMillan and Creelman, 1991: 147-159). In roving designs, stimuli are equally spaced in some physical units along a single continuum and presented within the same block. An analysis based on a decision procedure called the differencing strategy is used in roving designs, where the two observations on a trial are subtracted, and the result is compared to a criterion (see especially Figures 6.1 and 6.3 in MacMillan and Creelman (1991) for further

It should be noted that, a d' value is obtained from this analysis can in principle range from 0 to any value. A d' value of 0 means no sensitivity to the signal. The d' values obtained in the present study varied from 0 to 4.21 (including correction; see below).

Table 5.6 shows an example frequency matrix that is used for the calculation of Participant K17's discriminability score for the doublet [pʰáktʰa] vs. [pʰákʊtʰa]. As can be seen from the matrix in Table 5.6, the participant responded with "same" only once for the doublet [pʰáktʰa] vs. [pʰákʊtʰa] (thus 9 hits and 1 miss). Likewise, the same participant yielded a false alarm rate of 1 by responding with "different" 1 out of 20 times when the doublet in question was not actually different (i.e., [pʰákʊtʰa] vs. [pʰákʊtʰa], or [pʰáktʰa] vs. [pʰáktʰa]).

---

explanation of this difference). In fixed experiments, a single stimulus pair is used and the analysis based on an Independent-Observation Model is employed, where the decision rule is to decide separately whether each item is S_1 or S_2, then report whether these subdecisions are the same or different. Since the test doublets were drawn randomly from a pool of exemplars, they were different from each other in the present study and since they were presented in the same block, it cannot readily be classified as a classic fixed design. But, since the words contained in the doublets were not intended to vary equi-distantly from one another in a continuum, it is also not a classic roving design. However, the intended listening strategy that was imposed on the subjects is compatible with the decision rule given by the Independent-Observation Model. Specifically, the AX paradigm was designed to induce a phonological analysis of the stimuli, where the optimal decision rule for subjects was to construct a percept of each word in the doublet and then report whether these percepts were the same or different. The basic premise of the Differencing Model, however, leads to an acoustic comparison of the test words, where the acoustic images of the two words are subtracted. Given the 1500 ms of ISI used in the experiment, such a strategy is unlikely. Therefore, the pairs in this experiment are best analyzed in accordance with the Independent-Observation Model.

Table 5.6: A frequency matrix for Participant K17's d' score for [p^hákt^{h}a] vs. [p^hákʊt^{h}a]

Hits	9
Misses	1
False alarms	1
Correct rejections	19
p(H) =	0.9
p(F) =	0.05
z(H) =	1.281551
z(F) =	-1.64485
z(H)-z(F)	2.926404
d' =	***3.567311***

It should be noted that there are cases where a subject may exhibit perfect proportions. These are the cases where p(H)=1 or p(F)=0. That is, if a subject has no misses, p(H) becomes 1. Likewise, if he/she has no false alarms, p(F) becomes 0 (0 divided by any number is 0). Consequently, the SDT statistics, which involve the calculation of z-scores for each of the probabilities, returns an error. The present experiment had both of these cases. That is, there were cases where some of the participants had perfect hit rates (i.e., p(H)=1) and/or false alarm rates (i.e., p(F)=0) for certain doublets. In other words, they had no trouble giving "different" responses when the words were different and "same" responses when they were the same. Different correction methods have been proposed in the literature to handle such cases. For instance, adding and subtracting 0.5 to the frequency matrix when necessary has been suggested in Kadlec (1999). Another alternative method is to add 0.5 to all cells (e.g., Snodgrass & Corwin, 1998). In order to calculate d' values in such cases in the present study, a value of 0.5 was substituted for misses and/or false alarms that were zero (Tables 5.7, 5.8, 5.9).

Table 5.7: An example corrected d' calculation matrix where the participant originally has no false alarms but receives 0.5 to correct for this.

Hits	9
Misses	1
False alarms	**0.5**
Correct rejections	20
p(H) =	0.9
p(F) =	0.02439
z(H) =	1.281551
z(F) =	-1.9705
z(H)-z(F)	3.252053
d' =	***3.863388***

Table 5.8: An example corrected d' calculation matrix where the subject originally has no misses or false alarms but receives 0.5 for each to correct for this.

Hits	10
Misses	**0.5**
False alarms	**0.5**
Correct rejections	20
p(H) =	0.952381
p(F) =	0.02439
z(H) =	1.668391
z(F) =	-1.9705
z(H)-z(F)	3.638893
d' =	***4.215908***

Table 5.9: An example corrected d' calculation matrix where the subject originally has no misses but receives 0.5 to correct for this.

Hits	10
Misses	**0.5**
False alarms	1
Correct rejections	19
p(H) =	0.952381
p(F) =	0.05
z(H) =	1.668391
z(F) =	-1.64485
z(H)-z(F)	3.313244
d' =	***3.919067***

Essential to the calculation of d' scores is the size of the false alarm rate (p(F)) compared to the hit rate (p(H)). If p(F) is equal to or greater than p(H), d' values are defined to be zero in STD statistics. This was typically the case for Korean participants, who had virtually no hits and no false alarms for certain doublets, leaving both p(F) and p(H) with a value of 0 (Table 5.10).

Table 5.10: An example frequency matrix where p(F)=p(H) and d' value is zero by definition.

Hits	0
Misses	10
False alarms	0
Correct rejections	20
p(H) =	0
p(F) =	0
z(H) =	-
z(F) =	-
z(H)-z(F)	-
d'	**0** (by definition)

The method for calculation of discriminability indices illustrated in this short overview above will provide the basis for the quantification of each subject's "same" and "different" responses in the experimental data, which are presented in the following chapter.

Chapter 6

RESULTS

6.1. Introduction

This chapter reports and discusses the results from the AX discrimination of consonant clusters by Korean and English listeners. The primary research question is whether perceptual epenthesis is caused by coda violations or contact violations. The consonant clusters tested in the experiment ([C_1C_2]) are either licit or illicit clusters in Korean. The illicit clusters contain either a possible coda consonant or an impossible one in the C_1 position, as summarized in Table 6.1 below.

Table 6.1: Consonantal contexts used in the experiment.

			C_2	
			Oral Stop (i.e., [t^h])	***Nasal (i.e., [n]or[m])***
C_1	**Licit**	**[k]**	Licit	Illicit
		[l]	Licit	Illicit
	Illicit	**[c]**	Illicit	Illicit
		[j]	Illicit	Illicit
		[g]	Illicit	Illicit

The discrimination task involves a same-different judgment on two nonce words that are different from one another by the presence ([C_1VC_2]) or absence ([C_1C_2]) of a vowel between the cluster. The assumption is that if Korean listeners hear illusory vowels due to the phonotactics of Korean, they should judge such doublets to be the same. Furthermore, the study also investigates the way L2 speech inputs are processed in the perceptual system. Specifically, the study asks whether Korean listeners represent the feature [voice] in L2 speech signals, which is non-contrastive and predictable in Korean by an intersonorant voicing rule, as discussed in Chapter 3, Section 3.3. Finally, some test doublets contain a comparison of illicit consonant clusters in Korean (e.g., *[km]) and their likely output forms in the Korean production grammar (e.g., [ŋm]). These are embedded in the experiment in order to test whether Korean listeners apply production rules of Korean to their perception.

In the following, I provide the descriptive statistics of the data. Three groups of consonant clusters are identified for the Korean group: (1) the strident-clusters, namely [cm], [jm], [ct^h], [jt^h], where there is near-indiscriminability, (2) [gm] and [km], where there is intermediate performance, and (3) [gt^h], [ln], [lt], [kt^h], where there is successful discrimination. However, further analyses, such as the Cost Analysis, will reveal only one truly significant condition: Korean strident clusters. This provides evidence for the Coda/Onset Identity and Phonological Processing Hypotheses.

6.2. Descriptive Statistics

Figure 6.1 below graphically presents the mean d' values for consonant clusters in the epenthesis condition for the English and the Korean groups. Appendix C gives a complete list of English and Korean mean and median d' values and standard deviations for all the experimental doublets. Each individual subject's d' scores for each test pair is provided in Appendix D.

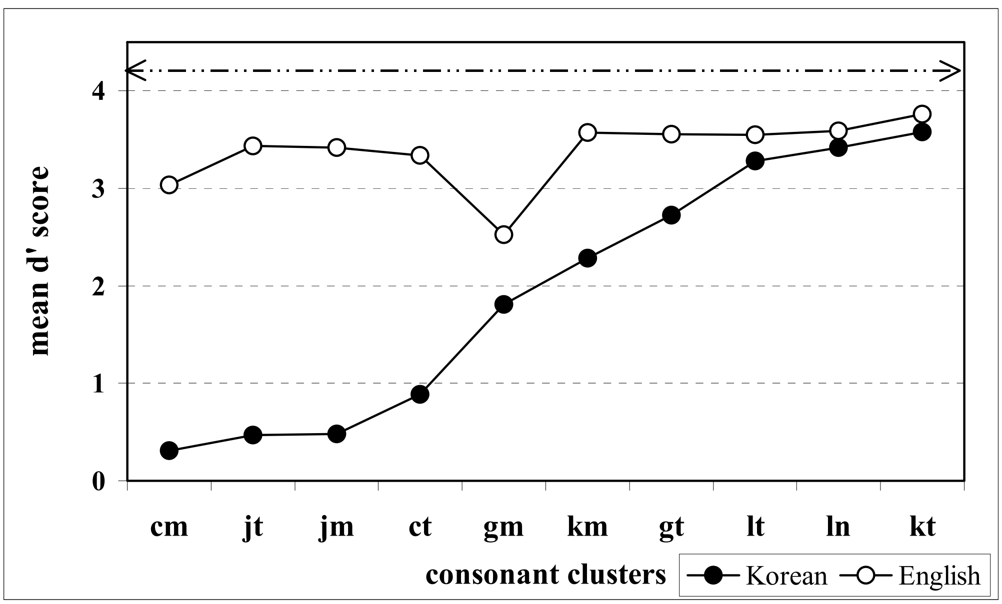

Figure 6.1: Korean and English *mean* d' scores for the discrimination of [CVCCV] vs. [CVCVCV] words (ordering by increasing d' values for Korean). The arrow at the top of the chart shows the perfect performance (4.21) in the experiment.

The highest and lowest attainable individual d' scores in both groups (with correction for instances of zero hits, misses, false alarms or correct rejections) is 4.21 and 0, respectively. The highest mean d' value for individual pairs in the subject groups is 3.77 (English, [kt^h]), while the lowest mean is 0.31 (Korean, [cm]).

We can make several observations from the d' scores in Figure 6.1 above. It is quite apparent that the English group, as expected, is successful at discriminating all forms with consonant clusters from those with vowels (e.g., [p^{h}act^{h}a] vs. [p^{h}ac^{h}it^{h}a]). All the mean d' scores in the English group exhibit an almost flat profile, ranging between 3 and 4 (average d'=3.38), except in the case of [gm] (average d'=2.53), which is discussed in more detail below. The Korean group's discrimination indices in the same case, however, range from 0.31 to 3.57.

Observationally, the Korean listeners' d' scores cluster in three groups: (1) [cm], [jm], [ct^h] and [ct^h], where the mean d' values are below 1; (2) [gm], [km] and [gt^h], which forms an intermediate category with d' values roughly between 2 and 3; and (3) [ln], [lt^h] and [kt^h], where the d' scores are very close to the English group's mean d' scores.

Given that the range of d' scores is 0 to 4.21, we can ask what ranges of d' values between the two extremes would constitute "moderate" and "successful" performance. For instance, should a d' value of 2.57, as in the case of English [gm], indicate "moderate" or “successful” discrimination ability? Similarly, should a d' value of 1.80, evidenced in the Korean [gm], be grouped with scores below 1, those that indicate almost no discrimination ability, or with values that are above 3? That classifying performance levels is not automatic in SDT should not be viewed as a drawback of the SDT quantification used in the analysis. The same question would arise even if the accuracy levels in discrimination were calculated as percentage correct.

Nevertheless, there is a clear division between lack of discrimination ability and good performance. On the one side, there is apparently poor discrimination (or rather lack of discriminability), evident in the Korean group's d' scores lower than 1 on certain clusters. On the other side, there is successful discrimination where d' scores reached higher than 3 in the same language group. We can conclude that those clusters that involve the affricate sounds, namely [cm], [ct^h], [jm] and [jt^h], that is the strident clusters, constitute the most troublesome cases for the Korean group, with an average d' score of 0.53, indicating an inability to discriminate. A mean d' value below 1 in a population as large as 25 raises the possibility that an overwhelming number of participants may have obtained d' scores of absolute 0 in the Korean data. It should be noted that when the hit rate is 1 and the false alarm rate is 0 (which is corrected by adding 0.5; see Section 5.5.1 for more information), the d' score is 0.89. When the hit rate is 0 and the false alarm rate is 0, however, the d' score automatically corresponds to 0. Everything else being equal, for those individuals who have a tendency give “same” responses when stimuli are different, there cannot be an intermediate score between 0 and 0.89 since a shift from no hit to 1 hit results in a difference of 0.89. This is what we observe in the Korean data: 19 out of 25 Korean participants have a d' score of 0 for the doublet [p^hacma] vs. [p^hac^{h}ima]. This is 18 out of 25 for [jt^h], 16 out of 25 for [jm], and 13 out of 25 for [ct^h].

The *median* d' scores plotted in Figure 6.2 below further demonstrate that all of the strident clusters in the Korean group have a median d' score of 0. Conversely, Korean median d' scores on the rest of the consonant clusters are values close to those of the English group. Especially, the median d' scores for [kt^h], [ln] and [lt^h] are very close in both groups. This suggests that the both groups' behavior on these clusters shows close similarities.

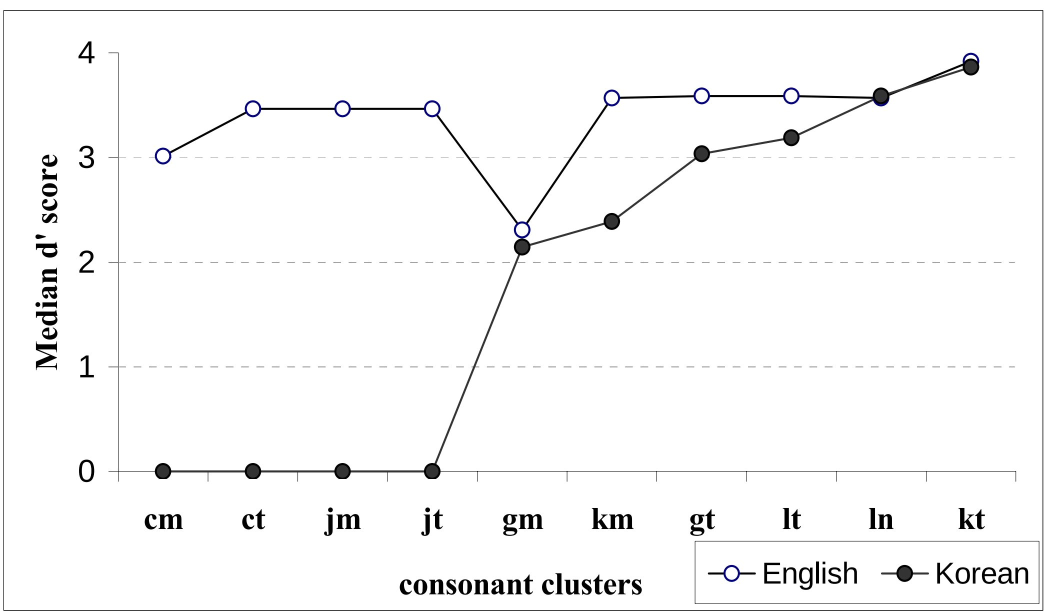

Figure 6.2: Korean and English *median* d' scores for the discrimination of [CVCCV] vs. [CVCVCV] words (ordering by increasing values for the Korean group).

Interestingly, the English group's performance on [gm] exhibits a sharp decrease, merging with the Korean group's performance. It is quite clear, however, that both groups' performances on this consonant cluster are not as bad as the Korean group's performance on the strident clusters. I will come back to this point below.

6.3. Analysis of Variance

To test the hypotheses raised earlier, a 5x2x2 (i.e., C_1 by C_2 by Language Group) repeated measures Analysis of Variance (ANOVA) was performed on the d' values in the epenthesis condition using an alpha level of .05. The within-participant factors are (1) C_1, the first member of the C_1C_2 cluster, ranging over [c], [j], [k], [g] and [l], and (2) C_2, which is either a plosive ([t^h]) or a nasal consonant ([m] or [n]). Language group (Korean and English) is the between-participant factor. The results from the ANOVA multivariate tests are provided in Table 6.2.

Overall, there is a significant effect of language ($F(1,48)=184.627$, $p<.0001$). This is due to the fact that the Korean participants have lower overall d' scores than the English participants. In particular, they have much lower d' scores on a number of items such as all strident clusters, as discussed above. The analysis also reveals overall main effects of the first consonant (C_1) in the cluster, ($F(4,45)=96.131$, $p<.0001$) and the second consonant (C_2) in the cluster, ($F(1,48)=55.555$, $p<.0001$), indicating that the d' values vary as a function of the first consonant as well as the second consonant when both groups were combined. The main effect of C_1 is not surprising given that performance on certain consonant clusters, particularly those with stridents as C_1, is much lower than those containing other consonants (e.g., [k] and [l]) in the Korean group.

Table 6.2: ANOVA Multivariate Tests (using alpha=.05)

Effect		F	Hypo. df	Error df	Sig.
C1	Pillai's Trace	96.131	4	45	.0001
	Wilks' Lambda	96.131	4	45	.0001
	Hotelling's Trace	96.131	4	45	.0001
	Roy's Largest Root	96.131	4	45	.0001
C1*Language	Pillai's Trace	57.871	4	45	.0001
	Wilks' Lambda	57.871	4	45	.0001
	Hotelling's Trace	57.871	4	45	.0001
	Roy's Largest Root	57.871	4	45	.0001
C2	Pillai's Trace	55.555	1	48	.0001
	Wilks' Lambda	55.555	1	48	.0001
	Hotelling's Trace	55.555	1	48	.0001
	Roy's Largest Root	55.555	1	48	.0001
C2*Language	Pillai's Trace	4.159	1	48	.047
	Wilks' Lambda	4.159	1	48	.047
	Hotelling's Trace	4.159	1	48	.047
	Roy's Largest Root	4.159	1	48	.047
C1*C2	Pillai's Trace	18.492	4	45	.0001
	Wilks' Lambda	18.492	4	45	.0001
	Hotelling's Trace	18.492	4	45	.0001
	Roy's Largest Root	18.492	4	45	.0001
C1*C2*Language	Pillai's Trace	6.080	4	45	.001
	Wilks' Lambda	6.080	4	45	.001
	Hotelling's Trace	6.080	4	45	.001
	Roy's Largest Root	6.080	4	45	.001

As for the effect of C_2, when this consonant is a nasal consonant, both groups suffer a decrease in d' values. This is especially evident in the combined effects of the clusters [gm] and [cm], which are lower than [gt^h] and [ct^h] in both groups, respectively. Furthermore, a two-way interaction effect is found for C_1 by Language ($F(4, 45)=57.871$, $p<.0001$). This effect is expected especially given that the Korean group's scores are substantially lower for all items that contained the strident consonants. Another two-way interaction emerges for C_2 by Language, although this barely reaches significance ($F(1, 48)=4.159$, $p<.047$). This is due to the fact that the d' scores of the Korean group for [cm], [km] and [gm] are lower than [ct^h], [kt^h], and [gt^h], respectively. The C_1 by C_2 comparison reveals yet another two-way interaction ($F(4, 45)=18.492$, $p<.0001$), suggesting that, overall, the performance on C_1 varies as a function of C_2. Ordinarily, these two-way interactions may be considered the first step evidence for an effect of consonantal contact. However, these constitute insufficient evidence to draw firm conclusions given the fact that a three-way interaction is also found for C_1 by C_2 by Language ($F(4,45)=6.08$, $p<.001$). This indicates that all three factors contribute to the variation in d' scores, which makes the interpretation of results complicated. In order to properly interpret the two- and three-way interactions, post-hoc tests were run to distinguish groups, which were significantly different from each other statistically.

Table 6.3 below reports the means, standard error and 95 percent confidence intervals for C_1 and C_2 for both groups. Figure 6.3 graphically illustrates the upper and lower bounds within 95% confidence level for both groups. A complete list of the results for the Scheffe and Tukey/Kramer post-hoc tests on all test clusters is provided in Tables E.1 and E.2 in Appendix E (for the English group) and Tables F.1 and F.2 in Appendix F (for the Korean group).

Table 6.3: Mean d' scores, standard errors, lower and upper bounds within 95% confidence rate for CVCCV vs. CVCVCV discrimination.

				95% Confidence Level	
Language	**C_1+C_2**	**Mean**	**Std. Error**	**Lower Bound**	**Higher Bound**
English	cm	**3.032**	0.125	2.781	3.284
	jm	**3.416**	0.141	3.132	3.700
	ct^h	**3.338**	0.163	3.010	3.665
	jt	**3.436**	0.134	3.167	3.706
	km	**3.575**	0.171	3.231	3.918
	gm	**2.526**	0.214	2.096	2.957
	kt	**3.764**	0.103	3.556	3.972
	gt^h	**3.554**	0.165	3.222	3.886
	ln	**3.587**	0.108	3.370	3.804
	lt^h	**3.552**	0.119	3.314	3.790
Korean	cm	**0.308**	0.125	0.056	0.559
	jm	**0.478**	0.141	0.195	0.762
	ct^h	**0.887**	0.163	0.559	1.215
	jt	**0.469**	0.134	0.199	0.738
	km	**2.282**	0.171	1.938	2.625
	gm	**1.807**	0.214	1.377	2.238
	kt^h	**3.580**	0.103	3.372	3.788
	gt^h	**2.724**	0.165	2.393	3.056
	ln	**3.421**	0.108	3.204	3.638
	lt^h	**3.280**	0.119	3.042	3.518

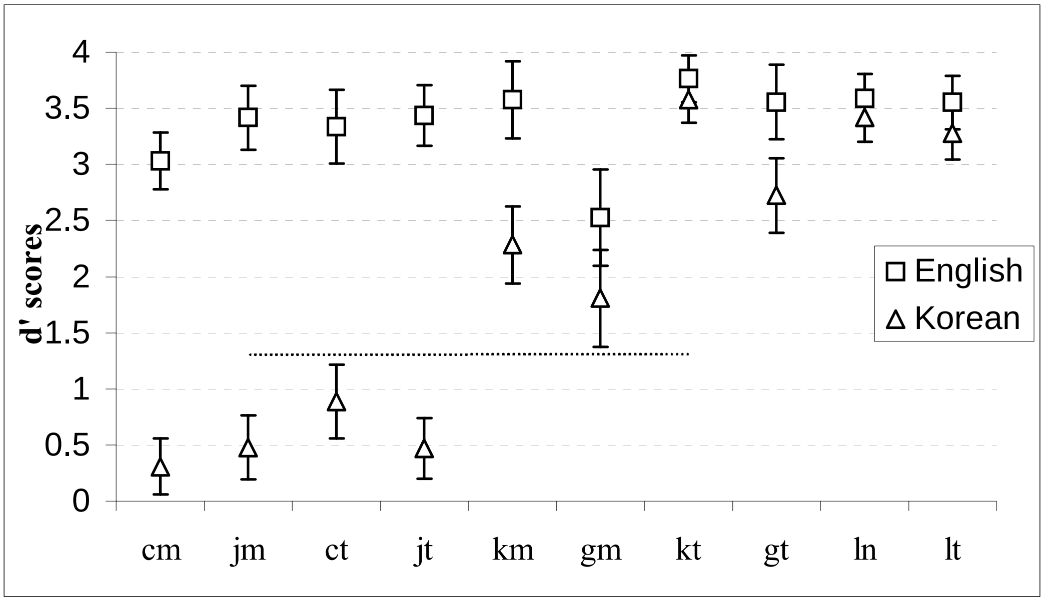

Figure 6.3: The upper and lower bounds for d' scores for CVCCV vs. CVCVCV discrimination with 95% confidence interval. The line indicates that the confidence intervals for [ct] and [gm] in the Korean group do not overlap.

Looking at Table 6.3 and Figure 6.3 above, we can say that many of the confidence intervals in the English group overlap and they have relatively small standard errors, except for [gm] whose upper and lower bounds do not overlap with those of any of the clusters other than [cm]. Interestingly, the confidence intervals for [cm] do not overlap with those for [kt^h], [ln] and [lt^h] in the English group. A post-hoc analysis using a Scheffe test with an alpha level of .01, a very conservative post-hoc test, further supports the fact that [gm] is significantly different from other clusters in the English group except for [cm]. However, under the same test, [cm] is only significantly different from [kt^h]. The way the consonant clusters are grouped based on the results from a Scheffe test in the English group is graphically illustrated in Figure 6.4 below.

[gm] [cm] [ct^h] [gt^h] [km] [lt^h] [ln] [kt^h]

Figure 6.4: Consonant cluster groupings in the English group based on a Scheffe post-hoc test

The structural reason why [cm] is statistically different from a few other clusters in the English group (i.e., from [kt^h] under the Scheffe test, and from [kt^h], [ln] and [lt^h] under the Tukey/Kramer test) is unclear. Given that the mean d' value on this item was 3.02, a high score, any speculation on statistical differences between this item and other items is unmotivated and uninteresting given the purposes of this dissertation.

The fact that [gm] is significantly different from all other clusters (except [cm]), however, deserves further consideration and speculation. The cluster in question has the lowest mean d' score in the English group and the highest standard deviation (SD). The SD for this item is 0.91, a very high number when compared to an SD of 0.67 for [cm], the second largest SD in the English group. This means that there are vast individual differences in the English group's d' scores for this item. A closer look at the individual data reveals that 7 participants in the English group scored below 2 on this cluster. The median is also low on this item (2.14), suggesting that about half of the subjects experienced some trouble with it. There is no obvious explanation for the differential behavior on the [gm] cluster, especially given the fact that English has quite a number of words containing this cluster, such as *pigment, sigma, pragmatic*, *dogma*, etc. Thus, the degraded performance on this cluster could not be due to frequency especially when we compare this cluster to [ct] and [cm], which do not occur in simplex words in English. Although there is moderate difficulty with this item, the English group's overall d' scores on this cluster is not as poor as the Korean group's d' scores on the strident clusters, where the mean d' values approximate 0. Therefore, I will consider the mean d' value on this cluster for the English group as an indication of, at least, moderately good discriminability. Other than the small problem with this cluster, the English group's performance on all other clusters shows close similarities. Therefore, there is nothing interesting in the English data. They convincingly show that English listeners, as the control group in this study, behave as successfully as expected.

6.4 Post Hoc Analyses on the Korean Data

Unlike the English group, the Korean group responds differently to many different clusters. Figure 6.5 below groups the consonant clusters based on the evaluation of the post hoc analyses from the Scheffe and Tukey/Kramer tests on the Korean group's d' scores.

[cm], [jt^h], [jm], [ct^h]	[gm], [km], [gt^h]	[ln], [lt^h],[kt^h]
low d'		*high d'*

Figure 6.5: Consonant cluster groupings in the Korean group based on post-hoc tests.

Figure 6.5 shows that the post hoc tests divide the consonant clusters into three different groups for the Korean listeners. On the one end, there are consonant clusters, all of which consist of strident codas with very low d' scores, and, on the other end, there are clusters, namely [ln], [lt^h],[kt^h] with very high d' scores. In the middle category, we have [gm], [km], [gt^h], which seem to cause problems for the Korean group, but these problems are significantly less than what we see with strident clusters.

It should be recalled that both of the main hypotheses, namely the Coda/Onset Identity Hypothesis and the Consonantal Contact Hypothesis, have some common predictions. First, they both expect the items that contained the strident consonants to produce poor results in the Korean group. Given the fact that all the items with the strident consonants as C_1 yield the lowest d' scores, the fundamental prediction of the two main hypotheses, find validation. Second, both of the main hypotheses predict that the sequences [kt^h] and [lt^h] would pose no problem for the Korean group. If we look back at the confidence intervals given in Figure 6.3, the Korean group's mean d' scores on the same items indeed overlap with the d' scores of the English group's discrimination. It should be remembered that these sequences result in a cluster situation that is possible in Korean and, necessarily, involve a consonant that can function as a permissible coda. Thus, one of the common predictions of both of the hypotheses is again confirmed. The high d' scores for these clusters in the Korean group confirm that the perceptual epenthesis effect does not mechanically arise with every consonantal sequence. That is, there are clusters that the Korean group could indeed discriminate at least as well as the English group.

With the data divided into three distinct groups, however, we cannot evaluate the conflicting hypotheses in any simple and meaningful way. First, we cannot say that the intermediate performance on [km] and [gm] supports the CC Hypothesis while the clusters at the high end conform to the predictions of the COI Hypothesis. This is because the cluster [ln], which contains a consonantal contact violation just like [km] and [gm], is among the top three successfully discriminated clusters in the Korean group. Both the English group's and the Korean group's mean d' scores for this cluster have overlapping confidence intervals. This means that the Korean group was able to discriminate the cluster as successfully as the English group. This finding, by itself, disputes the CC Hypothesis. This finding is discussed further in Section 6.6 below.

Second, the hypotheses were conceptualized in such a way that they allow for only two distinct levels of performance on the test consonant clusters: those that would yield bad performance and those that would yield good performance. To better evaluate the hypotheses, we now need to distinguish between relative degrees of badness. One approach to this problem is to pose the following question: How can we group the consonant clusters if only two different groups of performance were to be made, specifically, those that are categorically bad and those that are categorically good? A statistical approach to such a problem is Cluster Analysis, to which we now turn.

6.5 Cluster Analysis

Cluster Analysis (CA), a multivariate analysis technique, seeks to organize information about variables to sort cases into clusters so that the degree of association is strong between members of the same cluster and weak between members of different clusters. Most CA techniques are hierarchical, which results in classifications that look like a phylogenetic taxonomy. CA uses a dendogram, a tree structure, which graphically clusters cases by typically starting with single member clusters, which are gradually fused until one large cluster is formed, a clustering technique named agglomeration. The clusters are joined at increasing levels of dissimilarity. The farther the clustering takes place, the more dissimilarity there is between the members of the cluster.

Using SPSS, a hierarchical Cluster Analysis of Korean d' scores was performed on the individual participants' d' scores in the Korean group using a squared Euclidean distance measure and an Average Linkage clustering algorithm. This algorithm calculated the dissimilarity between clusters using cluster average values. Figure 6.6 presents the Cluster Analysis dendogram for the Korean data.

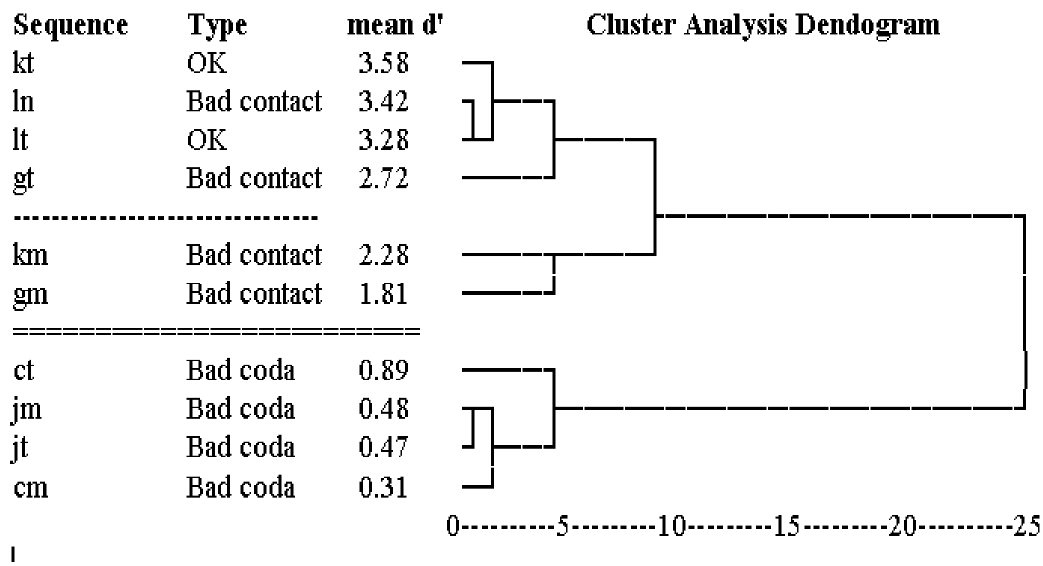

Figure 6.6: Cluster analysis dendogram

The ruler at the bottom of the dendogram is an index of an arbitrary measure of dissimilarity. Starting from the left side of the dendogram, we see that the analysis divides the consonant clusters first into 8 groups, and then into 6 groups. As we move further to the right, we get 3 groups at about 5 units on the ruler. Finally, at about 10 units, the consonant clusters are divided into two large groups. The analysis necessarily groups all the items into one big group in the end.

The 8-way clustering and the 6-way clustering that initially emerge in the analysis do not constitute a coherent group in any phonologically meaningful sense. The 3-way clustering and the 2-way clustering, however, reveal some important generalizations about the data relevant to the hypotheses at issue. The 3-way clustering of the consonant clusters is given in Figure 6.7 below.

[cm], [jt^h], [jm], [ct^h]	**[gm], [km]**	**[gt^h], [ln], [lt^h],[kt^h]**

Figure 6.7: Three-way clustering of the consonant clusters using Cluster Analysis.

When we compare this clustering with the way the ANOVA post-hoc tests have grouped the consonant clusters in Figure 6.5 above, we see that there is a disagreement on the status of [gt^h]. While the bad contact cases, namely [gm] and [km], form the middle category in both analyses, [gt^h] is agglomerated with the successfully discriminated consonant clusters in the Cluster Analysis. That [g.t^h] grouped together with [kt^h] suggests that voicing information on C_1 does not matter for the Korean listeners, which rules out the Phonetic Processing Hypothesis. That is, the feature [voice] on C_1 does not cause as much of a significant problem as the feature [nasal] on C_2, although the problem with [nasal] is significantly less than the feature [strident].

The 2-way clustering that emerges out of the Cluster Analysis is provided in Figure 6.8.

[cm], [jt^h], [jm], [ct^h] ‖ **[gm], [km], [gt^h], [ln], [lt^h],[kt^h]**

Figure 6.8: Two-way clustering of the consonant clusters using Cluster Analysis.

This clustering groups all the strident codas together in one group, and all other cases in another group. While the 3-way clustering cannot choose one set of hypotheses over the others, the 2-way clustering finds only one match in the sets of hypotheses. The two groups given in Figure 6.8 above fits the predictions of Column D of the evaluation table repeated in Table 6.4 below, where only strident codas are expected to be bad. The way the Cluster Analysis divides the data into two groups, thus, confirms only the Coda/Onset Hypothesis in combination with the Phonological Processing Hypothesis.

Table 6.4: Evaluation table for the hypotheses

	CC Hypothesis		COI Hypothesis	
	Phonetic	Phonological	Phonetic	Phonological
$k.t^h$				
$g.t^h$	X		X	
k.m	X	X		
g.m	X	X	X	
l.n	X	X		
$l.t^h$				
$c.t^h$	X	X	X	X
c.m	X	X	X	X
$j.t^h$	X	X	X	X
j.m	X	X	X	X
	A	B	C	D

The degradation in the Korean listeners' performance on [gm] and [km], that together form the potential intermediate category, may be taken as evidence for the effect of consonantal contact on their perception. Since these clusters contain a nasal segment as C_2, it could be that the contact actually mattered for the Korean listeners. At this point, it is crucial to draw our attention to the Korean groups' performance on the cluster [ln], which also contains a nasal segment as C_2. The [l]-clusters, namely [ln] and [lt^h], constitute a replication of the same question that the clusters [km] and [kt^h] aim to test in the experiment. Therefore, these two clusters can be used as an independent test for the evaluation of the main hypotheses of the present study.

6.6 The Analysis of the Clusters [ln] and [lt^h]

Let us recall what the main hypotheses predicted about [ln] and [lt^h]. In Korean the sequence [l.t^h] is a licit heterosyllabic contact while [l.n] is illicit. The Consonantal Contact Hypotheses, therefore, predicts the cluster [l.n] to be problematic for the Korean group. Since [l] is a legitimate coda consonant in Korean, both clusters should not cause any difficulty under the Coda/Onset Identity Hypothesis. Table 6.5 gives the mean and the 95% confidence intervals of both the English and the Korean groups' d' scores.

Table 6.5: Mean d' scores, standard errors, lower and upper bounds within 95% confidence rate for [ln] and [lt^h]

				95% Confidence Level	
Language	**C_1C_2**	**Mean**	**Std. Error**	**Lower Bound**	**Higher Bound**
English	ln	**3.587**	0.108	3.370	3.804
	lt^h	**3.552**	0.119	3.314	3.790
Korean	ln	**3.421**	0.108	3.204	3.638
	lt^h	**3.280**	0.119	3.042	3.518

As shown Table 6.5 above, the confidence intervals for the clusters [ln] and [lt^h] overlap in both groups, showing that these clusters are not different from each other. Furthermore, both group's scores on both of these consonants are not significantly different from one another with a 95% confidence rate. A power analysis reveals that any real difference in d' scores for [lt^h] and [ln], if such a difference exists, must be less than 0.6 units of d' with 90% confidence, and must be less than 0.8 units of d' with 99% confidence. With such small differences, the statistics on the clusters [ln] and [lt^h], thus, provide no support for the Consonantal Contact Hypothesis.

Separating the clusters with sonorant segments as C_1, namely [ln] and [lt^h], from the rest of the clusters gives us a more homogenous group of clusters, all of which have obstruent segments as C_1. In fact, we now have a complete 2x2x2x2 design. For C_1, we have palatal vs. velar and voiced vs. voiceless; for C_2, we have nasal vs. oral; and for Language, we have Korean vs. English. Figure 6.9 compares the groupings of obstruent clusters by the ANOVA post-hoc analyses and the Cluster Analysis (CA).

ANOVA:	**[cm], [jt^h], [jm], [ct^h]**	**[gm], [km], [gt^h]**	**[kt^h]**
CA:	**[cm], [jt^h], [jm], [ct^h]**	**[gm], [km]**	**[gt^h], [kt^h]**

Figure 6.9: Three-way clusterings using Cluster Analysis and ANOVA.

Having independently found support for the Coda/Onset Identity Hypothesis with the sonorant clusters, we are still left with the same problem as above when we look at the obstruent clusters. Can the intermediate performance on [gm] and [km] be due to the effect of a contact violation? One could perhaps attribute the successful discrimination of [ln] due to the fact that C_1 is sonorant and suggest that consonantal contact effects still hold for the obstruent consonants because when a cluster has an obstruent C_1 and a nasal C_2, the performance is degraded. Similarly, at least under the ANOVA post-hoc analyses, the same argument can also be made about voicing: when C_1 is a voiced segment, d' scores are lower.

So far, it looks as if both voicing and nasality are adding some difficulty to the discrimination of obstruent clusters for the Korean group. But, how large an effect are these features responsible for? One way to answer this question is to use the amount of variation in scores to contextualize the difference. If we have an idea of the amount of variation found within a group, we can use this as a ruler against which to compare the difference. This idea is quantified in the calculation of the effect size. For the present study, we can employ effect size statistics to understand the effect of the features [strident], [nasal] and [voice] on the distribution of d' scores of the obstruent clusters. In the following, two different types of analyses are carried out to try to measure the effect size. The first analysis is a statistical test, using a stepwise multiple regression model and analyzing the proportion of variance covered by various models under consideration. The second analysis calculates the cost of a given feature in units of d' by looking at the differences between consonant clusters that differ from one another by that feature.

6.7 Effect Size: Statistical Tests on Obstruent Clusters

Multiple regression provides a measure of the effects of more than one independent variable on a dependent variable without the interference of the effects of the other independent variables. The degree to which two or more independent variables are linearly related to the dependent variable is indicated by the correlation coefficient R, which is a number between -1 and 1 (if R= -1 then perfect negative correlation, if R=1 then perfect positive correlation). The strength of the relationship between the variables is expressed by squaring the size of the correlation coefficient. The resulting statistic is known as variance explained, or R^2, which can take a value between 0 and 1. The R^2 value gives the proportion of the variance in one variable that can be explained or predicted by the other variable.

In order to obtain R^2, we will employ a univariate multiple regression analysis, which is legitimate if the multivariate data do not depart significantly from sphericity. The statistics from the multivariate sphericity test on the Korean d' scores for obstruent clusters are given in Table 6.6. The null hypothesis states that sphericity holds. The test is not significant (p= 0.15) thus we cannot reject the null hypothesis. That is, the sphericity assumption is not violated, and, therefore, we can proceed with employing a univariate multiple regression analysis of this data in addition to the repeated measures analysis discussed above. Table 6.7 provides the results of a stepwise multiple regression analysis.

Table 6.6: Sphericity Test

Mauchly Criterion	0.20
ChiSquare	34.40
df	27
Prob>ChiSquare	0.15

Table 6.7: Sequence of hierarchically included models from the General Linear Model statistical analysis.

Step	Parameter	Sig. Prob	Seq SS	R-square	Cp	p
1	*strident*	0.0000	212.9635	0.5079	52.932	2
2	*nasal*	0.0000	24.2208	0.5657	25.71	3
3	*strident*nasal*	0.0023	8.454272	0.5858	17.509	4
4	*voice*	0.0028	7.777568	0.6044	10.126	5
5	*strident*voice*	0.0378	3.655808	0.6131	7.7148	6
6	*voice*nasal*	0.0605	2.942738	0.6201	6.1644	7
7	*strident*voice*nasa l*	0.6856	0.136242	0.6205	8	8
(8)	*(subject (random effect))*	-	-	(0.6991)	-	-

The important number that needs to be obtained from Table 6.7 is the difference in R^2 between hierarchically included models. The R^2 reveals the proportion of variance covered by the model (i.e., the statistical effect size of that model). Step (1) in Table 6.6 gives us the R^2 for the feature [strident]. Since it is the only feature included in the model in Step (1), the value corresponding to R^2 in this row is equal to the effect size of [strident] (i.e., 0.5079). Since the stepwise regression embeds models as it moves down the columns, the difference between Step (2) and Step (1) gives the additional R^2 covered for [nasal] given [strident] (i.e., 0.5657-0.5079 = 0.0578). Step (3) adds an interaction term for [strident] x [nasal], and the difference between Step (4) and Step (3) gives the addition of a term for [voice] to the model (i.e., 0.6044-0.5858 = 0.0186). It should be noted that Step (8) is added to the table from a separate analysis on the full mixed-effects univariate model that included the subjects as a random effect variable. Accordingly, the R^2 for the full model is 0.6991, accounting for about 70% of variance in the data. The R^2 for subjects (i.e., the variance observed between subjects) is the difference between Step (8) and Step (7) (i.e., 0.6991- 0.6205= 0.0786). Table 6.8 below summarizes the R^2 for each individual feature in the complete model.

Table 6.8: The effect size (R^2) for each model

Model	R^2
Strident	0.51
Nasal	0.06
Voice	0.02
Subjects	0.08

The feature [strident] accounts for more than 50% of the variance in the data. Compared to the features [nasal] and [voice], which account for 6% and 2% of the variation, respectively, that feature is an order of magnitude more important than any other in covering the observed variance. Interestingly, individual differences among subjects account for an additional 8% of the variance, suggesting that there is more between-subjects variance than the variance accounted for by differences in [nasal]. In conclusion, the feature [strident] has a vastly larger effect than [nasal] and [voice]. The multiple regression analysis suggests that a model with [strident] but without effects for [nasal] or [voice] may be adequate statistically. In the next section, another estimate of effect size is obtained by trying to measure the d' cost of each feature.

6.8 Effect Size: Cost Analysis

The effect size analysis using a multiple regression analysis shows that voicing and nasality effects, compared to stridency, account for much smaller amounts of observed variance in d' scores. While the regression analysis provides the proportion of variance covered by each feature in the data, we can also look at the cost of the same features for individual subject in units of d'. The question is how comparable the cost of voicing and nasality is to the cost of stridency for the Korean group, and how these costs compare to those for the English group.

The cost analysis for [voice], [strident], and [nasal] is carried out in the following way. For the feature [voice], the difference between the mean d' score for each consonant cluster where C_1 is a voiced obstruent (e.g., [gt^h]) and the mean d' score for its corresponding cluster where C_1 is a voiceless obstruent (i.e., [kt^h]) is calculated for each subject (i.e., [kt^h]-[gt^h]; [km]-[gm]; [ct^h]-[jt^h]; [cm]-[jm]). The numbers obtained for each difference are then averaged, giving the mean cost of voicing for that subject. As for the feature [strident], the difference between the mean d' score for each consonant cluster where C_1 is a strident obstruent and the mean d' score for its corresponding consonant cluster where C_1 is a non-strident obstruent is calculated for each subject (i.e., [kt^h]-[ct^h]; [km]-[cm]; [gt^h]-[jt^h]; [gm]-[jm]). The numbers from this calculation are averaged, yielding the mean cost of stridency for that subject. The same procedure is carried out for the cost of nasality; however, this time, the difference between the mean d' score for each consonant cluster is based on C_2. Specifically, the difference between the mean d' score on a consonant cluster where C_2 is nasal and the corresponding consonant cluster where C_2 is oral is calculated for each subject (i.e., [kt^h]-[km]; [gt^h]-[gm]; [ct^h]-[cm]; [jt^h]-[jm]; [lt^h]-[ln]). Averaging these numbers gives the mean cost of nasality for that subject. Using the individual mean scores, the group mean cost of each feature for each language group is calculated. Appendix C provides the cost of each feature for each subject in the present study.

Table 6.9 gives the mean cost of [voice], [strident], and [nasal] in units of d' and the statistics from a one-sample t-test procedure that was carried out to see if the cost of each feature is statistically different from 0. If the cost is statistically different from 0, then we can conclude that there is a cost in d' to the presence of the feature.

Table 6.9: The mean cost of features and one-sample t-test statistics

	Test Value= 0; p<0.005					
	English			Korean		
Feature	**Mean**	**t**	**Sig (2-tailed)**	**Mean**	**t**	**Sig (2-tailed)**
Voice	0.19	3.146	.004*	0.39	4.133	<.0005*
Strident	0.05	.508	.616	2.06	12.800	<.0005*
Nasal	0.30	4.854	<.0005*	0.53	5.718	<.0005*

An overall alpha level of $1-(1-0.005)^6=0.03$ is used to correct for the multiple t-tests with a Bonferroni method. Accordingly, the one-sample t-test statistics reveal that the mean cost of any feature for both groups, except for [strident] in the English group, are quite unlikely to be 0. Therefore, we can assume that [strident] did not cost any units of d' for the English group while [voice] and [nasal] decreased the mean d' scores by 0.19 and 0.30 units of d', respectively, in the same group. In the Korean group, the latter two features decreased the mean d' scores by about 0.39 and 0.53 units of d'. The feature [strident], however, dragged the Korean group's d' scores down by about 2.06 units of d' in average. Obviously, this cost is very high compared to the cost of [voice] and [nasality] for the same group. The question remains as to whether the cost of these features is different between the two language groups.

A 2x3 (language by features ([voice], [strident] and [nasal])) repeated measures ANOVA performed on the mean cost of the features using an alpha level of .05 reveals that there is an overall effect of feature ($F(2, 47)=18.310$, $p<.0005$). Furthermore, a two-way interaction emerges for language by feature ($F(2, 47)=29.984$, p<.0005) suggesting that the mean cost of features varied as a function of the feature and the language group. Given that the score for Korean [strident] is very high, these interactions are not surprising. Table 6.10 below reports the means, standard error and 95% confidence intervals for features and for both language groups. Figure 6.10 graphically illustrates the mean cost of each feature with the positive error bars showing the upper bounds with the 95% confidence interval.

Table 6.10: Mean cost of features in units of d', standard errors, lower and upper bounds within 95% confidence rate

				95% Confidence Level	
Language	**Feature**	**Mean**	**Std. Error**	**Lower Bound**	**Higher Bound**
English	**Voice**	0.194	0.080	0.03	0.355
	Strident	0.05	0.133	-0.218	0.317
	Nasal	0.302	0.079	0.143	0.460
Korean	**Voice**	0.394	0.080	0.233	0.556
	Strident	2.063	0.133	1.795	2.330
	Nasal	0.529	0.079	0.370	0.687

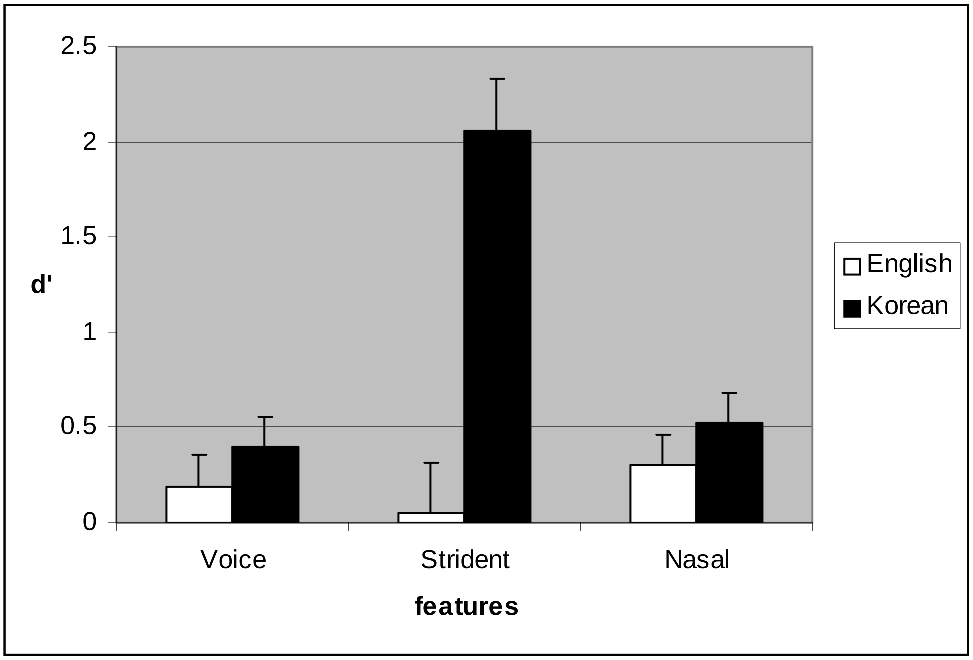

Figure 6.10: Mean cost of features and error bars with the 95% Confidence Interval

Figure 6.10 above shows that, although the cost of [voice] and the cost of [nasal] are higher on average for the Korean group than for the English group, the differences are not significant given the overlapping bounds for the 95% Confidence Interval. A post-hoc analysis using a Scheffe test with an alpha level of .05 further confirms the fact that the cost of [voice] for the English group is not statistically different from the cost of [voice] for the Korean group (p=0.850). Similarly, the same test also shows that the cost of [nasal] is not significantly different for both groups (p=0.768).

While both the English and the Korean groups are similar in the cost of the features [nasal] and [voice], the opposite is true for the feature [strident]. The effect of [strident] is the most for the Korean group but the least for the English group. The cost of [strident] for the Korean group is indeed significantly higher than all other features in the same group as well as in the English group.

These statistics suggest that both groups suffer some decrease in d' scores due to voicing and nasality, and the cost of these features in both groups are not different from one another. This analysis confirms the results from the regression analysis regarding the minute effect sizes for [voice] and [nasal] compared to a vastly larger effect for [strident]. Furthermore, it shows that the lower performance on [gm] and [gt^h], compared to [km] and [kt^h] in the Korean group is not likely to be due to a language-specific effect. Although voicing is phonemic, and voiced consonants are allowed in coda position, in English, the presence of [voice] in a consonant cluster also affects the English mean d' scores in the same way as it does the Korean mean d' scores. If we assume a phonetic processing strategy, we would expect different performance between the two languages for clusters with voiced C_1s. However, voicing is equally problematic for both groups.

Likewise, the presence of [nasal] in a cluster causes a similar decrease in performance for both groups. Despite the fact that English does not have any restrictions regarding C+N sequences, the English group also experiences some difficulty in the discrimination of such sequences due to the presence of the feature [nasal]. Therefore, the degradation of Korean d' scores on these clusters should not be attributed to a structural reason that is specific to Korean phonology. Hence, any claim that associates the Korean group's degraded performance on [gm] and [km] to an effect of consonantal contact would not explain the English group's performance on the same clusters, and is, therefore, a less preferable explanation. A better overall explanation is that there is a small universal cost to the presence of the features [nasal] and [voice]. We can then also conclude that the d' of 2.53 exhibited by the English group for [gm] represents good phonological performance on the contrast.

As for stridency, however, the significantly higher cost of stridency for the Korean group can simply be explained based on a structural difference that exists between the two languages. In Korean, the feature [strident] is not allowed in coda position, therefore the cost of stridency is expected to be higher for the Korean than for the English group, which does not have such a restriction. The Cost Analysis highlights the difference in the cost of stridency compared to the other features.

More importantly, the Cost Analysis leads us to a new hypothesis about the root of the perceptual problem with the feature [nasal] and [voice]. The fact of the matter is that both voicing and nasality pose similar degrees of difficulty for both groups in perception. The best working hypothesis is that the perceptual difficulty is due to a small universal effect of voicing and nasality. This speculation finds evidence especially in the case of [gm] where both features are converged and both groups' performance showed degradation. We can further speculate that the effect could be purely due to an increase in the acoustic complexity of the speech signal. Voiced obstruents combine periodic signals with noise signals, and these signals may be auditorially more complicated, or they may universally degrade the detection of components in the signal. Likewise, nasal sounds introduce pole/zero pairs into the spectra, resulting in more complex acoustic signals, perhaps with concomitant degradation of the detection of auditory components. Although this speculation cannot be empirically tested in the present work, it makes clear predictions for other languages, and could easily be tested in future studies.

6.9 The Perception of the Likely Output Forms of Illicit Clusters

It should be recalled that the experiment also includes a number of doublets which compare some of the test clusters, namely [km] and [ln], with their likely output forms in Korean production grammar, namely [ŋm] and [ll]/[nn], respectively. One of the reasons that these comparisons are included in the experiment is to guard the conclusions on perceptual epenthesis from the effects of other types of phonological adjustments on consonant clusters in Korean, that is, nasalization and lateralization. The mean d' scores for these comparisons are given in Table 6.11 below. The d' scores are all relatively similar and have very high means in both the English and the Korean groups.

Table 6.11: Korean mean d' scores for the discrimination of test clusters with their likely output forms

	mean d' score	
Test Doublet	**Korean**	**English**
p^hakma vs. p^haŋma	3.73	3.69
p^halna vs. p^halla	3.56	3.74
p^halna vs. p^hanna	3.48	3.75

These high d' scores indicate that the Korean listeners do not confuse the illicit consonant clusters with their likely output forms in Korean. Consequently, we can conclude that not all phonological processes are relevant for the Korean listeners' perception. This is an important finding that provides evidence against the P-map hypothesis (Steriade, 2001a). As will be discussed in Chapter 7, the P-map hypothesis predicts perceptual similarity between underlying forms and output forms that are derived from them. The difference between [km] and [ŋm], as well as [ln] and [ll], is very noticeable to Korean listeners, yet the alteration from these input forms to output forms is made in production anyway.

6.10 Summary and Conclusions

The results from the present study are summarized in the following:

(1) All the consonant clusters are successfully discriminated from their vowel-present counterparts by the English group, with d' values typically over 3 (average d'=3.38).

(2) Extremely low d' values emerge in the Korean group for the consonant clusters [ct^h], [jt^h], [jm], [cm] (average d'=0.53).

(3) Very high d' values for [lt^h], [ln] emerge in the Korean group, which are not statistically different from each other and from the same consonant clusters in the English group. This is inconsistent with the Consonantal Contact Hypothesis.

(4) Intermediate d' values for [km], [gt^h], [gm] emerge in the Korean group, suggesting that voicing and nasality add some problems for perception. The multiple regression analysis and the Cost Analysis, however, consistently show that the effect of the feature [nasal] and [voice] is very small compared to the effect of [strident]. Furthermore, both [voice] and [nasal] result in similar effects for both the English and the Korean group, suggesting possibly an auditory effect caused by the acoustic properties of these features.

(5) The Cluster Analysis into two groups divides the data into two different levels of discriminability: (1) lack of discriminability on the clusters [cm], [ct^h], [jm], [jt^h] and [km], and (2) successful discrimination on the clusters [gm], [kt^h],

[gt^h], [ln] and [lt^h]. This clustering can only be explained if we assume the predictions of the Coda/Onset Identity Hypothesis in combination with the Phonological Processing Hypothesis.

(6) The Korean listeners do not confuse the illicit consonant clusters [km] and [ln] with their likely output forms [ŋm] and [ll], respectively, in Korean.

Taken together, these results indicate that a given consonantal sequence C_1C_2 containing the consonants [k] or [l] as C_1 is distinguishable from its epenthetically adjusted counterpart C_1VC_2 for the Korean group regardless of whether C_2 is a plosive or a nasal consonant. When the same sequence contains an illicit coda consonant in the C_1, however, the cluster becomes indiscriminable for the Korean group from its epenthetically adjusted counterpart.

The Coda/Onset Identity Hypothesis predicts that if each member of a consonant cluster satisfies the syllable structure conditions, this cluster does not pose perceptual problems for the Korean group. In the experiment, the clusters [km] and [ln] test this hypothesis, where both [k] and [l] are possible coda consonants in Korean. Even though these consonants are followed by nasal consonants, which cause a consonantal contact violation, the Korean group successfully discriminates these clusters. Consonantal contact restrictions, therefore, do not explain the Korean group's performance in the experiment. Rather, the response patterns can only be explained if we assume that the L1 syllable structure violations, rather than consonantal contact violations, play an important factor in the perception of consonant clusters.

Finally, the voiced coda consonant [g] does not cause perceptual epenthesis in the Korean group. This suggests that voicing information is suppressed in the perception of the clusters that contain this segment and indicates phonological processing of features based on their abstract underspecified representations in the L1 of the listener.

Chapter 7

CONSEQUENCES AND IMPLICATIONS

7.1. Introduction

This chapter recapitulates the results of the present study and examines their various implications for existing phonological theories, and models advanced in the field of speech perception. First, in light of the present findings, I address the idea that perception is not equivalent to production, showing that certain illicit consonant clusters in the Korean production grammar do not cause any perceptual problem for Korean listeners. I argue that these findings provide evidence against accounts of segmental phonological alternations based on context-dependent perceptibility, such as the P-Map and Licensing by Cue hypotheses (Steriade, 1999, 2001 a, b). In addition, analyses that are purely based on familiarity or frequency of the segments and segment combinations likewise cannot explain the outcome of the present study. Instead, I argue that perceptual epenthesis is phonological and syllable dependent and thus analyses must refer to the syllable structure in order to account for the findings reported here. Last but not least, I propose a perceptual algorithm that employs syllabically conditioned detectors. The generalizations about perceptual epenthesis are given in syllabic terms and their implications are discussed. Finally, directions for future research are discussed in light of the findings and implications of this dissertation.

7.2. Overview

In this dissertation, I have motivated two competing hypotheses based on two different views on phonotactics. These hypotheses make different predictions about how L1 phonotactic knowledge influences the perception of L2 consonant clusters. The Consonantal Contact Hypothesis is formulated following a string-based approach towards phonotactics, which predicts perceptual epenthesis as a result of perceiving consonantal sequences that are illicit in the L1 of listeners. Contrasting with this hypothesis, the Coda/Onset Identity Hypothesis assumes that perceptual epenthesis arises if the illicit consonantal sequence violates the L1 syllable structure conditions. Korean provides the appropriate contexts to test these two hypotheses. It bans certain consonants to surface in coda position (e.g., *[c]) while allowing others (e.g., [k], [l]). Korean also imposes restrictions on certain other heterosyllabic contacts (e.g., *[ln] and *[km]) although the first member of the cluster can be a possible coda (i.e., [l] and [k] are permissible in coda position).

A secondary inquiry in the dissertation addresses the way L2 input forms are represented in the course of speech perception. I have postulated another pair of hypotheses that make conflicting predictions regarding the detail of L2 representations. The Phonetic Processing hypothesis claims that the gradient acoustic information in the speech signal is mapped on to corresponding phonetic categories in the L1. Accordingly, the phonetic information (including allophonic information) in the acoustic signal is represented on the basis of the corresponding feature in the L1 of the listener. For example, listeners who have the feature [voice] in their L1, regardless of its phonemic status, will represent that feature in the L2. The Phonological Processing Hypothesis, on the other hand, claims that the representation of L2 forms is constructed using only the contrastive features of the L1. Accordingly, the predictable and redundant information in the L1 phonology is also suppressed in L2 representations.

To test these two hypotheses, the first member (C_1) of the consonantal sequences (C_1C_2) employed in the experiment varied with regard to the feature [voice], which is non-contrastive in Korean but varies allophonically in a rule-governed way: plain stops become voiced intersonorantly. The consonants [j] and [g], the voiced allophones of /c/ and /k/, respectively, in Korean, were combined with the following nasal and oral stop onsets (C_2) to yield yet another set of consonantal sequences that are impossible on the surface in Korean: [jm], [jt], [gm] and [gt].

Furthermore, extra conditions were included in the experiment that contained sequences such as [ŋm], [ll] and [nn] to test if they are misperceived as [km] and [ln] by the Korean listeners because the earlier forms are the likely output forms of /km/ and /ln/ in Korean. The relevance of these pairs lies in an interesting question as to whether Korean production rules, which resolve illicit sequences in Korean such as /km/ and /ln/, also apply to Korean listeners' perception of the very same sequences.

As discussed in Chapter 6, three different levels of performance on the test clusters are observed: bad, intermediate and good. We take the English group's discriminability indices to indicate good performance. Only significantly worse performance is counted as bad. Accordingly, the Cluster Analysis divides the Korean group's responses to various consonant clusters into two main groups, as summarized in Table 7.1 below.

Table 7.1: Summary of Korean Results in the CC vs. CVC condition

Sequence	**Type**	**mean d'**
kt^h	OK	3.58
ln	Bad contact	3.42
lt^h	OK	3.28
gt^h	Bad contact	2.72
km	Bad contact	2.28
gm	Bad contact	1.80
=================================		
ct^h	Bad coda	0.89
jm	Bad coda	0.48
jt^h	Bad coda	0.47
cm	Bad coda	0.30

Note: (The line (==) denotes the separation determined by the Cluster Analysis into two clusters).

Various statistical tests and analyses carried out in Chapter 6 suggest that the discrimination of the clusters with strident consonants, namely [c] and [j], is significantly worse than those with [g], [k], and [l] regardless of the legitimacy of the consonant cluster in Korean. These results provide strong evidence against the view that advocated the role of sequential restrictions in perception: that is, they falsify the Consonantal Contact Hypothesis. Rather, the results are consistent with the view that language-specific syllable structure restrictions on coda consonants determine consonant clusters.

The present study reveals another important finding with regard to the way certain features are processed in the course of speech perception. Looking at Table 7.1 above, the Phonological Processing Hypothesis in combination with the COI Hypothesis predicts a difference between the two voiced consonants, based on the principle of the suspension of allophonic features in perception. In this case, since voicing is allophonic in Korean, the hypothesis predicts that it will be ignored in Korean speech perception. In particular, as [j] is thus mapped to /c/, it will still induce perceptual epenthesis because it is still strident, whereas [g], being mapped to /k/, will not cause epenthesis. The Phonetic Processing Hypothesis can be maintained with the present data if one assumes that voicing is represented, but there are no constraints banning voiced codas. That is, though the feature is present, no constraints actively refer to it. The present results offer evidence for a small effect of voicing, which can be viewed successfully as a universal effect, as the effect was of a similar size for both English and Korean listeners. Therefore, we have no evidence that voicing is attended to in the course of Korean perception, and these results are consistent with a view of speech perception that assumes that mental representations of speech segments are underspecified for their features that are predictable and redundant.

As discussed in Chapter 4, the successful perception of voiced segments in coda position by Korean listeners can be attributed to the lack of an explicit rule that specifically bans voiced segments from coda position in Korean. This is particularly due to the fact that the exclusion of voiced segments from coda position in Korean is epiphenomenal: voiced segments happen to not surface in coda position (see Section 4.2.3.2 for more information). Having confirmed that consonantal contact restrictions are irrelevant for the present results, this explanation would then predict that only those features that are explicitly banned from coda position in Korean, such as stridents, would cause perceptual epenthesis. However, as discussed earlier, this approach calls for substantial modifications in the current frameworks within Optimality Theory, where constraints violated by an observed grammatical structure is demoted in the hierarchy so that they are dominated by the constraints violated by competing structures (e.g., the Biased Constraint Demotion (BCD) algorithm (Prince and Tesar, 1999; Tesar, 2002). Accordingly, there is no piece of positive evidence, a grammatical structural description, in Korean that would violate a constraint banning voiced segments from coda position. There is no doubt that such a constraint is a universal constraint given the existence of many languages that have coda devoicing. Since OT assumes that constraints are universal, an equally valid argumentation here would be that such a constraint must be undominated in Korean. Furthermore, the crucial difference between the alternative explanation and the Phonological Processing Hypothesis is that while the latter is reflective of the difference between the status of voiced and voiceless segments in the minds of Korean listeners, the earlier overlooks this important generalization. The Phonological Processing Hypothesis builds the very same facts based on a causal relationship between Korean listeners' knowledge of what linguistic function the feature [voice] carries in their L1 and how their representations are reflective of that knowledge. In the absence of any further evidence that can make the alternative explanation more advantageous, I adhere to the Phonological Processing Hypothesis, which is conceptually more insightful.

7.3. <u>Speech Perception is not Isomorphic to Speech Production</u>

One of the most crucial findings of the present study is that not all illicit consonant clusters behave the same. Specifically, Korean listeners successfully discriminate illicit consonantal sequences such as /km/ and /ln/ from their likely misperceptions [ŋm], [ll] or [nn], respectively, although /km/→ [ŋm] and /ln/→ [ll] alternations are all mandatory in Korean production (see section 3.2 in Chapter 3). Although the successive occurrences of [k] plus [m], and [l] plus [n] are banned on the surface and Korean employs assimilatory processes to accommodate these segments when they occur in underlying forms, the processes of assimilation do not play the same role as epenthesis in perception. Korean listeners seem to be content with perceptual representations that are impossible in their L1.

In the case of similarly impossible [strident] codas, however, Korean listeners *do* perceptually alter their input. However, the perceptual strategy they employ is not in line with the way this same feature is accommodated in the production grammar. The different mechanisms employed by the perceptual and production systems in Korean are schematically shown in Figure 7.1.

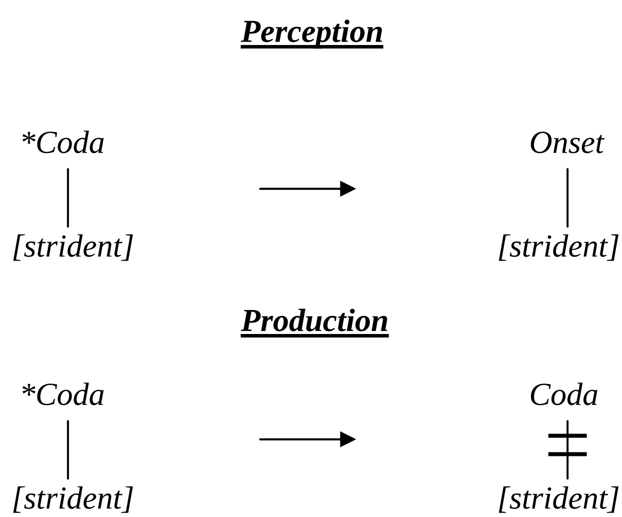

Figure 7.1: Korean perception vs. production grammar[19]

As shown in Figure 7.1, Korean listeners perceive coda [strident] sounds as being in the onset position via epenthesis although in production Korean simply

---

[19] It should be noted that the status of the assumed intermediate representation is different in these two illustrations. While in production, the feature [strident] is in a coda and then undergoes change, the same feature is never in a coda position in perception. Rather, the original stimulus, that is, the English production, has the feature [strident] linked to a coda.

delinks the feature [strident] from the coda position (e.g., /c/ →[t]). The perceptual system does not appear to use any repair strategies that mirror assimilatory processes that are active in the production grammar. Having empirically shown that perceptual phenomena are not simple inversions of phonological rules or constraints used by the production system, the findings of the present study are difficult to interpret in models that incorporate listeners' knowledge of perceptibility of sound contrasts to predict phonological alterations. This view is an essential part of Steriade's model (e.g., Steriade 1997, 1999, 2001a, b), which argues that phonetics drives much of phonology and that context-dependent phonetic cues determine phonological contrasts.

According to Steriade, there is a mechanism, which she calls to be the P(erceptual)-map, that relates rankings between correspondence constraints to perceived differences in degrees of similarity. As I will argue below, the perceived degree of similarity should be inherently related to the discriminability index attained by the Signal Detection Theory. The P-map idea has its roots in Steriade's (1999) Licensing by Cue hypothesis, which states that "the likelihood that distinctive values of the F-contrast will occur in a given context is a function of the relative perceptibility of the F-contrast in that context" (p. 4). Accordingly, the probability that a consonant deletes, triggers epenthesis, or blocks vowel deletion correlates with the quality and quantity of the auditory/ phonetic cues associated with it in a given context.

Steriade (2001b) employs the P-map to provide a synchronic explanation for directional asymmetries in phonological processes. She defines the P-map as "the repository of speakers' knowledge, rooted in observation and inference, that certain contrasts are more discriminable than others, and that the same contrast is more salient in some positions than others" (p. 236). The P-map attempts to explain, for instance, why illicit consonantal clusters are more likely to be repaired by altering the postvocalic consonant than by altering the prevocalic consonant; it claims that postvocalic consonants have a lower perceptibility.

In essence, the P-map is employed by speakers to determine when articulatory alterations and simplifications can be carried out without the listener noticing a deviation from accepted norms of pronunciation. Hence, the P-map is considered to be a set of statements, each of which assigns a similarity value to a perceived difference between two sound strings, that surface in a specific context without regard for their phonemic status in the language.

Since the P-map assumes that perception affects phonology directly, a strong correlation is predicted to hold between the way heterosyllabic consonant clusters are perceived and the way Korean phonology constrains its realizations of them. As the present study showed, there is no evidence for a direct and clear link between the Korean perceptual results on consonant clusters and their realization in the Korean production grammar. The missing link between perception and production in the present findings needs to be discussed with specific reference to Steriade's formulation of the likelihood of assimilation based on perceptual grounds.

(1) Perceptual similarity to input (Steriade, 2001b, p. 222):

"The likelihood that a lexical representation R will be realized as R' is a function of the perceived similarity between R and R'."

According to (1), an underlying phoneme /A/ surfaces as [B] rather than *[C] because perceptually [B] is more similar to /A/ than [C] is to /A/. The reference to perceived similarity in (1), according to Steriade (2001b), conveys "that perceptual factors -among them cue distribution- play a critical role in defining degrees of similarity between lexical forms and their conceivable modifications" (p. 222).

By definition, the degree of similarity between any two features should be related to their discriminability, and thus it can be empirically tested. As discussed in Chapter 5, the SDT analysis computes a discriminability value by comparing the probability of "same" responses when listeners are presented with A vs. B and B vs. A (i.e., the false alarm rate) to A vs. A and B vs. B (the correct rejection rate). This is illustrated in Table 7.2 below using the two experimental items in question.

Table 7.2: An example SDT analysis scheme ([p^hacma] vs. [p^hac^{h}ima])

Trial: ([p^hácma]-[p^hácima])	**RESPONSE**	
	DIFFERENT	**SAME**
[p^hácma]-[p^hác^{h}ima] [p^hác^{h}ima]-[p^hácma]	***Hits***	***Misses***
[p^hác^{h}ima]-[p^hác^{h}ima] [p^hácma]-[p^hácma]	***False Alarms***	***Correct Rejections***

Therefore, "perceived similarity" can be calculated with reference to the index of discriminability (d') such that discriminability must be the inverse of similarity. Accordingly, in the context of Korean lateralization, the likelihood that /…ln…/ will be realized as […ll…] rather than something else must be a function of the *perceived similarity* between /…ln…/ and the assimilated variant […ll…]. Thus, (1) can be mathematically reformulated as (2)[20].

(2)

$$P([\ldots ll\ldots] \mid /\ldots ln\ldots/) \sim \frac{1}{10^{d'([\ldots ll\ldots],\, [\ldots ln\ldots])}}$$

[20] Since division by 0 is undefined (i.e., if d'=0, the probability must be 1), 10 is used as the base of the logarithm for convenience. Naturally, any number can be used instead. The values will vary only by the constant factor (i.e., $\log_{10}(x) = \log_2(10)*\log_2(x)$).

The formula in (2) states that the probability of saying [ll] given /ln/ is proportional to the reciprocal of the discriminability score on [ll] vs. [ln]. Since discriminability is the inverse of similarity, as the d' value approaches 0, the probability approaches 1. Conversely, when d' approaches infinity, the probability approaches 0. Therefore, the probability that /…ln…/ will be realized as […ll…] increases as the discriminability score (d') attained on the items that contains these sequences gets close to zero.

The data in the present study showed that the Korean group's d' score on [p^halna] vs. [p^halla] is 3.56. According to (2), if d' ([palna], [palla]) > 0 then p([…ll…] | /…ln…/)≈0. That is, since the average d' score on [ln] vs. [ll] (or vice versa) is 3.56 in the present study, the /ln/→[ll] alteration in Korean production is not predicted since the probability of that alternation is close to zero. That is, /ln/→[ll] alteration in Korean cannot be explained by Steriade's Perceptual Similarity to Input idea in (1).

Similarly, the same statement in (1) predicts that Korean speakers' discrimination index on [km] vs. [ŋm] must be close to 0 in order to account for the /km/ → [ŋm] alteration in Korean production. However, the Korean listeners' average d' score on [p^hakma] vs. [p^haŋma] was 3.73. By the equation in (4), a d' score that is as high as 3.73 results in a probability score that is close to 0, which contradicts Steriade's statement in (1).

In the case of Korean neutralization of strident sounds in coda position, the likelihood that, /c/ will be realized as [t] rather than something else must be a function of the perceived similarity between /c/ and [t] in the coda position. That is, it must be the case that d'([p^hac], [p^hat])≈0 in Korean. In this case, the present results show that d'([p^hac.ma], [p^ha.c^hi.ma])≈0. The same is also true for [p^hac.t^ha] vs. [p^ha.c^hi.t^ha]. That is, the study has convincingly showed that a coda [c] is indistinguishable from [c^hi], not from anything else. It should be noted that the present study did not compare, for instance, [p^hacma] and [p^hamma], the expected surface realization of a potential underlying form /p^hacma/. However, there should be no doubt that [c] is distinguishable from [m]. First, anecdotally, Korean listeners do not have any trouble distinguishing these two. In addition, Korean speakers never produce loan words that contain such clusters, such as [cm] in *Richmond*, as having [mm]. Furthermore, given that the Korean listeners' ability to perceive the difference between [km] and [ŋm], they should also be able to distinguish [c] from [m] without any problems. Thus, on conceptual grounds, we can say that the neutralization of strident segments in coda position do not follow from Korean listeners' perceiving strident segments such as [c] as [t].

In short, if assimilatory patterns in consonant clusters were really predicted by an index of perceived similarity between the assimilated variant and the underlying representation, Korean listeners should confuse those clusters like [km] with their likely output forms. However, the present study shows that neither [km] vs. [ŋm] nor

[ln] vs. [ll] are confused. Furthermore, on those clusters where confusion arises, the Korean phonology should be expected to exhibit the same direction of assimilation. For instance, [strident] codas were indistinguishable from [strident] onsets; therefore, the P-map hypothesis would have to say that [strident] codas should be realized via epenthesis in the Korean production grammar. However, no such alternation is attested in the synchronic phonology of Korean.

Hence, the Perceptual Similarity to Input hypothesis does not have the means to account for why the present study found no perceptual repair strategies for coda [k] and [l] in the illicit clusters on the one hand, but did for coda [c] on the other hand. If, as Steriade (2001a, b) suggests, speakers' knowledge of the perceptibility of contrasts influences phonology, then there must be physical differences in the way /c/, /k/, and /l/ are realized in the coda position that cause the differing perceptual behavior on the pairs that contained these consonants. Accordingly, the P-map hypothesis (as well as all the sub-hypotheses that are assumed in this theory such as the Perceptual Similarity to Input Hypothesis and Licensing by Cue Hypothesis) attributes the susceptibility of a certain set of consonants to change in a particular position to their articulatory and acoustic characteristics that determine their relative perceptibility in that position. Accordingly, the hypothesis would predict that the acoustic cues that correspond to [c] are perceptually less salient word-finally than those of [k] and [l] because neither /k/ nor /l/ necessarily undergoes a change but /c/ does in this position. Furthermore, in the case where the very same consonants occur in heterosyllabic consonant clusters (e.g., /c.m/, or /k.m/), the P-map hypothesis would have to show that /c/ is more similar to [c^hi] than any other potential alteration, attested or unattested, in the coda position (e.g., the deletion of /c/, or manner of articulation change such as /c/→ [t], etc.). However, at the same time, it would also predict that the same does not hold for /k/ vs. [kɯ] or /l/ vs. [lɯ]. However, there is no reason to assume that the acoustic information pertaining to [c] as opposed to those for [k] and [l] is less perceptible in the coda position. On the contrary, the feature is so easy to hear, in fact, that the Korean listeners retain it in perception and perceive an illusory vowel to accommodate it.

7.4. Frequency Information and Perceptual Illusion

The results of the study constitute an important piece of evidence against views that attribute perceptual preference for certain consonant clusters to listeners' differing experience with them and the frequency with which those clusters occur in the language. Probabilistic phonotactics refers to the frequencies of segments and sequences of segments in syllables and words (Trask, 1996; Vitevitch and Luce, 1998; 1999). As discussed in Chapter 2, recent research has suggested that listeners are sensitive to probabilistic differences among phonotactic patterns (e.g., Vitevitch, Luce, Charles-Luce, and Kemmerer, 1997). Models, such as the TRACE model of word recognition (e.g., McClelland & Elman, 1986) does not rely on explicitly stored knowledge about the phonology of the language but rather employs a frequency-sensitive mechanism to recognize words. Accordingly, the bias towards perceiving legal sequences emerges from the fact that there are more lexical entries with legal than illegal sequences. That is, sensitivity towards frequency replaces what seems to be the function of the phonological knowledge about the sound distributions that the listener possesses.

Looking at the present study, if epenthesis were a means by which the perceptual system biases processing of clusters that have zero frequency, then, all the illicit consonant clusters in the present study would be more susceptible to epenthesis. Present findings, however, show that only on a subset of the illicit sequences, namely the ones where a strident sound is a coda, do the Korean groups' performance reflect near-indiscriminability although all the illicit consonantal sequences have zero frequency of occurrence in Korean production. For instance, although "a sequence of l and n in either order is absolutely not permissible on the surface in Korean, and a complete assimilation of n to l occurs whenever such a sequence arises" (Kim-Renaud, 1995, p. 223), the consonantal sequence [ln] was not susceptible to perceptual epenthesis in Korean listeners. These results suggest a phonological influence of L1 phonotactic knowledge, rather than an effect of frequency, on L2 speech perception and disconfirms approaches that attribute listeners' perceptual preferences for certain consonant clusters over others to the unattestedness or the rarity of the dispreferred cluster in the L1 of the listener (e.g., Hallé et al., 1998; Pitt, 1998; Pitt& McQueen, 1998).

Recently, Moreton (2002) claimed that a perceptual bias exists against clusters with more structural violations than those with fewer violations. In his study, Moreton controls for frequency separately from structural differences. For instance, according to Moreton, [dl] onset clusters incur more structural violations than [bw] onset clusters in English. Although both [dl] and [bw] violate a constraint against having onset clusters that share the same place of articulation, the consonants in [bw] are more distant from one another on the sonority scale than the consonants in [dl]. Moreton claims that English allows consonant clusters with the same place of articulation when the consonants in the cluster come from opposite extremes of the sonority scale (e.g., [dr] with the combination of [coronal] plus [coronal], and [gw] with [dorsal] plus [dorsal]). Accordingly, the ban on [dl] is stronger than [bw] since [d] is closer to [l] in sonority than [b] is to [w]. Moreton also claims that both of these clusters have zero frequency of occurrence in English, compared to [bl] and [dw], as summarized in Table 7.3 below.

Table 7.3: Frequency of occurrence and structural well-formedness of stop-consonant clusters (adapted from Moreton (2002))

Cluster	Frequency (by token)[21]	Structure
[bw]	zero	marginal
[bl]	389 (word-initial) 890 (syllable-initial)	OK
[dw]	10 (word-initial) 16 (syllable-initial)	OK
[dl]	zero	impossible

Moreton makes the prediction that if perceptual differences are due to listeners' differing experience of onset clusters in the language, then there should not be any difference between [bw] and [dl] since they are both unattested in English. If structural differences are behind perceptual differences, there should be a perceptual bias against [dl] as compared to [bw].

To test these predictions, Moreton had English listeners judge synthetic arrays of stop-sonorant clusters ambiguous among arrays of [dlæ-dwæ] (d-array) and [blæ-bwæ] (b-array) syllables. To measure the stop and sonorant judgments separately, both consonants in the clusters varied incrementally between ambiguous endpoints. That is, the stop was ambiguous between [b] and [g] in the b-array, and [d] and [g] in

[21] Values represent occurrences in an 18.5-million-word London-Lund corpus of written and spoken British English (see Moreton, 2002: 58 for more information).

the d-array. Likewise, the second consonant was ambiguous between [l] and [w] in both cases. This enabled Moreton to measure the dependence of "l"/"w" judgments on "g"/"d" and "g"/"b" judgments in these syllables. Accordingly, upon hearing each syllable, the participants were told to press one of "dw, dl, gw, gl" or "bw, bl, gw, gl" depending on which array was presented. The results of the study showed that when, for instance, the stop was identified as "d", the sonorant was more likely to be called "w" than "l"; this consequently shows a perceptual bias against [dl] onsets, but not against [bw]. This provides evidence against models that attribute perceptual differences between consonant clusters to their frequency of occurrence in a given language.

The present findings go along with Moreton's (2002) conclusions. Having ruled out the consonantal contact violations, it is clearly the case that strident-clusters in the present study, compared to [km] and [ln], induce a syllable structure violation in Korean. This explains the perceptual epenthesis effects found only with the strident clusters although all three types of clusters have zero probability of occurrence in Korean.

It should be remembered that the mean d'-scores on [km] (2.28) are somewhat degraded compared to [ln] (3.42) for Korean listeners. Again, the behavioral difference cannot be directly linked to frequency effects. There are obvious physical differences between these clusters. While [ln] is a combination of two sonorant sounds, [km] is an obstruent followed by a sonorant. It is possible that some minute difference of sonority between the members of these clusters, *a là* Moreton (2002), is responsible for the disparity between the d' scores on these clusters. Heterosyllabic consonantal contact relations do not seem to seek a specific sonority profile in Korean. However, sonority restrictions that try to maximize distance between the members of the clusters typically apply for consonants in the same syllable. Since heterosyllabic clusters belong to different syllables, such structural restrictions do not apply to [k.m] and [l.n].

Even if the same principles were also true for heterosyllabic consonants, a sonority-based explanation would fail to account for somewhat degraded performance on [k.m] compared to [l.n]. This is because consonants of [l.n] are closer to each other in sonority than are those of [k.m]. Furthermore, while the consonants of [l.n] share the same place of articulation, the consonants of [km] do not. The sequence [k.m] should, therefore, be more preferable than [ln] if Korean has the same structural requirements as Moreton argues for in English.

Future studies should look further into the effects of features in the perception of heterosyllabic consonant clusters. In the present study, a clear pattern has emerged from the Cost Analysis, where voicing and nasality share a relatively small universal cost compared to the massive cost of the feature [strident] for Korean listeners. For instance, the poorer performance obtained for [gm] and [gt^h], compared to [km] and [kt^h], seems to find an explanation since the earlier clusters contain at least one of these features. The difference between [km] and [ln], both of which cause a contact violation in Korean, however, remains to be unknown. I will leave this issue to future studies, which should specifically test the effect of phonological features on the misperception of heterosyllabic consonant clusters.

7.5. <u>Onset/Coda Detection Theory</u>

The syllable has long been considered as the unit that is instrumental in speech processing and was assigned a ubiquitous role in psycholinguistic models (e.g., Mehler, Domergues, Frauenfelder, & Segui, 1981; Frauenfelder & Kearns, 1996; Mehler, Dupoux, & Segui, 1990; Cutler & Norris, 1988; Cutler, Demuth & McQueen, 2002). In phonological theory, the role of the syllable has been extensively discussed and arguments for it were derived from the existence of phonological processes which are sensitive to a domain that is larger than a segment, smaller than the word (see Blevins, 1995 for a review; though see Blevins, 2002 for counter-arguments).

How exactly are the syllables involved in speech segmentation so that the syllable-based positional restrictions can apply during speech perception? The vocalic segments that typically determine the nucleus position of each syllable can be considered to designate the left and right anchoring points of the syllable. Therefore, the margins of a syllable to the left and right of nuclei, namely, the onset and coda, can constitute reliable syllabic segmentation points in the speech signal.

The Coda/Onset hypothesis tested in the dissertation has incorporated the idea of syllabically conditioned epenthesis by considering the coda- and onset-hood of consonantal elements in a given language. It must be noted that certain consonants may be allowed to occur both in the coda and in the onset position in a language (e.g., [k] can be both coda and onset consonant in Korean as well as in English). For simplicity, consonantal features that are designated for a specific syllable margin can be thought of as having only one of the two possible unary syllable-based prosodic positions: [Onset] or [Coda], or no such specification. For example, the feature bundles corresponding to the English velar nasal [ŋ] and English [h] are only specified for [Coda] and [Onset], respectively. Those consonantal features that can occupy both of these positions, however, are simply not specified for such prosodic positions. That is, only those consonants that are specifically reserved for a specific position need to be marked for that position.

The prosodic positions [Onset] and [Coda] are obviously in complementary distribution in this approach. That is, if a consonantal feature is banned from coda position, it must be associated with an onset. Likewise, if it is banned from onset position, that feature must be a coda. Although the latter case is quite rare in that single member onsets tend to be unrestricted across languages, there are many languages where the segments that may occupy coda position are highly limited.

In Korean, strident phones (i.e, /c, c', c^h/) are banned from the coda position. The "non-strident" condition on Korean codas in perception is described in Figure 7.2.

IF [+strident]

|

THEN ONSET (i.e., σ[[+strident]…]

Figure 7.2: Korean coda constraint (perceptual)

The restriction in Korean described in Figure 7.2 above is employed predictively in perception in the following way: if a [strident] feature is present then it must be placed in onset position. I will show that the statements illustrated in Figure 7.2 can easily be adopted as an algorithm for detecting and accommodating Korean stridents (i.e., "onset-only" cases) in the speech signal. The algorithm developed here crucially employs syllabic detectors to check for one of the two syllable-based prosodic positions (i.e., [Onset] or [Coda]) on consonants. In developing this algorithm, I do not adopt a particular theory of syllable. Any theory that employs onsets and codas would go along with the predictions of the algorithm.

The perceptual algorithm proposed in this dissertation builds the syllabic representations of input forms with reference to universal tendencies pertaining to syllable structure (e.g., sonority) as well as language-specific restrictions on the syllable (coda restrictions, allowance for branching codas and onsets, etc.). The perceptual algorithm detects [Onset] and [Coda] information in the acoustic flow by the use of onset and coda detectors. These detectors look out for the acoustic correlates of phonological features that are associated with the prosodic positions in the language. As the algorithm builds the syllabic representation of the speech stream, it assigns [Onset] interpreted features to onset positions and [Coda] interpreted features to coda positions. For instance, in Korean onset detectors look out for the feature [strident] due to the restriction given in Figure 7.2; consequently, the algorithm associates the [Onset] interpreted feature with an onset position in the syllable structure. In a language that excludes obstruent sounds from syllable final positions, (e.g., Mandarin), for instance, the same detectors would look out for [-sonorant], which is linked to [Onset] in this language.

Having designated the feature [strident] as [Onset] via the constraint given in Figure 7.2 above in Korean, the perceptual algorithm faces two options: (1) start a new syllable by using the resources available, such as linking the onset to a following syllabic element (i.e., vowel) that will serve as the nucleus of the syllable, or (2) in the absence of vocalic resources, posit a nucleus position to form one. That is, unless there is a vowel immediately following the consonant in question, the parser will create a default vowel for the nucleus. Failure to do so will result in an ill-formed syllable because the [Onset]-interpreted feature would otherwise have to merge with a following consonantal feature, creating an illicit branching onset. Since onsets are non-branching in Korean, an epenthetic vowel emerges as the [Onset]-interpreted strident feature requires, by default, a nucleus to satisfy the basic syllable structure condition: syllables must have a nucleus. In the absence of a vocalic element in the physical speech stream, the parser posits an empty nucleus, which is understood as the vowel /i/ in the context of affricates.

In contrast, consonants such as /k/ and /m/ in Korean are not specified for [Onset] since they do not necessarily require onsets. Features that are not linked to any prosodic position simply float until they fill in excess slots in the syllabic template postulated by the parser. That is, if there is a vowel in the speech string following such features, an onset will become available and the features will then be linked to it. This can be seen as a consequence of whatever principle in UG maximizes onsets (e.g., the Universal Core Syllable Condition, Itô, 1986; 'MAX ONSET', Prince and Smolensky, 1993).

It must be noted that there are consonants, such as /s/ in Korean, that may seem to appear in coda position in Korean only before /s/ or /s'/ (e.g., /os.s'ok/ 'inside of clothes'). Such a sequence, however, constitutes a geminate structure, which should be distributed heterosyllabically between the coda and the onset. This is analogous to Japanese, where stop consonants only surface in coda position provided that they are also linked to an onset position (e.g., [kap.pa] 'a legendary being'). Likewise, /s/ in Korean cannot independently occur in coda position; therefore, the feature [strident] must be linked to an [Onset]. The coda portion in the syllable structure, however, is filled by the same feature if durational cues indicate that /s/ is a geminate.

The reason that the algorithm is build up on the assumption that certain features desire certain positions, rather than certain features are banned from a specific prosodic position, is precisely due to the cases where features might be divided between an onset and a coda due to gemination, as shown above. That is, if we simply said that the feature [strident] is banned from coda position, this would make incorrect predictions since the feature can indeed occupy a coda position. This, however, happens only when the same feature occupies the onset of the following syllable. Table 7.4 lays out Korean phonemes and the specific syllable positions they are allowed to occupy.

Table 7.4: The Onset/Coda specification of Korean consonants

Korean Segment	Onset	Coda	Specification
p, t, k, m, n	YES	YES	--
p', t', k', p^h, t^h, k^h, s, s', c, c', c^h, h, ɾ	YES	*	ONSET
l, ŋ	*	YES	CODA

When coda detectors find features associated with [Coda], the perceptual algorithm licenses them via adjunction to a preceding V, as a coda. As can be seen in Table 7.4, the Korean lateral liquid [l] is an example of a coda-only consonant. An underlying /l/ in Korean is realized as an alveodental lateral liquid [l] in the coda position (e.g., [yəl] 'ten'; [mal] 'horse'). If /l/ occurs in onset position, as in the case of /l/-final morphemes that are inflected with vowel- or glide-initial morphemes, the phoneme is realized as a flap (e.g., /yəl+ahop/ → [yə.ɾa.hop] ('ten'+'nine') 'nineteen', /mal+i/→ [maɾi] 'horse-Nominative', /mal+yək/→[ma.ɾyək] 'horsepower'). It seems, therefore, that [l], or rather the feature [lateral], constitutes a coda-only feature since [l] is restricted to coda positions. Interestingly, there are cases where [l] may also appear in the onset position provided that the segment is simultaneously associated with a coda position. This is the case where [l] is a geminate (e.g., [mul.li] 'physics'). This is the reverse of /s/ in Korean, which is an onset-only consonant due to the feature [strident], but which may occur in the coda position under the condition that it is also linked to an onset (e.g., [os.s'ok] 'inside of clothes'). That is, [s] can never be an independent coda consonant. Similarly, it is more precise to say that the feature [lateral] can never be *independently* linked to an onset position.

Korean loan word adaptations of English forms that end in [l] quite clearly show that [lateral] is preserved in the coda position (e.g., [hol] 'hall', [pel.bet] 'velvet'). That is, it is not the case that [lateral] is simply recovered as long and thus must be doubly linked to both an onset and a coda position (i.e., *[hol.lɯ], *[pel.lɯbet] are not likely interpretations of 'hall' and 'velvet', respectively). Just as with the syllable-final [l] cases, when [l] occurs in intervocalic positions in loan words, the feature [lateral] is preserved instead of being replaced by [ɾ]. However, preserving laterality, in this case, is costly. The feature [lateral] is realized as a geminate (e.g., [ɾul.let] 'roulette', [kil.lo] 'kilo', [p'ol.li.e.sɯ.t'e.rɯ] 'polyester'). That is, instead of mirroring a production rule, and, thereby, changing the sound into the relevant allophone ([ɾ]), the feature [lateral] is preserved in the only way it can be intervocalically. Likewise, the

English /r/ in the word-final position in 'polyester' is pronounced with the flap [ɾ], instead of [l], in the Korean. Since [ɾ] occurs only in onset position in Korean, it conditions vowel epenthesis. In the case of the present study, perceptual epenthesis, rather than assimilation, is again the only way to preserve the feature [strident].

Another important piece of evidence for the preservation of the feature [lateral] comes from loan words where [lateral] occurs in *word-medial onset* positions. This is different from the examples given above in that the lateral segment occurs in an onset position following coda consonants in English. In such a case, the feature is also geminated (e.g., [igɯl.lu] 'igloo'; [a.t^hɯl.lan.t^ha] 'Atlanta') instead of changing into a flap (i.e., *[ig.ɾu], *[at.ɾan.t^ha]). This is, again, analogous to the preservation of the feature [strident] in the present study, as well as in loan words in Korean, where the only way to accommodate it is to link the feature to an onset position (e.g., [p^hacma]→ [p^hac^hima], or, in loan words, 'match'→ [mac^hi] instead of [mat]). In the case of [lateral], however, the only way to preserve the feature [lateral] is to associate it with a coda position. Given [VlV], it is possible to preserve [lateral] by positing a segment in the coda, /Vl.V/. However, although codas do not have to be present in Korean (i.e., [V.CV] is permissible), onsets cannot generally be absent (i.e., *[VC.V]). That is, the onset maximization principle obtains along with the coda requirement on [lateral], which together produce a geminate representation /Vl.lV/. When we look at those cases where [l] follows a consonant as in 'igloo' or 'Atlanta', the coda position that the feature [lateral] requires is already filled with another segment. An extra syllable is, thus, posited, which provides a vacant coda position for the feature [lateral] (i.e., [igɯl.lu], [a.tʰɯl.lan.tʰa]).

It should be noted that in word-initial situations, Korean exhibits a different strategy with regard to the feature [lateral]. [l]-initial words of foreign origin, which, together with [r]-initial foreign words, constitute all of the /l/-initial words in the Korean lexicon, are typically pronounced with a flap (e.g., [ɾemon] 'lemon', [ɾeisɯ] 'lace'). Since the only way to preserve [lateral] is to link it to an extra syllable would have to be created at the beginning of the word (e.g., *[ɯl.le.mon], *[ɯl.le.i.sɯ]). In this case, a prosodic constraint banning prothesis can be said to outweigh whatever constraints are designated for the preservation of the feature [lateral][22]. Indeed, prothesis before a lateral geminate would result in a structure that is very unusual in Korean: there are perhaps only two words that start with /ɯl.lV/ in the entire Korean lexicon. The infrequency of such words may force Korean speakers to simply interpret the [lateral] initial words as containing the flap, which is a possible onset consonant.

[22] This point has been brought to my attention by John Kingston.

Alternatively, it could also be that when /l/-initial words were written in the Korean alphabet as being /l/-initial, these words were automatically regarded as beginning with an [ɾ] phonetically by virtue of analogy to word-medial onset positions where /l/ is realized as [ɾ]. These speculations naturally imply some limited role for frequency.

In the absence of any perceptual experimental evidence as to how coda-only features are perceptually accommodated when they occur in onset positions in the L2, any proposal that regards the way these features are accommodated in perception cannot go beyond speculation. The loan word evidence strongly suggests that a coda-only feature such as [lateral] is preserved and the syllable structure is modified to accommodate it when necessary[23]. The perception of such forms may, however, not

---

[23] The Korean velar nasal is also an example of a [Coda] consonant. However, it behaves differently from other consonants in syllabification in that it cannot be syllabified as the onset of the following syllable while most other consonants are normally syllabified as the onset of the next syllable (Chung, 2001). That is, the language prefers $C_1V_1\mathbf{C_2}.V_2C_3$. over $C_1V_1\mathbf{.C_2}V_2C_3$ when C_2 is a velar nasal (e.g., /saŋo/→[saŋ.o], *[sa. ŋo] 'morning'). Compared to the Seoul dialect (SD), in the Kyungsang dialect (KD), a velar nasal in coda position is deleted, but the neighboring vowels are still nasalized (e.g., /talpʰeŋi/→[tal.pʰẽŋ.ĩ] (SD), [tal.pʰẽ.ĩ] (KD) 'snail'. In both dialects of Korean, it seems that the velar nasal is accommodated in different ways to avoid Onset formation, which is further evidence for the [-Onset] property of coda-only consonants. It could be that in the KD, nasal deletion is actually motivated in intervocalic contexts, in which case it is hard to tell whether the velar nasal occupied the coda or the onset position (i.e., /talpʰeŋi/→[tal.pʰẽŋ.ĩ] / [tal.pʰẽ.ŋĩ] →[tal.pʰẽ.ĩ]). Since the status of [ŋ] and the syllabic behavior of the features that make up this segment are poorly understood, any discussion of this matter is nothing more than speculation.

be symmetrical to the way they are produced, a point that has been made throughout this study.

In summary, the perceptual algorithm developed here is consistent with the experimental data collected here. It motivates perceptual epenthesis in Korean from a necessity to build a CV syllable when an [Onset]-linked feature is perceived. Accordingly, a vowel must follow to fill the nucleus since onsets must be non-branching in Korean. With features that are not linked to [Onset] in Korean (e.g., /k/), however, the algorithm does not postulate an empty nucleus, even at the expense of creating an illegal consonantal contact with subsequent consonantal features (e.g., *[k.m]). The present findings as well as the observations on loan word adaptations also suggest that the perceptual system has a desire not only for well-formed syllables but also for the preservation of features. This may suggest that assimilation did not arise as a perceptual repair strategy in the perception of illegal clusters such as [km] and [ln] because if these forms were to change into [ŋm] and [ll], the features corresponding to [k] and [n] would not be preserved. In the case of [strident] occurring in the coda position, the feature can only be preserved if it were interpreted as an onset, which results in epenthesis. Similarly, when the feature [lateral] occurs in the onset position, the feature can only be preserved if it is linked to a coda position, which results in its gemination, leading to epenthesis if the coda position is not available.

The present study has investigated the onset-only cases, leading to an important finding that prioritizes the role of syllables in perception. Naturally, further experimental research on the perception of coda-only segments is necessary. Especially, testing the perception of [l]-initial syllables in various situations by Korean listeners suggests a promising research enterprise.

Chapter 8

CONCLUSIONS AND REMAINING QUESTIONS

This dissertation has revealed several important findings. First, based on empirical evidence, I have shown that having a sequentially illicit consonantal sequence is not sufficient to induce perceptual epenthesis. Rather, I have argued that the first member of the cluster must be actively banned from the coda in the L1 in order to evoke perceptual epenthesis effects. The present research, therefore, constitutes substantial evidence against any analyses that employ syllable-independent, string-based and linear statements to explain consonantal phonotactics (e.g., Steriade 1999; 2001a; Dziubalska-Kolaczyk, 1994; Blevins, 2002).

Second, the present study has shown that L2 representations are made on the basis of the abstract phonological properties of the L1 system of contrast. I have claimed that the perceptual system suspends featural information in the speech signal if the detected values correspond to features that are underspecified in the L1. Supporting this notion, the present study has shown that [voice] was suspended in the perceptual representations of Korean listeners. That is, a voiced segment such as [g], which cannot surface in the coda position, was interpreted as /k/ by the Korean listeners, thus did not lead to perceptual epenthesis.

Third, I have shown that other types of phonological processes such as neutralization, lateralization and nasalization do not constitute possible perceptual repair strategies in the perception of illicit consonant clusters by Korean listeners. I have taken these findings to provide strong evidence against models that prioritize the role of perception in explaining synchronic phonological alterations (e.g., Steriade, various). Accordingly, I have argued that not all phonology is relevant to speech perception and that speech perception is not isomorphic to speech production.

Fourth, the perceptual epenthesis effects found in the present study cannot be explained by the frequency of consonant clusters in Korean. Despite their zero probability of occurrence, certain consonant clusters such as [km] and [ln] were successfully perceived by the Korean listeners.

Fifth, having shown that neither consonantal contact violations nor frequency can explain the present findings, I have proposed a syllable-based perceptual algorithm to capture the perceptual epenthesis effects in a simple and straightforward manner. The algorithm employs syllabically-conditioned detectors, which evaluate language-specific onset or coda identity of features or feature bundles. Accordingly, certain (sets of) features are prespecified to signal [Onset], and certain others to [Coda]. I have claimed that the desire of the perceptual system for well-formed syllables is the cause of perceptual epenthesis. At this point, it would be interesting to test in what other circumstances perceptual repair strategies may emerge and what the role of syllable structure for such strategies can be .

The present findings raise further questions that can lead to potential research topics. One of these questions regards the way voicing is represented cross-linguistically. In the present study, I have argued that the perceptual system treated the feature [strident] differently from the feature [voice]. While the former feature is linked to [ONSET], the latter does not induce "onsetness" although in the production grammar voiced segments always occupy onset positions. I have considered the reason that the feature [voice] is not linked to [ONSET] in Korean as an indication of phonological processing, rather than phonetic processing, of L2 strings. Second, it would be interesting to investigate whether phonological processing applies to other allophonic features in the L1. Can all allophonic, predictable features be suppressed? If not, what are the (phonetic or acoustic) factors that determine a feature's immunity to suppression?

Finally, the question of whether consonants that are only specified for [Coda] ever evoke perceptual epenthesis promises to be very fruitful for future studies. An enterprise in this direction will be especially crucial for speech perception models that make reference to the syllable and prosodic positions within it.

Future research should look further into those cases where perceptual errors do not necessarily mirror production phenomena in the L1. Studies that investigate the discrepancy between speech production and perception will have important consequences and posit crucial implications for speech processing models. Finally, future research should further explore the possibility that imperfect production may be due to imperfect representations that result from the filtering done by the L1 phonological system.

APPENDIX A

CERTIFICATE OF INFORMED CONSENT

What do we do when listening to foreign language words?

Investigator: Baris Kabak
University of Delaware, Department of Linguistics

1. WHAT IS THIS RESEARCH STUDY ABOUT?

This experiment is part of a cross-linguistic research study that investigates the ways in which the native language influences the perception of words from a foreign language. It is well known that having acquired a language, people have relatively difficult experiences in learning a foreign language. This difficulty usually depends on the extent to which there are differences between the native language and the target language in question. This particular experiment in which you are about to participate involves native speakers of different language groups such as Korean, Turkish and German and their perception of words from a foreign language.

2. WHAT ARE THE EXPERIMENTAL PROCEDURES?

You will be asked to sit in a quiet room in front of a computer and listen to a series of nonsense words from a foreign language with a pair of headphones. You will hear the words at a comfortable volume. Your participation is required only once in this experiment. The experiment takes about 45-50 minutes. You will have a brief session in which you are given instructions as to what you need to do and go through a brief practice session. Please do not hesitate to ask the investigator if you have any questions or concerns before the experiment.

a. **What do I need to do during the experiment?**

You will hear pairs of words uttered by an American speaker. (S)he will say a word and then after a brief tone the same speaker will say another word. The two words will constitute a pair.

Your task is to determine, for each pair, whether the speaker has repeated what (s)he has said before, or whether (s)he said a different word. In other words, you are asked to decide if the speaker has repeated the same word, or not. If you think that the speaker has repeated the same word, you will press “A” on the keyboard in front of you. If, on the other hand, you think that (s)he has said a different word, you need to press “L”.

For each pair you will be given approximately 6 seconds to respond. Therefore, you will be asked to respond as fast as possible. If you do not respond in time, the computer will present you another pair. There are a number of breaks during the experiment. The computer will periodically prompt you a notice and ask you if you would like to take a break. At the end of the experiment, there will be a prompt on the computer screen to take off the headphones and leave the room.

3. WHAT ARE THE BENEFITS AND RISKS?
Your participation is considered to be voluntary. There is no penalty from withdrawing from this research. You may refuse to participate before the start of the experiment. You may also discontinue the experiment at any time and any stage during the experiment. You will bear no responsibility for any consequence that your withdrawal from the experiment may cause to the investigators. Your personal information, responses to experimental items, and the analysis of the results that are relevant to your performance will remain strictly confidential.

There are no anticipated risks that may arise by participating in this experiment.

4. DATA USE
The data obtained from this study may be used in the investigator’s relevant publications, conference posters and presentations and dissertation. They will be kept in a safe place in an indefinite period of time by the investigator.

5. CONTACT INFORMATION
If you have any questions about this study, please contact:
Investigator:
Baris Kabak, PhD student
Department of Linguistics
46 E. Delaware Ave.
Newark, DE 19716
kabak@udel.edu

Academic Advisor:
Dr. William J. Idsardi, Associate Professor
Department of Linguistics
46 E. Delaware Ave.
Newark, DE 19716
idsardi@udel.edu

If you have any questions and concerns regarding your rights, please contact:

Dr. T.W. Fraser Russell, Acting Vice Provost for Research
University of Delaware, Newark, DE 19716-- Phone: 302-831-4007

My signature below certifies that the experiment I am about to complete has been explained to me and that all my questions have been answered satisfactorily. I voluntarily agree to complete this experiment. Furthermore, I recognize that I may refuse to answer any questions without penalty.

Name: -----------------------------------

Signature: -----------------------------------

Date: -----------------------------------

APPENDIX B

QUESTIONNAIRE

Please take your time to fill in this form. The information you provide will be very useful in interpreting the results of this study and it will strictly remain confidential.

1. Age:

2. Sex:

Languages spoken

3. a. First language:

3. b. Other languages: (please indicate your level of proficiency as one of "elementary", "intermediate", or "advanced"):

4. Country of Birth:

For non-native speakers of English:

5. How long have you studied English?

6. How long have you lived in the USA (or any other major English speaking countries)?

7. Did you go to school in the USA? If yes, what schools/courses have you attended?

8. Do you speak any particular dialect of your native language? If yes, please tell us what dialect it is.

APPENDIX C

ENGLISH AND KOREAN MEDIAN AND MEAN D-PRIME SCORES AND STANDARD DEVIATIONS FOR EXPERIMENTAL DOUBLETS

	ENGLISH (n=25)			KOREAN (n=25)		
Doublet	**Median d'**	**Mean d'**	**Std. Dev.**	**Median d'**	**Mean d'**	**Std. Dev**
p^{h}acma-p^{h}ac^{h}ɨma	3.01	3.03	0.670	0	0.31	0.577
p^{h}act^{h}a-p^{h}ac^{h}it^{h}a	3.46	3.35	0.561	0	0.89	1.01
p^{h}ajma-p^{h}ajima	3.46	3.42	0.640	0	0.48	0.77
p^{h}ajt^{h}a-p^{h}ajit^{h}a	3.46	3.44	0.499	0	0.47	0.81
p^{h}akma-p^{h}ak^{h}ʊma	3.57	3.58	0.563	2.39	2.28	1.07
p^{h}akt^{h}a-p^{h}ak^{h}ʊt^{h}a	3.92	3.77	0.366	3.86	3.58	0.66
p^{h}agma-p^{h}agʊma	2.31	2.53	0.905	2.14	1.81	1.22
p^{h}agt^{h}a- p^{h}agʊt^{h}a	3.59	3.56	0.578	3.03	2.73	1.01
p^{h}alna-p^{h}alʊna	3.57	3.59	0.486	3.59	3.42	0.59
p^{h}alt^{h}a-p^{h}alʊt^{h}a	3.59	3.55	0.569	3.19	3.29	0.62
p^{h}akma-p^{h}aŋma	3.92	3.69	0.55	3.92	3.73	0.48
p^{h}agma-p^{h}aŋma	3.59	3.66	0.42	3.59	3.64	0.42
p^{h}alna-p^{h}alla	3.86	3.74	0.45	3.59	3.56	0.43
p^{h}alna-p^{h}anna	3.92	3.75	0.47	3.59	3.48	0.62
p^{h}akma-p^{h}ak^{h}ima	3.92	3.85	0.40	3.92	3.83	0.48

APPENDIX D

INDIVIDUAL D-PRIME VALUES AND COST OF FEATURES

Table D.1: Individual English d' scores and cost of features

	cm / cim	ct / cit	jm / jim	jt / jit	km / kum	kt / kut	gm / gum	gt / gut	ln / lun	lt / lut	km / ŋm	ln / ll	ln / nn	km / kim	Cost Voi.	Cost Strd	Cost .Nas
1	3.57	3.46	3.46	3.86	3.92	3.92	2.05	3.86	3.17	3.19	3.92	3.24	3.92	3.24	0.4	0.2	0.4
2	2.93	3.46	3.46	4.22	3.24	3.59	0.85	3.04	3.92	3.59	3.59	3.92	3.92	3.92	0.4	0.8	0.7
3	2.93	2.93	3.17	2.93	3.01	3.92	2.28	3.92	3.92	3.92	3.37	3.86	4.22	3.59	0.1	-0.3	0.5
4	3.17	3.86	3.24	3.19	4.22	3.24	2.12	3.86	4.22	3.92	4.22	4.22	4.22	4.22	0.5	0.0	0.2
5	1.78	2.46	2.15	2.20	3.57	3.92	3.17	2.63	3.17	1.88	3.01	3.46	4.22	3.92	0.4	-1.2	-0.2
6	3.46	3.17	3.86	3.57	3.24	4.22	2.31	3.92	3.57	3.17	3.37	3.57	3.37	3.92	0.1	0.1	0.3
7	2.83	3.59	3.59	3.01	3.59	3.37	2.83	3.37	2.53	3.04	3.19	2.77	2.41	3.37	0.1	0.0	0.2
8	2.63	3.17	2.93	2.88	3.24	3.86	3.24	3.24	3.57	4.22	3.19	3.92	3.24	3.59	0.2	-0.5	0.4
9	3.24	3.86	3.86	3.17	3.59	3.59	1.21	2.61	3.57	4.22	3.92	4.22	4.22	3.92	0.9	0.8	0.4
10	1.72	2.54	1.98	3.17	2.54	3.37	1.91	3.37	3.04	2.90	1.85	3.37	3.59	3.04	-0.1	-0.4	0.8
11	3.17	2.83	3.92	3.59	3.92	3.92	2.61	3.90	3.37	3.04	3.24	3.92	3.57	4.22	0.0	-0.2	0.1
12	3.92	3.17	3.86	3.37	3.57	3.92	1.85	3.57	3.57	3.37	3.59	3.92	3.92	3.92	0.5	0.4	0.1
13	2.39	3.86	3.92	4.22	4.22	3.92	2.54	3.19	4.22	4.22	3.92	4.22	4.22	4.22	0.1	0.1	0.4
14	4.22	3.92	4.22	3.86	4.22	4.22	4.22	4.22	3.59	4.22	4.22	3.59	3.59	4.22	0.0	-0.2	0.0
15	3.17	3.86	2.88	2.69	3.17	3.17	3.46	3.92	3.57	3.24	3.37	4.22	3.92	3.92	0.1	-0.3	0.1
16	2.69	2.69	4.22	3.57	4.22	3.86	2.93	3.92	4.22	3.24	4.22	4.22	3.92	4.22	-0.3	-0.4	-0.2
17	2.54	2.54	2.54	3.46	3.46	3.17	1.35	3.57	3.86	3.86	4.22	3.57	3.57	4.22	0.2	-0.1	0.6
18	3.86	3.92	3.17	3.17	3.24	3.92	1.51	3.24	3.92	3.57	3.92	3.57	3.57	3.01	1.0	0.6	0.4
19	4.22	3.92	3.24	3.92	3.37	3.92	2.54	4.22	4.22	3.92	3.59	4.22	4.22	3.19	0.4	0.3	0.5
20	3.01	4.22	4.22	3.92	4.22	3.59	1.95	3.59	3.59	3.86	4.22	3.59	3.59	4.22	0.3	0.5	0.4
21	2.83	3.86	3.59	3.92	3.92	4.22	3.59	4.22	2.77	3.04	4.22	2.77	2.90	3.92	-0.1	-0.4	0.5
22	3.46	3.86	3.92	3.86	4.22	3.86	3.92	3.59	4.22	3.92	4.22	4.22	4.22	4.22	0.0	-0.1	-0.1
23	2.04	2.88	3.17	3.14	2.12	3.01	2.22	1.85	2.83	3.17	3.57	3.19	3.19	3.92	-0.1	0.5	0.3
24	3.46	2.69	4.22	3.17	4.22	4.22	4.22	3.86	3.24	3.92	4.22	3.59	3.92	4.22	-0.2	-0.7	-0.3
25	2.56	2.69	2.63	3.86	3.17	4.22	2.28	4.22	3.86	4.22	3.92	4.22	4.22	3.92	-0.1	-0.5	0.9

Table D.2: Individual Korean d' scores and cost of features

	cm / cim	ct / cit	jm / jim	jt / jit	km / kum	kt / kut	gm / gum	gt / gut	ln / lun	lt / lut	km / ŋm	ln / ll	ln / nn	km / kim	Cost Voi.	Cost Strd	Cost .Nas
1	0.00	0.00	0.00	0.00	2.12	3.37	2.68	2.41	3.37	2.41	2.77	3.19	3.15	2.65	0.1	-2.6	0.0
2	0.00	0.00	0.00	0.00	1.54	2.99	2.20	2.69	3.86	3.92	4.22	3.86	4.22	4.22	-0.1	-2.4	0.4
3	0.00	1.41	0.99	0.00	2.61	4.22	3.17	3.17	2.68	4.22	4.22	2.83	3.19	4.22	0.2	-2.7	0.7
4	0.00	0.99	0.00	1.10	3.17	3.92	1.78	3.30	4.22	2.77	3.24	4.22	4.22	3.59	0.5	-2.5	0.6
5	0.00	1.88	0.00	0.00	2.39	4.22	1.41	2.12	1.80	2.77	3.92	3.19	2.65	3.92	1.2	-2.1	1.1
6	0.00	0.00	0.00	0.00	2.15	3.86	2.39	4.22	3.59	3.04	4.22	3.92	3.24	3.92	-0.1	-3.2	0.6
7	0.00	2.39	2.93	2.39	3.17	3.86	3.37	3.04	2.54	3.37	3.19	2.90	2.77	3.92	-0.6	-1.4	0.6
8	0.00	0.00	1.88	1.19	3.92	3.92	3.24	3.92	3.59	3.59	3.92	3.59	3.59	4.22	-0.6	-3.0	0.0
9	0.00	0.00	0.00	0.00	2.93	4.22	0.00	2.65	3.19	3.59	3.86	3.37	2.90	4.22	1.1	-2.4	0.9
10	0.00	0.00	0.00	0.00	2.46	3.17	0.00	2.93	3.86	3.86	3.86	3.86	4.22	4.22	0.7	-2.1	0.7
11	0.00	2.20	0.00	0.00	1.88	4.22	0.00	3.37	3.59	3.19	4.22	3.24	3.59	4.22	1.2	-1.8	1.5
12	1.79	0.00	0.00	1.54	2.61	3.86	2.54	3.19	3.92	2.90	3.59	3.92	3.86	3.37	0.2	-2.2	0.1
13	0.00	0.00	0.00	0.00	0.00	4.22	0.00	3.17	3.59	2.77	4.22	3.59	3.59	4.22	0.3	-1.8	1.3
14	1.41	0.00	0.00	0.00	0.00	3.17	1.10	3.17	4.22	3.92	3.92	3.86	3.59	3.92	0.1	-1.5	0.7
15	1.10	2.15	0.85	0.00	1.85	3.59	1.95	2.28	2.90	3.04	3.92	3.04	2.41	3.59	0.9	-1.4	0.5
16	0.00	0.00	0.00	0.00	1.41	2.69	0.00	0.00	3.92	4.22	3.86	3.86	4.22	3.86	1.0	-1.0	0.3
17	0.00	1.41	0.00	0.00	3.17	3.57	3.17	3.59	3.87	3.37	3.92	3.92	3.24	3.92	0.3	-3.0	0.3
18	1.41	0.00	0.99	0.00	1.49	1.52	0.00	0.99	3.17	3.57	2.93	3.92	4.22	3.37	0.6	-0.4	-0.2
19	0.99	0.00	1.08	0.00	4.15	4.20	3.24	3.04	3.54	2.86	3.82	2.83	3.24	3.56	0.5	-3.1	-0.6
21	0.99	1.88	0.99	1.88	2.39	3.46	2.20	2.04	3.86	3.17	3.17	4.22	4.22	3.92	0.4	-1.1	0.4
22	0.00	1.88	0.00	1.41	4.22	3.86	3.17	3.92	3.92	3.92	4.22	4.22	4.22	4.22	0.4	-3.0	0.7
23	0.00	0.00	0.00	0.00	2.69	3.86	1.41	3.86	3.37	4.22	4.22	3.37	2.90	4.22	0.3	-3.0	0.9
25	0.00	2.69	0.71	2.20	1.78	3.17	2.15	1.10	3.37	3.04	3.92	3.37	3.04	3.92	0.4	-0.6	0.8
26	0.00	2.20	0.00	0.00	1.48	2.80	2.63	2.42	2.42	2.28	3.15	3.37	2.41	2.39	0.4	-1.8	0.6
27	0.00	1.10	1.54	0.00	1.49	3.57	1.41	1.54	3.17	2.04	2.88	3.37	4.22	3.92	0.4	-1.3	0.1

Note: Subjects 20 and 24 are excluded from the analysis due to reasons given in Chapter 5, Section 5.3.

APPENDIX E

POST-HOC TEST TABLES FOR THE ENGLISH GROUP

Table E.1: The Scheffe Test for English d' scores (Significance Level: 1%)

Clusters	Mean Diff.	Crit. Diff.	P-Value	Significance
cm, ct^h	-0.304	0.681	0.8756	
cm, jm	-0.385	0.681	0.6218	
cm, jt^h	-0.405	0.681	0.5447	
cm, km	-0.545	0.681	0.1188	
cm, kt^h	-0.734	0.681	0.003	**S**
cm, gm	0.506	0.681	0.2025	
cm, gt^h	-0.524	0.681	0.1593	
cm, ln	-0.557	0.681	0.0987	
cm, lt^h	-0.522	0.681	0.1628	
ct^h, jm	-0.08	0.681	>.9999	
ct^h, jt^h	-0.1	0.681	>.9999	
ct^h, km	-0.24	0.681	0.9711	
ct^h, kt^h	-0.429	0.681	0.4506	
ct^h, gm	0.81	0.681	0.0004	**S**
ct^h, gt^h	-0.22	0.681	0.9846	
ct^h, ln	-0.253	0.681	0.9596	
ct^h, lt	-0.218	0.681	0.9854	
jm, jt^h	-0.02	0.681	>.9999	
jm, km	-0.16	0.681	0.9986	
jm, kt^h	-0.349	0.681	0.7509	
jm, gm	0.89	0.681	<.0001	**S**
jm, gt^h	-0.139	0.681	0.9995	
jm, ln	-0.172	0.681	0.9975	
jm, lt^h	-0.138	0.681	0.9996	
jt^h, km	-0.14	0.681	0.9995	
jt^h, kt^h	-0.329	0.681	0.8128	
jt^h, gm	0.91	0.681	<.0001	**S**

Table E.1: continued

jt^h, gt^h	-0.119	0.681	0.9999	
jt^h, ln	-0.152	0.681	0.9991	
jt^h, lt^h	-0.118	0.681	0.9999	
km, kt^h	-0.189	0.681	0.9949	
km, gm	1.05	0.681	<.0001	**S**
km, gt^h	0.021	0.681	>.9999	
km, ln	-0.012	0.681	>.9999	
km, lt^h	0.022	0.681	>.9999	
kt^h, gm	1.239	0.681	<.0001	**S**
kt^h, gt^h	0.21	0.681	0.989	
kt^h, ln	0.176	0.681	0.997	
kt^h, lt^h	0.211	0.681	0.9883	
gm, gt^h	-1.03	0.681	<.0001	**S**
gm, ln	-1.063	0.681	<.0001	**S**
gm, lt^h	-1.028	0.681	<.0001	**S**
gt^h, ln	-0.033	0.681	>.9999	
gt^h, lt^h	0.002	0.681	>.9999	
ln, lt^h	0.035	0.681	>.9999	

Table E.2: The Tukey/Kramer test for English d' scores (Significance Level: 1%)

Clusters	Mean Diff.	Crit. Diff.	Significance
cm, ct^h	-0.304	0.536	
cm, jm	-0.385	0.536	
cm, jt^h	-0.405	0.536	
cm, km	-0.545	0.536	**S**
cm, kt^h	-0.734	0.536	**S**
cm, gm	0.506	0.536	
cm, gt^h	-0.524	0.536	
cm, ln	-0.557	0.536	**S**
cm, lt^h	-0.522	0.536	
ct^h, jm	-0.08	0.536	
ct^h, jt^h	-0.1	0.536	
ct^h, km	-0.24	0.536	
ct^h, kt^h	-0.429	0.536	

Table E.2: continued

ct^h, gm	0.81	0.536	**S**
ct^h, gt^h	-0.22	0.536	
ct^h, ln	-0.253	0.536	
ct^h, lt	-0.218	0.536	
jm, jt^h	-0.02	0.536	
jm, km	-0.16	0.536	
jm, kt^h	-0.349	0.536	
jm, gm	0.89	0.536	**S**
jm, gt^h	-0.139	0.536	
jm, ln	-0.172	0.536	
jm, lt^h	-0.138	0.536	
jt^h, km	-0.14	0.536	
jt^h, kt^h	-0.329	0.536	
jt^h, gm	0.91	0.536	**S**
jt^h, gt^h	-0.119	0.536	
jt^h, ln	-0.152	0.536	
jt^h, lt^h	-0.118	0.536	
km, kt^h	-0.189	0.536	
km, gm	1.05	0.536	**S**
km, gt^h	0.021	0.536	
km, ln	-0.012	0.536	
km, lt^h	0.022	0.536	
kt^h, gm	1.239	0.536	**S**
kt^h, gt^h	0.21	0.536	
kt^h, ln	0.176	0.536	
kt^h, lt^h	0.211	0.536	
gm, gt^h	-1.03	0.536	**S**
gm, ln	-1.063	0.536	**S**
gm, lt^h	-1.028	0.536	**S**
gt^h, ln	-0.033	0.536	
gt^h, lt^h	0.002	0.536	
ln, lt^h	0.035	0.536	

APPENDIX F

POST-HOC TEST TABLES FOR THE KOREAN GROUP

Table F.1: The Scheffe Test for Korean d' scores (Significance Level: 1%)

Clusters	Mean Diff.	Crit. Diff.	P-Value	Significant
cm, ct^h	-0.58	1.07	0.6804	
cm, jm	-0.171	1.07	>.9999	
cm, jt^h	-0.161	1.07	>.9999	
cm, km	-1.975	1.07	<.0001	**S**
cm, kt^h	-3.273	1.07	<.0001	**S**
cm, gm	-1.501	1.07	<.0001	**S**
cm, gt^h	-2.418	1.07	<.0001	**S**
cm, ln	-3.114	1.07	<.0001	**S**
cm, lt^h	-2.974	1.07	<.0001	**S**
ct^h, jm	0.409	1.07	0.9512	
ct^h, jt^h	0.419	1.07	0.9431	
ct^h, km	-1.396	1.07	<.0001	**S**
ct^h, kt^h	-2.693	1.07	<.0001	**S**
ct^h, gm	-0.921	1.07	0.0616	
ct^h, gt^h	-1.838	1.07	<.0001	**S**
ct^h, ln	-2.534	1.07	<.0001	**S**
ct^h, lt	-2.395	1.07	<.0001	**S**
jm, jt^h	0.01	1.07	>.9999	
jm, km	-1.804	1.07	<.0001	**S**
jm, kt^h	-3.102	1.07	<.0001	**S**
jm, gm	-1.33	1.07	0.0002	**S**
jm, gt^h	-2.247	1.07	<.0001	**S**
jm, ln	-2.943	1.07	<.0001	**S**
jm, lt^h	-2.804	1.07	<.0001	**S**
jt^h, km	-1.814	1.07	<.0001	**S**
jt^h, kt^h	-3.112	1.07	<.0001	**S**
jt^h, gm	-1.34	1.07	0.0001	**S**

Table F.1: continued

jt^h, gt^h	-2.257	1.07	<.0001	**S**
jt^h, ln	-2.953	1.07	<.0001	**S**
jt^h, lt^h	-2.814	1.07	<.0001	**S**
km, kt^h	-1.298	1.07	0.0003	**S**
km, gm	0.474	1.07	0.8807	
km, gt^h	-0.442	1.07	0.9204	
km, ln	-1.138	1.07	0.0037	**S**
km, lt^h	-0.999	1.07	0.0251	
kt^h, gm	1.772	1.07	<.0001	**S**
kt^h, gt^h	0.855	1.07	0.1193	
kt^h, ln	0.159	1.07	>.9999	
kt^h, lt^h	0.298	1.07	0.9947	
gm, gt^h	-0.917	1.07	0.0645	
gm, ln	-1.613	1.07	<.0001	**S**
gm, lt^h	-1.474	1.07	<.0001	**S**
gt^h, ln	-0.696	1.07	0.3984	
gt^h, lt^h	-0.557	1.07	0.7316	
ln, lt^h	0.139	1.07	>.9999	

Table F.2: The Tukey/Kramer test for Korean d' scores (Significance Level: 1%)

Clusters	**Mean Diff.**	**Crit. Diff.**	**Significance**
cm, ct^h	-0.58	0.841	
cm, jm	-0.171	0.841	
cm, jt^h	-0.161	0.841	
cm, km	-1.975	0.841	**S**
cm, kt^h	-3.273	0.841	**S**
cm, gm	-1.501	0.841	**S**
cm, gt^h	-2.418	0.841	**S**
cm, ln	-3.114	0.841	**S**
cm, lt^h	-2.974	0.841	**S**
ct^h, jm	0.409	0.841	
ct^h, jt^h	0.419	0.841	
ct^h, km	-1.396	0.841	**S**
ct^h, kt^h	-2.693	0.841	**S**

Table F.2: continued

ct^h, gm	-0.921	0.841	**S**
ct^h, gt^h	-1.838	0.841	**S**
ct^h, ln	-2.534	0.841	**S**
ct^h, lt	-2.395	0.841	**S**
jm, jt^h	0.01	0.841	
jm, km	-1.804	0.841	**S**
jm, kt^h	-3.102	0.841	**S**
jm, gm	-1.33	0.841	**S**
jm, gt^h	-2.247	0.841	**S**
jm, ln	-2.943	0.841	**S**
jm, lt^h	-2.804	0.841	**S**
jt^h, km	-1.814	0.841	**S**
jt^h, kt^h	-3.112	0.841	**S**
jt^h, gm	-1.34	0.841	**S**
jt^h, gt^h	-2.257	0.841	**S**
jt^h, ln	-2.953	0.841	**S**
jt^h, lt^h	-2.814	0.841	**S**
km, kt^h	-1.298	0.841	**S**
km, gm	0.474	0.841	
km, gt^h	-0.442	0.841	
km, ln	-1.138	0.841	**S**
km, lt^h	-0.999	0.841	**S**
kt^h, gm	1.772	0.841	**S**
kt^h, gt^h	0.855	0.841	**S**
kt^h, ln	0.159	0.841	
kt^h, lt^h	0.298	0.841	
gm, gt^h	-0.917	0.841	**S**
gm, ln	-1.613	0.841	**S**
gm, lt^h	-1.474	0.841	**S**
gt^h, ln	-0.696	0.841	
gt^h, lt^h	-0.557	0.841	
ln, lt^h	0.139	0.841	

REFERENCES

Ahn, S.-C. (1998). *An Introduction to Korean Phonology*. Seul, Korea: Hanshin.

Archangeli, D. (1984). *Underspecification in Yawelmani Phonology and Morphology*. PhD dissertation. MIT: Cambridge, Mass. Published by Garland Press, New York, 1988.

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology, 5*, 183-207.

Beckman, M. E. (1982). Segmental duration and the mora in Japanese. *Phonetica, 39*, 113-135.

Blevins, J. (1995). The syllable in phonological theory. In J. A. Goldsmith (ed.). *The Handbook of Phonological Theory*. Blackwell.

Blevins, J. (2002). The independent nature of phonotactic constraints: an alternative to syllable-based approaches. In C. Féry and R. van de Weijer (eds.). *The Syllable in Optimality Theory*, Cambridge University Press.

Boersma, P. and Weenink, D. *PRAAT: doing phonetics by computer*. Website: http://www.praat.org, 1992-2002.

Boersma, P. (1998). Functional phonology: Formalizing the interactions between articulatory and perceptual drives. [*LOT International Series* 11]. PhD dissertation, University of Amsterdam. The Hague: Holland Academic Graphics.

Chung, C.-W. (2001). The behavior of velar nasal and syllabification. *Studies in Phonetics, Phonology and Morphology, 7*, 1, 177-189.

Christie, W. (1974). Some cues for syllable juncture perception in English. *Journal of the Acoustical Society of America, 55*, 819-821.

Church, K. W. (1987). Phonological parsing and lexical retrieval. *Cognition, 25*, 53–70.

Cohen J.D., MacWhinney B., Flatt M., and Provost J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25*, 2, 257- 271.

Cutler, A. & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 113-121.

Cutler, A., Demuth, K. & McQueen, J. M. (2002). Universality versus language-specificity in listenning to running speech. *Psychological Science*, *13*, 3, 258-262.

Crystal, T. H., and House, A. S. (1988). Segmental durations in connected speech signals: Syllabic stress. *Journal of the Acoustical Society of America*, *83*, 1574-1585.

Dehaene-Lambertz, G., Dupoux, E., & Gout, A. (2000). Electrophysiological correlates of phonological processing: a cross-linguistic study. *Journal of Cognitive Neuroscience*, *12*, 635-647.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 6, 1568-1578.

Dupoux, E., Pallier, C., Kakehi, K., & Mehler, J. (2001). New evidence for prelexical phonological processing in word recognition. *Language and Cognitive Processes*, *16,* 5, 491-505.

Dziubalska-Kolaczyk, K. (1994). Acquisitional evidence against the phonological syllable. *Papers and Studies in Contrastive Linguistics*, *29*, 17-30.

Ewen, C. & Van der Hulst, J. (2001). *The Phonological Structure of Words: An Introduction*. Cambridge University Press.

Fitzpatrick J. & Wheeldon, L. (2000). Phonology and phonetics in psycholinguistic models of speech Perception. In N. Burton-Roberts, P. Carr, & G. Docherty (eds.). *Phonological Knowledge: Conceptual and Empirical Issues*. Oxford University Press.

Frauenfelder, U. H., & Kearns, R. (1996). Sequence monitoring. *Language and Cognitive Processes*, *11*, 665-674.

Frauenfelder, U. H., & Lahiri, A. (1989). Understanding words and word recognition: Does phonology help? In W. Marslen-Wilson (ed.), *Lexical Representation and Process*. pp. 319-341. Cambridge, MA: MIT Press.

Friederici, A. D., & Wessels, J.M.I. (1993). Phonotactic knowledge and its use in infant speech perception. *Perception & Psychophysics*, *54*, 287-295.

Ganong, W. F., (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 110-125.

Green, D. M., and Swets, J. A. (1974). *Signal Detection Theory and Psychophysics*. Huntington, NY: Krieger.

Hallé, P., Segui, J., Frauenfelder, U. and Meunier, C. (1998) Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance, 24*, 2, 592-608.

Itô, J. (1986). *Syllable theory in prosodic phonology*. PhD Dissertation, University of Massachussetts, Amherst.

Itô, J. (1989). A prosodic theory of epenthesis. *Natural Language and Linguistic Theory, 7*, 217-259.

Jusczyk, P., Friederici, A., Wessels, J., Svenkerud, V., & Jusczyk, A. (1993). Infants' sensitivity to the sound pattern of native language words. *Journal of Memory and Language*, *32*, 402-420.

Jusczyk, P., Luce, P., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, *33*, 630-645.

Jusczyk, P., Houston, D. & Goodman, M. (1998). Speech perception during the first year. In A. Slater (ed.), *Perceptual development: Visual, Auditory, and Speech Perception in Infancy*. Psychology Press.

Kadlec, H. (1999). Statistical properties of d' and [beta] estimates of signal detection theory. *Psychological Methods, 4*, 22-43.

Kang, H. (1995). English loanwords in Korean. *Studies in Phonetics, Phonology and Morphology*, *2*.

Kang, Y. (2001). Adaptation of English post-vocalic word-final stops to Korean Loanwords. Paper presented at NELS-32, CUNY/NYU: New York, NY.

Keating, P., & Hoffman, M. (1984). Vowel variation in Japanese. *Phonetica*, *41*, 191-207.

Kim-Renaud, Y-K.(1995). *Korean Consonantal Phonology*. Hanshin Publishing Company.

Lahiri, A. & Jongman, A. (1990). Intermediate Level of Analysis: Features or Segments? *Journal of Phonetics*, *18*, 435-443.

Lahiri, A. & Marslen-Wilson, W. D. (1991). The mental representation of lexical form: A phonological approach to recognition of lexicon. *Cognition*, *38*, 245-294.

Lahiri, A. & Marslen-Wilson, W. D. (1992). Lexical Processing and Phonological Representation. In G. J. Docherty and D. R. Ladd (eds.) *Papers in Laboratory Phonology II*, 229-254. Cambridge: CUP.

Lahiri, A.& Reetz, H. (1999). The FUL Speech Recognition System. *Journal of the Acoustical Society of America*, 105, 1091.

Luce, P. A. & Pissoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*, 1-36.

Macmillan, N. A. and Creelman, C. D. (1991). *Detection Theory: A User's Guide*. Cambridge, UK: Cambridge University Press.

Macmillan, N. A. (2002). Signal Detection Theory. In H. Pashler & J. Wixted (eds.) *Handbook of Experimental Psychology: Methodology in Experimental Psychology*. Vol. 4., 3rd edition.

Massaro, D. W., & Cohen, M. M. (1983). Phonological constrainst in speech perception. *Perception & Psychophysics*, *34*, 338– 348.

Massaro, D. W., & Oden, G. C. (1980a). Evaluation and integration of acoustic features in speech perception. *Journal of the Acoustical Society of America*, *67*, 996-1013.

Massaro, D. W., & Oden, G. C. (1980b). Speech perception: A framework for research and theory. In N. J. Lass (ed.), *Speech and Language: Advances in basic research and practice* (Vol. 3). New York: Academic Press.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86.

Mehler, J., Dupoux, E., & Segui, J. (1990). Constraining models of lexical access: The onset of word recognition. In G. Altmann (ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. pp. 236-261. Camridge, MA: MIT Press.

Mehler, J., Dommergues, J.-Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, *20*, 298-305.

Moreton, E. (2002). Structural constraints in the perception of English stop-sonorant clusters. *Cognition*, *84*, 55-71.

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition,* 52, 189–234.

Pitt, M. A. (1998). Phonological processes and the perception of phonetically illegal consonant clusters. *Perception and Pschycophysics*, *60*, 941-951.

Pitt, M. A. & McQueen, J. M. (1998). Is compensation for coarticulation mediated between monosyllabic spoken words. *Journal of Experimental Psychology: Learning, memory and Cognition, 15*, 378-387.

Polivanov, E. (1931). La perception des sons d'une langue etrangere. *Travaux du Cercle linguistic de Prague*, *4*, 79-86.

Prince, A. & Smolensky, P. (1993). *Optimality Theory: constraint interaction in generative grammar*. Ms., Rutgers University, New Brunswick and University of Colorado, Boulder.

Prince, A. and Tesar, B. (1999). Learning Phonotactic Distributions. Technical Report RuCCS-TR-54, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ.

Samuel, A. G. (1997). Lexical activation produces potent phonetic percepts. *Cognitive Psychology*, *32*, 97-127.

Sapir, E. (1933). La réalite psychologique des phonemes. *Journal de Psychologie normale et Pathologique*, *30*, 247-265.

Sawusch, J. R. (1996). Instrumentation and methodology for the study of speech perception. In N. J. Lass (ed.), *Principles of Experimental Phonetics*. Mosby.

Silverman, D. (1992). Multiple scansions in loanword phonology: evidence from Cantonese. *Phonology*, *9*, 289-328.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*, 34–50.

Sohn, H-M. (1994). *Korean*. Routledge.

Sohn, H-M (1999). *The Korean Language*. Cambridge University Press.

Steriade, D. (1997). Phonetics in phonology: The case of laryngeal neutralization. ms. UCLA.

Steriade, D. (1999). Alternatives to syllabic analysis of consonantal phonotactics. In O. Fujimura, B. Joseph, & B. Palek (eds.), *Proceedings of the 2nd Linguistics and Phonetics Conference*, 1998 (pp. 205-245). Prague: Charles University, Press.

Steriade, D. (2001a). The Phonology of Perceptibility Effects: The P-map and its consequences for constraint organization. Ms. UCLA. (to appear in *A Festschrift for Paul Kiparsky*).

Steriade, D. (2001b). Directional asymmetries in place assimilation: a perceptual account. In E. Hume & K. Johnson (eds.), *The Role of Speech Perception in Phonology*. (pp. 219-250). New York: Academic Press.

Tesar, Bruce. (2002). Enforcing grammatical restrictiveness can help resolve structural ambiguity. *The Proceedings of WCCFL, 21*, 443-456.

Trubetzkoy, N. (1969). Principles of Phonology. C. A. Baltaxe (trans.). Berkeley, CA: University of California Press (Original work published in 1939).

Trask, R. L. (1996). *A dictionary of Phonetics and Phonology*. London: Routledge.

Vitevitch, M. S., Luce, P. A., Charles-Luce, J., & Kemmerer, J. (1997). Phonotactics and syllable trees: Implications for the processing of spoken nonsense words. *Language and Speech*, *40*, 47-62.

Vitevitch, M. S. and Luce, P. A. (1998). When words compete: levels of processing in perception of spoken words. *Psychological Science*, *9*, 325-329.

Vitevitch, M. S. and Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, *40*, 374-408.

Werker, J. and Logan, J. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics*, *37*: 35-44.