

THESE
4.
2007
v.1

This is to certify that the
dissertation entitled

KOREAN SPEAKERS' PRODUCTION OF ENGLISH
CONSONANT CLUSTERS: ARTICULATORY AND
PERCEPTUAL ACCOUNTS

presented by

BO-YOUNG KWON

has been accepted towards fulfillment
of the requirements for the

Ph.D degree in LINGUISTICS



Major Professor's Signature

8-15-2006

Date

MSU is an Affirmative Action/Equal Opportunity Institution

LIBRARY
Michigan State
University

PLACE IN RETURN BOX to remove this checkout from your record.
 TO AVOID FINES return on or before date due.
 MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE
SEP 01 2008		
102308		
122808		

KOREAN SPEAKERS' PRODUCTION OF ENGLISH CONSONANT CLUSTERS:
ARTICULATORY AND PERCEPTUAL ACCOUNTS

VOLUME I

By

Bo-Young Kwon

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Linguistics and Germanic, Slavic, Asian, and African Languages

2006

The

arti

clus

the

adja

acqu

Mod

arti

theo

Go

arti

prod

first

exam

stimu

onset

ABSTRACT

KOREAN SPEAKERS' PRODUCTION OF ENGLISH CONSONANT CLUSTERS: ARTICULATORY AND PERCEPTUAL ACCOUNTS

By

Bo-Young Kwon

The central goal of this dissertation is to investigate the roles and interactions of articulatory and perceptual factors in Korean speakers' production of English consonant clusters. Second language learners' variable degree of production difficulty according to the cluster type has previously been accounted for in terms of sonority distance between adjacent segments. However, I draw attention to the second language consonant cluster acquisition data that have resisted a satisfactory account under the sonority-based model.

As an alternative, I propose a *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM) in which consonant cluster markedness is defined based on the articulatory and perceptual factors associated with each consonant sequence. On the theoretical side, PCCAM draws on insights from Articulatory Phonology (Browman & Goldstein 1986, et seq.) and the P-map (Steriade 2001a, 2001b) to determine how articulatory and perceptual factors influence L2 learners' non-native consonant cluster production.

This dissertation reports on two experiments to test the validity of PCCAM. The first experiment addresses the role of perception in the patterns of vowel insertion by examining Korean speakers' auditory similarity judgments between unepenthesized stimuli and epenthesized stimuli. Results show that an epenthetic vowel in stop + liquid onset clusters is perceptually less obtrusive than an epenthetic vowel in /s/ + C onset

cl

cl

vo

the

exp

Res

13

spe

unde

facto

sensi

patte

cont

*Con

*Con

interl

clusters. In addition, an epenthetic vowel in voiced obstruent + voiced obstruent coda clusters is perceptually less obtrusive than an epenthetic vowel in voiceless obstruent + voiceless obstruent coda clusters. These observations are incorporated into PCCAM in the form of the Perceptual Stridency Effect and Perceptual Voicing Effect. In the second experiment, Korean speakers' production of English consonant clusters was examined. Results show that most of the relative consonant cluster markedness relationships (out of 13 pairwise relative markedness comparisons) in PCCAM are supported in Korean speakers' production of English consonant clusters.

Korean speakers' production of English consonant clusters is, then, formalized under Optimality Theory (Prince & Smolensky 1993). The articulatory and perceptual factors in PCCAM are included in the interlanguage grammar in the form of context-sensitive faithfulness constraints (Steriade 2001a, 2001b). In particular, the attested patterns of vowel insertion and consonant deletion are formalized by relatively ranking context-sensitive faithfulness constraints with respect to the markedness constraint

*Complex. This dissertation shows that the relative position of the markedness constraint *Complex in the constraint ranking successfully captures the types of Korean-English interlanguage grammars in the present study.

Copyright by
Bo-Young Kwon
2006

n

c

in

W

an

co

Sh

infl

Mod

and

statis

many

comf

was a

comm

after I

excited

workin

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Yen-Hwei Lin for her guidance and intellectual inspiration. The term paper that I wrote in her phonology II class has been developed into this dissertation. As you can imagine, the idea was undeveloped and immature at that time, and I did not imagine that it would evolve into my dissertation. What leads me to expand my thoughts and keep searching for the evidence for my argument was Yen-Hwei's vast and profound knowledge on phonetics and phonology, constant encouragement, and unbelievably prompt, yet insightful comments on my draft. She also introduced me to the work on phonetically-based phonology, which crucially influenced my development of the *Phonetically Based Consonant Cluster Acquisition Model* in this dissertation. I was so lucky to have her as my advisor.

I am also honored to have the guidance of Dennis R. Preston, Cristina Schmitt, and Anne Violin-Wigent. I have learned a lot from Dennis about experimental phonetics, statistics, and sociolinguistics. He was always available when I needed his advice. As many who know Dennis would agree, he has the ability to make a student feel comfortable and welcomed, in particular with his delightful jests and warm personality. It was a real pleasure to work with him.

Cristina is especially well versed in language acquisition and provided keen comments on my draft. More importantly, I decided to do my Ph.D studies in linguistics after I took an introductory syntax course with her. I still remember how much I was excited about the beauty of syntax, and I did not mind the lack of sleep when I was working on syntax problem sets. I would like to let Cristina know that I am really

th

cer

pro

wo

Eve

she

con

than

sugg

of li

Gro

invo

deve

of h

heav

mu

I do

thankful to her and her class, which introduced the world of linguistics to me and certainly changed my life.

Anne has contributed to this dissertation by carefully reading the draft and providing meticulous comments on every aspect of this dissertation. She suggested better words, phrases, and sentences and caught a lot of typos and errors that I could not find. Every time I received her comments on my draft, I could see how thorough she was when she was reading my ideas. I really appreciate her time. Anne also provided insightful comments and corrections on the OT analysis part of this dissertation. I would also like to thank Gabriela Alfaraz, who served as an outside reader of my dissertation, for her suggestions and praise on this dissertation.

I would like to acknowledge the financial support I received from the department of linguistics, College of Arts and Letters, and Graduate School. My special thanks go to Grover Hudson, who helped me to find balance between the research project that I was involved with and my own academic work. Over the years that I worked with him, I developed many ideas and research papers. I believe that this was made possible because of his deep generosity and concern for my academic development.

On a personal note, I would like to thank my family, friends, and the one from the heaven above. I believe that they know their contributions to this dissertation and how much I am thankful to them for being my family, friends, and my eternal rest even though I do not write each of their names in here.

LIS

LIS

Cha

1.1.

1.2.

1.3.

1.4.

Cha

2.1.

2.2.

2.3.

2.4.

2.5.

Chap

3.1.

3.2.

3.3.

3.4.

Chap

4.1.

4.2.

4.3.

4.4.

4.5.

4.6.

4.7.

4.8.

4.9.

4.10.

4.11.

4.12.

4.13.

4.14.

4.15.

4.16.

4.17.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. Introduction	1
1.1. Overview of the Dissertation	1
1.2. Overview of the <i>Phonetically Based Consonant Cluster Acquisition Model</i>	5
1.3. Research Questions	12
1.4. Organization of the Dissertation	13
Chapter 2. Issues in L2 Acquisition of Consonant Clusters	16
2.1. Relationship between Consonant Cluster Markedness and Error Rate	17
2.2. Relationship between Speech Style and Error Type (and Error Rate)	19
2.3. Influence of Learners' First Language on Error Type	22
2.4. Patterns of Vowel Insertion and Consonant Deletion	23
2.5. Summary of Chapter 2	23
Chapter 3. Previous Models of L2 Acquisition of Consonant Clusters	25
3.1. Sonority and the Sonority Sequencing Principle	25
3.2. Models of L2 Acquisition of Consonant Clusters	28
3.2.1. Minimal Sonority Distance Parameter	28
3.2.2. Typological Markedness Hypothesis	34
3.3. Empirical Data on L2 Acquisition of Consonant Clusters	39
3.3.1. Consonant Cluster Markedness in MSD and TMH	40
3.3.2. Empirical Data <i>Supporting</i> MSD or TMH	42
3.3.3. Empirical Data <i>Against</i> MSD or TMH	44
3.4. Summary of Chapter 3	51
Chapter 4. Phonetically Based Consonant Cluster Acquisition Model	53
4.1. Articulatory Factors in PCCAM	54
4.1.1. Theoretical Framework: Articulatory Phonology (AP)	56
4.1.1.1. Representation of Sounds in AP	56
4.1.1.2. Gestural Coordination	58
4.1.1.3. Gestural Coordination and an Epenthetic Vowel	61
4.1.1.4. Phonological Acquisition under AP	65
4.1.1.5. AP and L2 Acquisition of Consonant Clusters	69
4.1.2. Degree of Gestural Overlap Parameter	73
4.1.3. Articulatory Release Effect	82
4.1.4. Articulatory Stridency Effect	85
4.1.5. Articulatory Voicing Effect	90
4.2. Perceptual Factors in PCCAM	99
4.2.1. Theoretical Framework: P-map	100

4.3

4.4

4.5.

4.6. S

Chap

5.1. I

5.2. E

5.3. R

5.4. S

Chap

6.1. K

6.2. E

6.3. R

4.2.2. Perceptual Stridency Effect	104
4.2.3. Perceptual Voicing Effect	106
4.2.4. Principle of Perceptual Saliency (Côté 2000)	108
4.2.5. Maximal Perceptual Contrast Parameter	112
4.3. Consonant Cluster Markedness under PCCAM	116
4.3.1. Three-way Computation of Consonant Cluster Markedness	117
4.4. Relative Markedness of Onset Consonant Clusters	124
4.4.1. sp, sk, st < sm, sn and sm sn < sl	126
4.4.2. sl < pl, pr, tr, kl, kr	130
4.4.3. pl, pr, tr, kl, kr < br, bl, dr, gl, gr	135
4.4.4. Summary of the Relative Markedness of Onset Consonant Clusters	137
4.5. Relative Markedness of Coda Consonant Clusters	139
4.5.1. ls, ns < lz, nz	139
4.5.2. ls, ns < lp, lt, lk, np, nt, mp	141
4.5.3. lp, lt, lk, np, nt, mp < lb, ld, nd	144
4.5.4. lp, lt, lk, np, nt, mp < lm	147
4.5.5. ps, ts, ks < sp, st, sk	150
4.5.6. sp, st, sk < pt, kt	155
4.5.7. pt, kt < bd, gd	158
4.5.8. bz, gz < zd	159
4.5.9. zd < bd, gd	160
4.5.10. Summary of the Relative Markedness of Coda Consonant Cluster	162
4.6. Summary of Chapter 4	164
 Chapter 5. Auditory Similarity Judgment	169
5.1. Introduction	169
5.1.1. Hypotheses	173
5.2. Experimental Design	174
5.2.1. Stimuli	174
5.2.2. Procedure	176
5.3. Results	177
5.3.1. Results by Type of C ₁	178
5.3.2. Results by Type of Onset and Coda Cluster	181
5.3.2.1. Results by Type of Onset Clusters	181
5.3.2.2. Results by Type of Coda Clusters	190
5.4. Summary of Chapter 5	194
 Chapter 6. Korean Speakers' Production of English Consonant Clusters	197
6.1. Korean Consonant and Syllable Structure	198
6.2. Experimental Design	201
6.2.1. Participants	201
6.2.2. Materials	202
6.2.3. Procedure	204
6.2.4. Measures	205
6.3. Results	211
6.3.1. Patterns of Vowel Insertion	211

6.4

Ch.

7.1

7.2

7.3

Char

APP

APP

APP

REF

6.3.1.1. Vowel Insertion in Onset Consonant Clusters	212
6.3.1.2. Vowel Insertion in Coda Consonant Clusters	213
6.3.2. Patterns of Consonant Deletion	216
6.3.3. Consonant Cluster Markedness Hierarchy	217
6.3.3.1. Relative Markedness of Onset Consonant Clusters	219
6.3.3.2. Relative Markedness of Coda Consonant Clusters	222
6.4 Summary of Chapter 6	229
 Chapter 7. Formal Analysis	231
7.1. Interlanguage Phonology	234
7.2. Optimality Theory	236
7.2.1. Overview	236
7.2.2. Phonetically-Based Constraints in Optimality Theory	242
7.3. OT Constraints	248
7.3.1. Phonological Acquisition Under OT	248
7.3.2. Patterns of Vowel Insertion in Onset Clusters	251
7.3.3. Patterns of Vowel Insertion in Coda Clusters	258
7.3.4. Patterns of Consonant Deletion in Coda Clusters	264
 Chapter 8. Conclusions	272
 APPENDIX A: Answer Sheet for the Auditory Similarity Judgment Experiment	282
 APPENDIX B: Questionnaire for the Production Experiment	285
 APPENDIX C: Word List for the Production Experiment	288
 REFERENCES	291

LIST OF TABLES

Table 3.1. Proposals on a sonority scale (less sonorant < more sonorant).....	27
Table 3.2. Broselow & Finer (1991): summary of the experiment.....	31
Table 3.3. Errors by cluster type: Korean (Broselow & Finer 1991: 45).....	32
Table 3.4. Scales based on typological markedness.....	34
Table 3.5. Markedness hierarchy in Broselow & Finer and Eckman & Iverson.....	36
Table 3.6. Complexity ranking (Clements 1990: 304).....	37
Table 3.7. Eckman & Iverson (1993): summary of the experiment.....	38
Table 3.8. Consonant cluster markedness hierarchy in MSD and TMH.....	40
Table 3.9. Empirical data <i>supporting</i> MSD or TMH.....	42
Table 3.10. Empirical data <i>against</i> MSD or TMH.....	46
Table 4.1. Vocal tract variables (Browman & Goldstein 1989: 207, 223).....	57
Table 4.2. Degree of gestural overlap in obstruent + obstruent clusters (Byrd 1996)	78
Table 4.3. Summary of clusters acquired by age 3 years.....	80
Table 4.4. Influence of an adjacent consonant on Japanese vowel devoicing.....	86
Table 4.5. Summary of articulatory factors in PCCAM.....	99
Table 4.6. P-map for vowel epenthesis (Steriade 2001b: 18).....	104
Table 4.7. Summary of perceptual factors in PCCAM.....	116
Table 4.8. Consonant cluster markedness hierarchy in PCCAM.....	118
Table 4.9. 75% levels of acquisition for word-initial sC clusters.....	127
Table 4.10. 75% levels of acquisition for word-initial clusters (Smit et al. 1990)....	133
Table 4.11. Relative markedness of onset consonant clusters.....	138

Tab

Tab

Tab

Tab

Tab

Tab

Tab

Tab

Tab

Tab

Tabl

Tabl

Tabl

Tabl

Tabl

Tabl

Tabl

Tabl

Tabl

.....

Tabl

Table 4.12. Evidence for ls, ns < lz, nz (Hansen 2000: 203-206).....	141
Table 4.13. Evidence for ls, ns < lp, lt, lk, np, nt (Hansen 2000: 203-206).....	143
Table 4.14. Evidence for lp, lt, lk, np, nt, mp < lb, ld, nd (Hansen 2000: 203-206) ..	145
Table 4.15. Evidence for lp, lt, lk, np, nt, mp < lb, ld, nd (Hansen 2001: 351-352) ..	146
Table 4.16. Evidence for lp, lt, lk, np, nt, mp < lm (Hansen 2000: 203-206).....	148
Table 4.17. Evidence for lp, lt, lk, np, nt, mp < lm (Hansen 2001: 351-352).....	149
Table 4.18. Summary of sonorant + C clusters with empirical data.....	149
Table 4.19. Developmental order in coda clusters (Templin 1957: 51).....	152
Table 4.20. Evidence for ps, ts, ks < sp, st, sk (Hansen 2000: 203-206).....	154
Table 4.21. Evidence for ps, ts, ks < sp, st, sk (Hansen 2001: 351-352).....	154
Table 4.22. Evidence for ps, ts, ks < sp, st, sk (Wiltshire 2005: 3).....	155
Table 4.23. Evidence for sp, st, sk < pt, kt (Hansen 2000: 203-206).....	157
Table 4.24. Evidence for sp, st, sk < pt, kt (Hansen 2001: 351-352).....	157
Table 4.25. Evidence for bz, gz < zd (Hansen 2001: 351-352).....	161
Table 4.26. Evidence for pt, kt < bd, gd; bz, gz < bd, gd (Hansen 2000: 203-206) ..	162
Table 4.27. Summary of coda consonant clusters with empirical data	163
Table 5.1. P-map for vowel epenthesis (Steriade 2001b: 18).....	170
Table 5.2. Stimuli for the auditory similarity judgment.....	176
Table 5.3. Consonant cluster markedness relationships vs. auditory similarity judgment	196
Table 6.1. Korean consonant chart (Sohn 1999: 153).....	199

Tab

Tab

Tab

Tab

Tab

Tab

Tab

Tab

Tab

Tab

Tab

prod

Tab

Tab

Tab

Tab

Table 6.2. Target Consonant clusters in the production experiment.....	203
Table 6.3. Number of tokens by cluster type.....	210
Table 6.4 Epenthesis by type of C ₁ in onset clusters.....	212
Table 6.5. Epenthesis by type of C ₁ in coda clusters.....	213
Table 6.6. Epenthesis by type of C ₂ in coda clusters.....	215
Table 6.7. Consonant deletion by type of coda cluster.....	217
Table 6.8. Consonant cluster markedness hierarchy in PCCAM.....	218
Table 6.9. Errors in onset clusters.....	219
Table 6.10. Errors in coda clusters (obstruent + obstruent).....	223
Table 6.11. Errors in coda clusters (sonorant + C).....	224
Table 6.12. Coda cluster markedness hierarchy in PCCAM vs. Korean speakers' production.....	225
Table 6.13. Summary: Korean speakers' production of English consonant clusters.....	227
Table 7.1. Epenthesis by type of C ₁ in onset clusters.....	251
Table 7.2. Epenthesis by type of C ₁ in coda clusters.....	258
Table 7.3. Epenthesis by type of C ₂ in coda clusters.....	259

LIST OF FIGURES

Figure 3.1. Sonority dispersion principle.....	37
Figure 4.1. Gestural score of English <i>span</i> [spæn].....	59
Figure 4.2. Examples of temporal relations.....	62
Figure 4.3. Optimal core syllable (Fikkert 1994: 65).....	113
Figure 5.1. Results by type of C _l	179
Figure 5.2. Results by type of onset cluster.....	182
Figure 5.3. Results by type of coda clusters.....	192
Figure 6.1. Waveform and spectrogram of /dret/ with an epenthetic vowel.....	206
Figure 6.2. Extended waveform of an epenthetic vowel in Figure 6.1.....	206
Figure 6.3. Waveform and spectrogram of /tobd/ with an epenthetic vowel.....	207
Figure 6.4. Waveform and spectrogram of /tesk/ with an epenthetic vowel.....	208
Figure 6.5. Waveform and spectrogram of /tesp/ with final stop deletion.....	209
Figure 7.1. Candidate generation and evaluation.....	237

CHAPTER 1

INTRODUCTION

1.1. Overview of the Dissertation

This dissertation is a study of roles that articulatory and perceptual factors play in Korean speakers' production of English consonant clusters. Previous studies on the production of English consonant clusters by second language (L2 hereafter) learners have reported that L2 learners insert a vowel or delete a consonant to break down non-native consonant clusters (e.g., Broselow & Finer 1991; Eckman & Iverson 1993; Hancin-Bhatt & Bhatt 1997; Kim 2000, among others). In particular, it has been reported that certain consonant sequences are more problematic for L2 learners than other sequences. For instance, Broselow & Finer (1991) found that Korean speakers insert an epenthetic vowel more often in voiced stop + liquid clusters than in voiceless stop + liquid clusters.

To account for the difficulties that L2 learners face in the production of non-native consonant clusters, particularly regarding the variable degree of production difficulty according to the cluster type, Broselow & Finer (1991) and Eckman & Iverson (1993) propose distinct consonant cluster acquisition models. The *Minimal Sonority Distance Parameter* proposed by Broselow & Finer (1991) refers to sonority distance between adjacent segments in determining consonant cluster markedness. On the other hand, Eckman & Iverson's model (1993) is based on the typological markedness of segments in a cluster. Note that in L2 research, consonant cluster markedness is assumed to determine learners' production difficulty with non-native consonant sequences: the

me

Da

sof

cof

mov

+ st

acco

erro

seve

few

sequ

child

(Dys

Well

L2 ac

expla

+ stop

referr

¹ Thro

word-1

propo

the wo

part, a

and kn

² Sonor

more marked the cluster is, the more likely L2 learners produce errors. However, Davidson (2001: 32) argued that defining consonant cluster markedness on the basis of sonority is not an adequate predictor of L2 learners' performance on non-native consonant clusters.

With respect to the contradiction between the predictions of sonority-based models and L2 learners' production data, two sets of data are worth highlighting. First, /s/ + stop onset clusters,¹ which are assumed to be the most difficult onset clusters in English according to the sonority-based model (due to sonority reversal), actually induce fewer errors than other onset clusters (Chen 2003; Kim 2000; Kwon 2005; Major 1996). Second, several studies reported that coda clusters with sonority reversal (e.g., /ps, ts, ks/) induce fewer errors than /s/ + stop coda clusters (/sp, st, sk/) which follow the sonority sequencing principle (Major 1996; Hansen 2000, 2001; Wiltshire 2005).² Note also that children seem to acquire stop + /s/ coda clusters earlier than /s/ + stop coda clusters (Dyson 1988; Jakielski 1998; Kirk & Demuth 2003; Olmsted 1971; Templin 1957; Wellman et al. 1931).

In this dissertation, I suggest that the near-failure of the sonority-based model in L2 acquisition of consonant clusters can be attributed to the lack of *phonetically natural* explanations for the consonant cluster markedness. Thus, the relative unmarkedness of /s/ + stop onset clusters may be problematic under the sonority-based model, but when referring to the phonetic factors associated with /s/ + stop onset clusters, it can be

¹ Throughout the dissertation, onset and coda clusters are used as synonyms of word-initial and word-final clusters, respectively. Justification for this is found in Steriade (1999a) where she proposed a *Word-Based Syllable* hypothesis that *syllabic intuitions can be deduced entirely from the word-edge differences*. For instance, "the possibility of parsing a string *abc* into *ab* and *c* is, in part, a function of the similarity between *b* and known word-finals and of the similarity between *c* and known word initials"(p. 5).

² Sonority sequencing principle is defined in § 3.1.

understood why vowel insertion and consonant deletion, two major error types in L2 learners' consonant cluster production, are not likely to occur in /s/ + stop sequences. Consider the phonetic factors accounting for the relative unmarkedness of /s/ + stop clusters.

(1) Phonetic factors associated with /s/ + stop onset clusters³

- a) *Articulatory Stridency Effect*: vowel insertion is less likely to occur after a voiceless strident because the aerodynamic conditions of /s/ remain conducive to a continued turbulence of a strident rather than vocal fold vibration.
- b) *Perceptual Stridency Effect*: vowel insertion is less likely to occur after a voiceless strident because an epenthetic vowel after a voiceless strident is perceptually more obtrusive than an epenthetic vowel after a stop. This is based on the P-map hypothesis proposed by Steriade (2001a, 2001b), which will be discussed in § 4.2.2.
- c) *Maximal Perceptual Contrast*: language learners prefer a sequence where two segments in an alternating sequence make a maximal perceptual contrast. In this sense, a voiceless stop in /s/ + stop is optimal (compared to other C₂ in onset clusters) in that a voiceless stop makes a maximal perceptual contrast with a following vowel. Justification for this is provided in § 4.2.5.

The computation of the three aforementioned phonetic factors determines that /s/ + stop onset clusters are least marked among onset clusters in English.

This dissertation will also demonstrate that the phonetic factors associated with

³ This dissertation does not make any claims about the representation of sC onset clusters (whether /s/ is an appendix). Instead, the aim is to identify the phonetic factors associated with consonant sequences rather than to make claims about how those sequences are prosodically represented in learners' grammars.

consonant sequences efficiently account for the asymmetry between /s/ + stop vs. stop + /s/. The relative unmarkedness of stop + /s/ over /s/ + stop coda clusters is grounded in two phonetic factors: a) degree of gestural overlap and b) the principle of perceptual saliency. Detailed justifications of each phonetic factor are provided in § 4.1 and § 4.2.

As was briefly demonstrated with the cases of a) the relative unmarkedness of /s/ + stop onset clusters and b) asymmetry between /s/ + stop and stop + /s/ coda clusters, the main purpose of this dissertation is to define consonant cluster markedness based on articulatory and perceptual factors associated with each consonant sequence. This model named the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM hereafter) will incorporate up-to-date research on speech production and perception as an analytical tool in L2 consonant cluster acquisition. It is assumed in this dissertation that consonant cluster markedness determines a) L2 learners' production difficulty with non-native consonant sequences and b) L2 learners' repair strategies dealing with a certain type of consonant sequence.

I first test the validity of PCCAM by referring to the published L1 and L2 consonant cluster acquisition data. It will be shown that a fair amount of published L1 and L2 consonant cluster acquisition data supports the consonant cluster markedness hierarchy in PCCAM. However, PCCAM cannot fully be tested by only using previously published data because these data were not intended to test this model. Thus, additional production and perception data were collected for this dissertation to test the validity of PCCAM on L2 acquisition of consonant clusters.

An auditory similarity judgment experiment was designed to collect Korean speakers' judgments on perceptual similarities between unpenthesized stimuli and

epenthesized stimuli. More specifically, this perception experiment tested the validity of the two perceptual principles in PCCAM, the *Perceptual Stridency Effect* and the *Perceptual Voicing Effect*. Thirty-eight native Korean speakers participated in the experiment.

Consonant cluster production data by thirty Korean speakers were also collected and analyzed to determine relative markedness of English consonant clusters in Korean speakers' production. It will be shown that most of the relative consonant cluster markedness relationships (out of 13 pairwise relative markedness comparisons) in PCCAM are supported in Korean speakers' production of English consonant clusters.

The following section presents a brief overview of PCCAM. It focuses on illustrating how the main tenets of Articulatory Phonology (Browman & Goldstein 1986, et seq.) and the P-map (Steriade 2001a, 2001b) are incorporated into the model.

1.2. Overview of the *Phonetically Based Consonant Cluster Acquisition Model*

The *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM) is based on the assumption that phonological markedness may be phonetically grounded. The main tenet of the phonetically based phonological markedness is that there is a connection between markedness patterns and constraints governing the production and perception of speech (Hayes & Steriade 2004: 5). Thus, the trigger for certain phonological processes (e.g., cluster simplification, place assimilation, or lenition) may be found in the demands of articulatory simplification, while the targets of such phonological processes are frequently influenced by perceptual factors.

Consider the case of place assimilation, in particular, characteristically regressive direction of place assimilation (e.g., /anpa/ → [ampa], /amta/ → [anta]). This feature of place assimilation, under the assumption that phonological markedness is phonetically grounded, is accounted for by referring to the asymmetries in the perceptibility of place of articulation in different segments and in different contexts.

More specifically, the basic mechanism of locating the target of place assimilation is that the weaker the perceptual cues of a segment are, the more likely the segment will undergo place assimilation (Jun 2004, Steriade 2001a). For instance, that C₁ is the target of place assimilation in a VC₁C₂V sequence (i.e., regressive assimilation) can be attributed to the observation that the place cues of a consonant are stronger in a prevocalic position than in a preconsonantal position.

From the perspective of the phonetically based markedness, second language learners' modification (i.e., errors) of non-native L2 consonant sequences can be viewed as being derived from learners' articulatory difficulty with producing such a sequence, while the specific contexts targeted by L2 learners (e.g., the site of vowel insertion, or a segment being deleted) are often attributed to perceptual factors. The main goal of this dissertation is to demonstrate that articulatory and perceptual factors, which constitute PCCAM, efficiently account for the patterns of vowel insertion and consonant deletion attested in Korean speakers' production of English consonant clusters.

The frameworks of PCCAM will be constructed based on Articulatory Phonology (Browman & Goldstein 1986, et seq.) and the P-map (Steriade 2001a, 2001b). Note that both Articulatory Phonology and the P-map incorporate phonetic information (articulatory factors for Articulatory Phonology and perceptual factors for the P-map)

into a unitary grammatical system.

Articulatory Phonology (AP hereafter) emerged as an attempt to relate speech activity to more general issues in motor behavior. Drawing upon principles and techniques from general motor behavior such as task dynamics, AP views phonological structures as an organization of the physical actions involved in speaking.

As was made clear in the successive work of Browman & Goldstein, AP provides a very useful vehicle that allows a direct connection between analytical investigations of articulatory structure and linguistic issues. Consider the words [mã^ʔ, de^{dn}, hɪn, ^mbõ, p^hɪn, ba^h, dauⁿ, buã] that a 15-month old girl produced in an attempt to say the word *pen* (Studdert-Kennedy 1987: 76). These words may sound nonsense to an adult's ear and linguistically insignificant, but as Studdert-Kennedy (1987: 76) notes, they can still be analyzed as the set of gestures: glottal opening, bilabial closure, tongue body lowering, alveolar closure, and velum lowering.⁴ That is, by looking at the gestures involved in the word *pen* and comparing those gestures with the 15-month old girl's production, we come to an understanding of what is going on with the 15-month old girl's speech. As Browman & Goldstein (1989: 205) point out, the child seems to know what the relevant gestures are for *pen*, but what he/she lacks is a *proper organization of such gestures*.

It is this aspect of AP that provides strong explanatory power, particularly for the issues of phonological acquisition. The assumption that phonological structures are organized patterns of gestural units leads one to analytically investigate first and second language learners' speech (as child speech is analyzed as a set of gestures). In addition, the assumption leads one to search for articulatory constraints which inhibit language

⁴ Gesture is a unit of representation in Articulatory Phonology. A review of Articulatory Phonology is presented in § 4.1.

learners from correctly producing certain segments or segmental sequences. For instance, we have seen that the seemingly nonsense production of *pen* by a young child results from the fact that the child has not mastered how the gestures associated with the word are supposed to be organized to produce a correct output.

In this dissertation, I propose that we can gain more insights into phonological acquisition when we follow the line of the gestural and analytical approach under AP rather than attributing production failure of first and second language learners to an abstract phonological representation available to language learners at a certain time of phonological development.⁵

PCCAM particularly embraces the aspect of AP where articulatory structures and their organizational mechanisms are referred directly to account for language learners' production data. Thus, L2 learners' articulatory difficulty with non-native consonant sequences is captured through the concept of relative markedness of gestural coordination, the core concept in AP.

Stated in another way, AP is particularly useful in understanding why L2 learners have difficulty producing consonant sequences that do not exist in their L1. AP takes the coordination relationships between adjacent gestures as important aspects of syllable structure. Therefore L2 learners' production difficulty in producing non-native consonant sequences, under the principles of AP, is viewed such that L2 learners may have not yet

⁵ One of the approaches to the acquisition of consonant clusters (both first and second) is to assume that production failure may stem from the fact that language learners do not assign segments to particular prosodic positions. Under this approach, successful acquisition of consonant clusters means that language learners acquire the correct prosodic structure of a consonant sequence at issue. Refer to Goad & Rose (2004) for child acquisition of consonant clusters, which attributes variable reduction in sC consonant clusters to the issue of the prosodification of /s/ in child grammar. Archibald (2003) uses a similar approach to L2 acquisition of consonant clusters in terms of phonological parsing.

mastered the constraints governing coordination relationships in non-native consonant sequences.

Based on the notion of gestural coordination in AP and its consequences in phonological acquisition, PCCAM develops the *Degree of Gestural Overlap* parameter: *a consonant sequence with less gestural overlap is easier for language learners to acquire than a consonant sequence with more gestural overlap*. The degree of gestural overlap parameter is theoretically grounded on the notion that well-formed sequences will be least overlapped and least variable (Browman & Goldstein 1990b; Byrd 1996). Furthermore, PCCAM develops the *Articulatory Stridency Effect*, the *Articulatory Voicing Effect*, and the *Articulatory Release Effect* as additional articulatory factors determining consonant cluster markedness.

Along with the four aforementioned articulatory factors, perceptual factors are considered to account for the patterns of consonant cluster modification. The perception-based accounts of vowel insertion and consonant deletion developed in this dissertation are in line with a growing body of work emphasizing the role of perception in phonological processes (Boersma 1998; Côté 2000, 2004; Steriade 1999a, 1999b, 2001a, 2001b; Wilson 2001; Wright 1996, 2004, among others). The central assumption in this dissertation is that L2 learners are guided by certain perceptual principles when dealing with consonant clusters that do not exist in their L1. In particular, I refer to the P-map hypothesis proposed by Steriade (2001a, 2001b). The P-map is a set of statements about perceived distinctiveness-differences between contrasts in various contexts.

According to the P-map hypothesis, when faced with a phonotactic constraint, speakers prefer “the least distinctive contrast whose modification resolves the violation”

(Steriade 2001b: 14). Applying the P-map hypothesis to L2 phonological acquisition, I suggest that the P-map knowledge guides L2 learners to choose the output form that differs minimally from the original form (i.e., input of the target language), while resolving the phonotactic constraints in the L2.

Based on the main tenet of the P-map that an alternation that involves a less perceptible change is more likely to take place than an alternation that involves a more perceptible change, PCCAM develops two perceptual principles, the *Perceptual Stridency Effect* and the *Perceptual Voicing Effect*. PCCAM also incorporates the *Principle of Perceptual Saliency* (Côté 2000) and the *Maximal Perceptual Contrast* as additional phonetic factors determining consonant cluster markedness.

As a formal analysis of the data, Optimality Theory (Prince & Smolensky 1993; OT hereafter) is adopted. Under OT, generator (GEN) creates a large set of possible structural changes (i.e., candidates) for optimal output, and the issue of which structural change is optimal is determined by the relative ranking of the Markedness constraints with respect to Faithfulness constraints. When it comes to Korean speakers' production of English consonant clusters, the markedness constraint *COMPLEX (no complex margins are allowed) interacts with faithfulness constraints MAX-IO and DEP-IO. MAX-IO requires that for every output segment, there is an input segment corresponding to it; deletion of a segment is not allowed. On the other hand, DEP-IO is violated if the output has a segment that lacks a correspondent in the input; insertion of a segment in the output is not allowed.

Yet, the basic markedness and faithfulness constraints are not sufficient to fully capture the patterns of vowel insertion and consonant deletion because the site of vowel

insertion and the segment being deleted crucially depend on the syntagmatic context where the segment occurs. Thus, the formal analysis in Chapter 7 develops context-sensitive faithfulness constraints which differently penalize MAX-IO and DEP-IO violation depending on the syntagmatic context.

Consider the case of word-final obstruent devoicing to illustrate how context-sensitive faithfulness constraints are projected from phonetic observation. Steriade (2001b) proposed two context-sensitive faithfulness constraints IDENT [$\pm$ nas]/ V _] and IDENT [$\pm$ voice]/ V _] to account for the word-final obstruent devoicing. IDENT [$\pm$ nas]/ V _] militates against a nasal feature change in word-final position. On the other hand, IDENT [$\pm$ voice]/ V _] militates against a voicing feature change in word-final position. According to Steriade (2001b), the grammar ranks higher the context-sensitive faithfulness constraint projected from the more perceptible contrast between input and modified output. Following this line of reasoning, IDENT [$\pm$ nas]/ V _] ranks higher than IDENT [$\pm$ voice]/ V _] because obstruency difference in word-final position is more perceptible than voicing difference in the same position (Greenberg & Jenkins 1965).⁶

I develop, following Steriade (2001a, 2001b), context-sensitive faithfulness constraints which differently penalize the violation of faithfulness constraints. Thus the attested patterns of vowel insertion and consonant deletion are incorporated into a grammar by relatively ranking context-sensitive faithfulness constraints with respect to the markedness constraint *COMPLEX.

The goals of this dissertation are summarized as follows. First, this dissertation aims to construct the *Phonetically Based Consonant Cluster Acquisition Model*

⁶ Steriade's analysis on word-final obstruent devoicing will also be discussed in § 7.2.

(PCCAM) in which consonant cluster markedness is defined based on articulatory and perceptual factors associated with consonant sequences. Second, this dissertation tests predictions of PCCAM on the relative markedness of English consonant clusters through Korean speakers' production of English consonant clusters as well as through the published L1 and L2 consonant cluster acquisition data. Third, this dissertation aims to show how the articulatory and perceptual factors in PCCAM are included in the interlanguage grammar and how Korean speakers' production data are formalized under Optimality Theory (Prince & Smolensky 1993).

The specific research questions for the dissertation are presented in § 1.3.

1.3. Research Questions

Research questions:

- 1) What are the articulatory factors resulting in vowel insertion and consonant deletion in Korean speakers' production of English consonant clusters?
- 2) What are the perceptual factors resulting in vowel insertion and consonant deletion in Korean speakers' production of English consonant clusters?
- 3) How can we define consonant cluster markedness that will determine a) language learners' variable degree of production difficulty with non-native consonant sequences and b) learners' repair strategies when dealing with a certain type of consonant sequence based on articulatory and perceptual factors?
- 4) Does the *Phonetically Based Consonant Cluster Acquisition Model* have an explanatory power beyond Korean speakers' production of English consonant

1.4

The

ove

cor

bet

var

pro

con

the

Mar

acqu

son

⁷ Eck
of an
occurs
Eckm.
used
Eckm.

clusters? In other words, can the *Phonetically Based Consonant Cluster Acquisition Model* be applied to L2 learners with other first language backgrounds and first language acquisition of consonant clusters?

- 5) How can we formalize the articulatory and perceptual factors that define consonant cluster markedness in the grammar?

1.4. Organization of the Dissertation

This section provides a roadmap to the dissertation. Chapter 2 presents a general overview of the studies on L2 acquisition of consonant clusters. The major issues of L2 consonant cluster acquisition to be reviewed in the chapter include a) the relationship between consonant cluster markedness and error rate, b) the influence of stylistic variation on the production of consonant clusters, c) the relationship between L2 learners' production and first language transfer, and d) the patterns of vowel insertion and consonant deletion in L2 learners' production of non-native consonant clusters.

Chapter 3 reviews two previous models of L2 acquisition of consonant clusters, the *Minimal Sonority Distance Parameter* (Broselow & Finer 1991) and the *Typological Markedness Hypothesis* (Eckman & Iverson 1993).⁷ These previous models of L2 acquisition of consonant clusters mainly rely on the notion of sonority, for instance, the sonority distance between the two segments in a cluster (the bigger the sonority distance

⁷ Eckman & Iverson (1993) define typological markedness based on the implicational universals of an occurrence of segments (e.g., the occurrence of a voiced stop in a language implies the occurrence of a voiceless stop). Refer to § 3.2.2 for the definition of typological markedness in Eckman & Iverson (1993). Also note that the *Typological Markedness Hypothesis* is not the label used by Eckman & Iverson (1993). I invented this term, for the sake of convenience, to refer to Eckman & Iverson's model.

be

pa

in

Cl

co

M

Ph

of

resp

mar

mar

sect

PCC

rela

mec

judg

justi

mark

PCC

spea

between the two segments, the easier it is to acquire; Minimal Sonority Distance parameter) or the sonority contour of consonant clusters (e.g., the steeper the sonority rise in onset clusters, the easier it is to acquire; Sonority Dispersion Principle). I will show in Chapter 3 that these sonority-based models of consonant cluster acquisition teem with conflicting data (e.g., Chen 2003; Davidson 2001, 2003; Hansen 2000, 2001; Kim 2000; Major 1996; Osburne 1996).

Chapter 4 presents a new model of L2 consonant cluster acquisition, the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM). The core principles of PCCAM, Articulatory Phonology, and the P-map, are reviewed in § 4.1 and § 4.2, respectively. The articulatory and perceptual factors defining consonant cluster markedness in PCCAM are also identified in this section. § 4.3 introduces the way markedness is computed in PCCAM (i.e., three-way computation of markedness). This section also presents the English consonant cluster markedness hierarchy generated by PCCAM. § 4.4 and § 4.5 are devoted to *empirically* justifying the relative markedness relationships in PCCAM based on the published L1 and L2 acquisition data.

To provide more direct empirical evidence for the role of the perceptual mechanism (i.e, P-map) in L2 phonology, Chapter 5 presents an auditory similarity judgment experiment. The auditory similarity judgment experiment provides justifications for the two perceptual factors defining the relative consonant cluster markedness (i.e., the perceptual voicing effect and perceptual stridency effect) in PCCAM. We can then proceed to test the validity of PCCAM as a whole through Korean speakers' production of English consonant clusters (chapter 6).

Chapter 6 presents Korean speakers' production of English consonant clusters. I

first discuss Korean syllable structure and the Korean consonant system to provide background for subsequent chapters (§ 6.1). § 6.2 presents the research design of the production experiment. Patterns of vowel insertion and consonant deletion attested in Korean speakers' production of English consonant clusters are presented in § 6.3.1 and § 6.3.2. Based on a count of vowel insertion, consonant deletion, and metathesis in Korean speakers' production of English consonant clusters, the relative consonant cluster markedness is established, and its ranking relations are compared with the predictions of PCCAM in § 6.3.3.

Chapter 7 provides a formal analysis of Korean speakers' production of English consonant clusters. § 7.1 defines the term "interlanguage phonology" and discusses the position taken by the present study on the issue of whether the basic mechanisms governing interlanguage phonology are similar in important ways to those of a first language. § 7.2 presents the main tenets of Optimality Theory and discusses how OT improves the way that phonetically-based constraints are expressed in a formal system. § 7.3 provides specific constraints which embody phonetic factors in PCCAM and provides formal analyses of the patterns of vowel insertion and consonant deletion found in Korean speakers' production of English.

Chapter 8 concludes the dissertation by summarizing the findings and analyses of L2 acquisition of English consonant clusters. Directions for future research are also suggested in Chapter 8.

CHAPTER 2

ISSUES IN L2 ACQUISITION OF CONSONANT CLUSTERS

The literature is replete with studies of second language acquisition of consonant clusters. A variety of native language groups have been involved in research (e.g., Arabic, Cantonese, Japanese, Korean, Mandarin Chinese, and Spanish, to name a few), and a range of hypotheses have been tested. Regarding the target language, although the acquisition of *English* consonant clusters occupies the major portion of the literature, several studies have tested how L2 learners acquire Czech, Polish, German, or Swedish consonant clusters.

In this chapter, I provide an overview of four major issues related to how L2 learners acquire the consonant sequences of their second language: a) the relationship between consonant cluster markedness and error rate, b) the relationship between speech style and error type (or error rate), c) the influence of the learners' first language on error type, and d) patterns of vowel insertion and consonant deletion in consonant cluster modification.

Note that this chapter is intended to give a general background of the studies on L2 acquisition of consonant clusters. Many of these studies are peripheral to the focus of this dissertation; therefore I will only present a brief summary of these studies in this chapter.

2.1. Relationship between Consonant Cluster Markedness and Error Rate

One of the hypotheses frequently tested in L2 acquisition of consonant clusters is that the more marked the cluster is, the more likely L2 learners are to have difficulty producing the cluster (Anderson 1987; Broselow & Finer 1991; Davidson 2001, 2003; Davidson et al. 2004; Eckman 1991; Eckman & Iverson 1993; Hancin-Bhatt & Bhatt 1997; Kim 2000). Some studies have tested whether the hypothesized order of consonant cluster markedness holds true in interlanguage (IL) phonology.⁸ For instance, Broselow & Finer (1991) investigated the mastery of particular English onset clusters (/pj, bj, fj, pr, br, fr/) by native speakers of Japanese and Korean. It was assumed that the clusters whose components are closer on the sonority index⁹ are more marked than clusters whose components are further apart on the index. The researchers then tested whether the hypothesized order of consonant cluster markedness (for instance, Cj is less marked than Cr) is related to L2 learners' difficulty with these clusters. The results showed that sonority distance between segments plays a role in determining consonant cluster markedness: Korean speakers demonstrated the most difficulty with producing /fr/ clusters.

Davidson (2001) provides different views on the role of sonority in determining consonant cluster markedness. She tested English speakers on their production of Polish words containing legal Polish clusters (/kp, kt, pt, tʃk, vz, tf, dv, zm, vn, sm, sn, zr, ʃl, ʃr, fr/). In the experiment, English speakers were asked to read aloud a sentence containing a

⁸ Chapter 3 discusses how the previous studies define consonant cluster markedness.

⁹ The sonority index (Broselow & Finer 1991: 38) assigns a sonority value for each class of sounds as follows: stops = 1, fricatives = 2, nasals = 3, liquids = 4, glides = 5

Polish word after hearing it produced by the Polish speaker.¹⁰

Based on the production results, Davidson (2001) concludes that sonority sequencing alone cannot account for the data. For instance, clusters like /zm/ and /vn/ do not differ in sonority distance, but English speakers produced /zm/ far more accurately (63% accuracy) than /vn/ (11% accuracy). The question then remains: what factors contribute to English speakers' performance difference between /zm/ and /vn/. Davidson (2001) finds a possible answer to this question in the relative markedness of /v/ over /z/: /z/ is [+coronal], whereas /v/ is a more marked labial fricative.

Based on the assumption that longer clusters are more marked than shorter clusters, some studies have examined the effect of consonant cluster length (CC vs. CCC) on consonant cluster markedness. For instance, Anderson (1987) observed that correct performance of both Arabic and Chinese speakers decreased significantly as English final clusters increased from two to three consonants. The effect of consonant cluster length was also reported by Carlisle (1998). He found that when tested on initial /sk/, /skr/, /sp/, and /spr/ clusters at two different stages of learning, 10 native Spanish speakers learning English demonstrated a greater rate of accuracy in producing the two-member consonant clusters than three-member consonant clusters at both times.

To summarize, L2 production data on non-native consonant clusters converges on the observation that the relative markedness of consonant clusters is correlated with L2 learners' difficulty with producing such clusters. However, the results of Davidson (2001) clearly show that the definition of consonant cluster markedness is still an issue of controversy. This issue of defining consonant cluster markedness is further elaborated in

¹⁰ To tease apart perception from production, English speakers were presented the target words both auditorily and with English-like orthography simultaneously.

C

2.

A

is

sy

19

cor

con

bec

wor

incr

part

eper

is lo

diff

term

princ

greate

(Wein

Chapter 3.

2.2. Relationship between Speech Style and Error Type (and Error Rate)

Among the robust findings to emerge from L2 learners' consonant cluster production data is that speech style (i.e., task type during data collection) influences learners' choice of syllable simplification strategies (Bayley 1996; Lin 2001, 2003; Major 1987; Weinberger 1987, 1994).

Lin (2001) tested 20 Chinese learners on their production of English onset consonant clusters in four different task types: minimal pairs, word lists, sentences, and conversation. It was observed in the study that the use of epenthesis increases as the task becomes more formal (epenthesis rate: conversation 4.38% < sentences 11.38% < wordlist 27.5% < minimal pairs 35.13%), whereas the use of deletion and replacement increases as the task becomes less formal. Lin (2003) replicated her 2001 study with 40 participants and found a similar pattern of epenthesis/deletion ratio: the proportion of epenthesis to deletion is highest in the minimal pairs, while the epenthesis/deletion ratio is lowest in sentences and conversation.

The question arises as to how the different degree of task formality leads to different syllable simplification strategies. Lin (2001, 2003) accounts for this pattern in terms of the principle of recoverability proposed by Weinberger (1987, 1994). The principle of recoverability states that “the proportion of epenthesis to deletion should be greater in tasks without linguistic context than in tasks with linguistic context” (Weinberger 1987: 408). The assumption underlying the recoverability principle is that

in t

rea

pos

the

inte

epet

to c

aud

lear

Chi

was

to of

t d in

Bas

spee

that

spee

Pres

patter

in Li

in the tasks that do not supply linguistic context such as minimal pairs and wordlist reading, deletion of a segment is less likely to occur because deletion causes a high possibility of ambiguity. By the same logic, deletion can be more frequently employed in the tasks with linguistic context such as paragraph reading and conversation because the interlocutor can recover what has been lost through deletion from the linguistic context.

Following the notion of recoverability, Lin (2001, 2003) proposed that the use of epenthesis increases in a more formal task that requires more attention to form rather than to content because learners hypercorrect their speech to make the target consonant audible. In contrast, more instances of deletion and replacement are induced when learners pay more attention to content rather than to form.

Speech style-dependent variation is also observed in Bayley's (1996) study on Chinese speakers' acquisition of an English variable phonological rule, -t/d deletion. It was found in the study that the more careful the style, the less likely Chinese speakers are to omit -t/d. More specifically, Chinese learners of English are least likely to delete final -t/d in reading style and most likely in the relatively informal conversation in interviews. Based on these results, Bayley (1996) concludes that the -t/d variable is sensitive to speech style in both native and non-native speech. Bayley's (1996) study typically shows that a) interlanguage displays the same type of style shifting found in first language speech, and that b) greater accuracy is obtained as style becomes more formal (refer to Preston 1989 for discussion of this topic).

When it comes to the relationship between speech style and error rates, the patterns are not as straightforward as that of speech style and error types (as we have seen in Lin 2001, 2003 and Weinberger 1987, 1994). That is, if we assume that greater

acc

inc

ap

pro

sig

Sin

Eng

stor

wor

23.8

in it

sent

(p. c

Cor

(20

stud

acc

tha

epe

for

accuracy is obtained as style becomes more formal, we would expect that error rates increase as the speech style becomes less formal. However, previous studies provide apparently conflicting data on the relationship between speech style and error rates.

For instance, Major (1999) tested four Brazilian learners of English on their production of English onset and coda consonant clusters and found that there were significantly more target-like occurrences in paragraph reading than in wordlist reading. Similarly, Weinberger (1987: 410) found that Chinese learners of English produced English word-final consonants and consonant clusters more accurately in less formal story-telling than in more formal paragraph reading, while the error rate was lowest in wordlist reading (error rate: wordlist 16.8% < story-telling 19.8% < paragraph reading 23.8%).

Lin (2001) also noticed that speech style does not explain much about error rates in itself. The observed order of error rates by speech style in Lin (2001) was as follows: sentence reading (fewer errors) < word list reading < conversation < minimal pair reading (p. 696). There seems to be no effect of task formality on error rates in Lin (2001). Comparing the results of her own study with Major (1999) and Weinberger (1987), Lin (2001: 703) argues that “such obvious differences in the error rate patterns of the three studies strongly support my hypothesis that, for consonant clusters, what varies in accordance with task formality is the strategies utilized rather than the error rates”.

To summarize, the finding regarding the influence of speech style on error type is that the use of more formal tasks during data collection leads to an increase in vowel epenthesis, while the percentage of deletion and replacements becomes higher in less formal tasks. The influence of speech style on error rate also demonstrates an interesting

pattern. Although it is generally assumed that greater accuracy is obtained by L2 learners as style becomes more formal, previous studies on L2 acquisition of consonant clusters have not found this straightforward pattern in L2 learners' production.

2.3. Influence of Learners' First Language on Error Type

Earlier studies of L2 acquisition of consonant clusters seem to focus on the question of whether errors found in L2 learners' production is a result of language transfer or of universal principles of markedness. For instance, Sato (1984) tested whether preference of repair strategies (e.g., deletion of a consonant vs. insertion of an epenthetic vowel) can be attributed to the syllable structure of the learners' native language. Sato (1984) observed that Vietnamese learners of English prefer deletion of a consonant when they produce English coda consonant clusters and interpreted the Vietnamese speakers' strategy as a result of language transfer. Vietnamese has a large proportion of CVC syllables in the lexicon; thus they prefer to delete one segment in a cluster which renders a CVC sequence.

However, Tarone (1987) takes a rather different view on why second language learners prefer certain repair strategies. In a study involving Cantonese, Portuguese, and Korean speakers, Tarone (1987: 239) found that Cantonese and Korean speakers prefer consonant deletion, while the Portuguese speakers favor epenthesis as a strategy in simplifying syllable structure. Based on these results, Tarone (1987) concludes that both first language transfer and universal principles of markedness play a role. That is, the dominant process influencing the syllable structure of IL phonology appears to be

language transfer, while a preference for the open (CV) syllable also seems to operate as a process independent of language transfer.

2.4. Patterns of Vowel Insertion and Consonant Deletion

Some studies have been aimed at specifically identifying the patterns of vowel insertion and consonant deletion in L2 learners' production of non-native consonant clusters (Davidson 2001, 2003, 2006; Davidson et al. 2004; Hancin-Bhatt & Bhatt 1997; Hansen 2001; Kim 2000; Tropic 1987). The findings of these studies can be summarized as follows: a) more sound modifications occur in coda clusters than in onset clusters (Hancin-Bhatt & Bhatt 1997: 345; Kim 2000: 207), b) vowel epenthesis is more likely to occur word-initially, whereas deletion is more likely to occur word-finally (Hancin-Bhatt & Bhatt 1997: 345), c) the less sonorant segment in a cluster is usually deleted (Hansen 2001: 356; Tropic 1987: 185), and d) vowel insertion is more likely to occur after voiced obstruents than after voiceless obstruents (Davidson 2003: 48; Kim 2000: 138).

2.5. Summary of Chapter 2

In Chapter 2, we have seen that previous studies have dealt with various aspects of L2 acquisition of consonant clusters such as native language transfer, consonant cluster markedness, stylistic variation, and patterns of vowel insertion and consonant deletion. We have also noticed that different interpretations of the results are possible depending on the perspective of the researchers.

What most attracts our attention in these previous studies is that the errors that occur when second language learners produce non-native consonant clusters are by no means random. That is, we have seen that not all combinations that are absent in the learners' native language cause the same degree of difficulty. Clear asymmetries are observed depending on the stylistic and linguistic context.

The notions of recoverability, sonority distance, and typological markedness and a number of hypotheses have been proposed in the literature to account for the variability in L2 acquisition of consonant clusters, including the patterns of vowel insertion and consonant deletion. Nevertheless, the question of the validity of these explanations still remains.

Chapter 3 takes up this issue, while reviewing two previous models of L2 acquisition of consonant clusters, namely the *Minimal Sonority Distance parameter* (Broselow & Finer 1991) and *Typological Markedness Hypothesis* (Eckman & Iverson 1993).

CHAPTER 3

PREVIOUS MODELS OF L2 ACQUISITION OF CONSONANT CLUSTERS

In this chapter, I review two previous models of L2 acquisition of consonant clusters, Broselow & Finer's (1991) *Minimal Sonority Distance Parameter* (MSD hereafter) and Eckman & Iverson's (1993) *Typological Markedness Hypothesis* (TMH hereafter).

A review of these two models of L2 consonant cluster acquisition will give a general background on how L2 consonant cluster production data have been accounted for in the literature. More importantly, it will be shown that these previous models are not adequate predictors of L2 learners' performance on non-native consonant clusters.

Postulating a new consonant cluster acquisition model, the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM), is thus justified throughout the chapter.

In the following section, I define two major concepts underlying MSD and TMH, namely, the *Sonority* and the *Sonority Sequencing Principle*. Even though MSD and TMH make different claims about how consonant cluster markedness should be defined, both models make crucial use of *Sonority* and the *Sonority Sequencing Principle*. Thus, it is necessary to define these two core concepts before introducing each model in detail.

3.1. Sonority and the Sonority Sequencing Principle

The number of consonants allowed to occur in syllable margins (i.e., onset and coda) differs across languages. Languages like Japanese, Korean, and Mandarin Chinese allow only one consonant in syllable margins, while languages like English, German, and

Swedish allow two or three consonants in each margin.¹¹

It has long been noted that there are constraints on the sequences of segments that can occur in a syllable. Many researchers have proposed that the position of a segment within a syllable is driven by the principle of sonority. This principle is referred to in the phonological literature as the *Sonority Sequencing Principle* (Clements 1992; Hooper 1976; Selkirk 1984; Steriade 1982, among others).

The *Sonority Sequencing Principle* (SSP hereafter) requires that the most sonorous segment constitutes the syllable peak, with all other segments organized around it in such a way that the relatively more sonorous segments are closer to the peak and the relatively less sonorous ones are farthest away from it. It has been claimed that SSP has explanatory power in accounting for a number of cross-linguistic tendencies about the distribution and sequencing of segments.

Despite the fairly uncontroversial role of SSP in defining permissible syllable margins, the question of what sonority is and how it can be defined still remains a highly controversial issue. Sonority is sometimes defined in terms of the amount of airflow in the resonance chamber (Bloch & Trager 1942; Goldsmith 1995), while some argue that a phonetic parameter such as perceptual salience or loudness of a particular sound should be incorporated to define sonority (Ladefoged 2001). Strong criticism of the notion of sonority is found in Ohala & Kawasaki-Fukumori (1997). The claim is that neither the parameters that are supposed to define sonority nor the notion of a syllable, which the sonority hierarchy is assumed to shape, have been defined in a way that could be

¹¹ Depending on the view on the status of glides in the language, some researchers may argue that Korean and Mandarin Chinese allow obstruent + glide onset clusters. This dissertation does not make any specific claims on the status of glides in these languages and assumes that the aforementioned languages do not allow any type of complex margins.

empirically tested (Ohala & Kawasaki-Fukumori 1997: 344). Note that there has been a recent attempt to define the sonority hierarchy and an optimal syllable structure in a phonetically natural way. The reader is referred to Wright (2004).

Several different proposals on the sonority scale have been put forward as shown in Table 3.1.

Table 3.1. Proposals on a sonority scale (less sonorant < more sonorant)

Source	Sonority Scale
Universal Sonority Scale	
Clements (1990)	obstruents < nasals < liquids < glides < vowels
Butt (1992)	vl obstruents < vd obstruents < nasals < liquids < glides < vowels
Selkirk (1984)	p, t, k < b, d, g < f, θ < v, z, ð < s < m, n < l < r
Blevins (1995)	for non-syllabic segments, voiceless plosives < voiced plosives < vl fricatives < vd fricatives < nasals < liquids < glides < non-low vowels < low vowels
Language-Specific Sonority Scale	
Steriade (1982)	p, b, k, g < t, d < f < s < m < n < l < r for Classical Latin; p, t, k < b, d, g < s < z < n, m < l < r for Attic Greek (for non-syllabic segments)
Davis (1990)	vl stops < vd stops < non-coronal fricatives < coronal fricatives < n < m < liquids < vowels for Italian

Note: vl = voiceless; vd = voiced

For instance, Clements’s (1990) sonority scale consists of five major natural classes of sounds and ranks obstruents, nasals, liquids, glides, and vowels in order of increasing sonority. On the other hand, Butt (1992) divides voiceless obstruents from voiced obstruents and considers voiced obstruents more sonorous than voiceless obstruents. Selkirk (1984) and Blevins (1995) make even finer distinctions on a sonority scale as seen in Table 3.1.

Finally, Steriade (1982) and Davis (1990) provide language-specific sonority scales. Steriade (1982) proposes that languages have a certain degree of freedom in the

assignment of sonority values to their segments. In particular, she states that “the onset rule of these languages (Attic, Mycenaean Greek, and Sanskrit) can be allowed to freely iterate on its own output: the number of permissible iterations per syllable - and therefore the resultant number of consonants in any given onset - will be restricted by a language specific constraint on the relative sonority of adjacent segments” (Steriade 1982: 80).

It is beyond the scope of this dissertation to discuss controversies over sonority and sonority scales in further detail. However, it should be noted that many of the L2 studies which make use of sonority and a sonority scale do not provide any accounts for why they choose one type of a sonority scale over others. For instance, Broselow & Finer (1991: 38) adopted the following sonority scale, stops < fricatives < nasals < liquids < glides, but provided no explanations for why they chose this specific scale.

In the following sections, I introduce two previous L2 consonant cluster acquisition models, MSD (Broselow & Finer 1991) and TMH (Eckman & Iverson 1993).

3.2. Models of L2 Acquisition of Consonant Clusters

3.2.1. Minimal Sonority Distance Parameter

The *Minimal Sonority Distance Parameter* (MSD) was originally proposed to explain language-specific patterns of consonant clustering (Harris 1983; Selkirk 1982; Steriade 1982). The basic idea of MSD is that sonority distance of allowable clusters characterizes the various types of consonant clusters allowed in different languages. For instance, Harris (1983: 21) proposes an *Onset Rule* in Spanish as follows.

(1) Onset rule (Spanish)

Construct a maximally binary branching tree of category O(nset) whose branches dominate [+consonantal] segments that are not adjacent on the universal sonority scale.

The onset rule in Spanish states that segments in a branching onset should not be adjacent on the universal sonority scale. This stipulation is to account for the fact that in Spanish only obstruent + liquid clusters constitute well-formed two segment onsets, while other possible onset clusters such as obstruent + nasal and nasal + liquid are not allowed. Refer to the *Sonority Index* in (2) to see how the onset rule works in Spanish.

(2) Sonority index (Broselow & Finer 1991: 38)

class	value
stops	1
fricatives	2
nasals	3
liquids	4
glides	5

It is well known that stops and fricative are combined into one category, i.e., obstruents. We can see that obstruents are adjacent to nasals in the sonority index, which explains the consonant cluster phonotactics of Spanish.

As was demonstrated with the onset rule in Spanish, the sonority index is designed to determine the sonority distance of allowable clusters. In the sonority index in (2), the most sonorous consonants, *glides* are assigned the sonority value of 5, whereas *stops* are assigned the value of 1. MSD states that, other factors being equal, languages

that require a greater difference in sonority between adjacent segments will have fewer kinds of consonant clusters in the onset. By the same reasoning, languages that allow a smaller minimal difference between adjacent consonants will have more kinds of consonant clusters in the onset.

For instance, a language with an MSD setting of 5 allows only a single consonant as an onset (there are no two consonants on the sonority index that differ in sonority by 5). A setting of 4 allows stop + glide clusters as well as a single consonant onset, and a setting of 3 allows any cluster that has at least a difference of 3 (e.g., stop + liquid and fricative + glide clusters) as well as all the onset types allowed by settings of 4 and 5. Note that the most restrictive parameter setting (MSD = 5) generates the least marked set of elements: no clusters.

Now that the principle of MSD has been explained, let us examine how the working mechanism of MSD applies to L2 consonant cluster acquisition. Consider how Broselow & Finer (1991) make use of MSD in making predictions on L2 learners' consonant cluster production.

Broselow & Finer (1991: 38) share the notion of the MSD parameter such that languages may vary in the degree of difference they require between the values of adjacent tautosyllabic segments, and this difference presumably reflects different settings of the multi-valued MSD parameter. Based on this assumption, Broselow & Finer propose that by comparing the MSD value of the target language and the learners' native language, it is predictable which clusters will be problematic for L2 learners. More specifically, Broselow & Finer propose that onset sequences with a lower MSD setting than the learner's first language (L1) would be difficult to acquire.

Suppose a learner's native language has an MSD value of 4. Any combination of consonants which generates an MSD value equal to or lower than 3 would be problematic for the learner. It is also assumed that the lower the MSD value of a consonant sequence, the more difficult it is to acquire that sequence. Thus, for those learners whose native language MSD value is 4, a consonantal sequence of an MSD value 1 is even more difficult than that of MSD value 2 or 3.

To empirically test the relevance of the MSD parameter to L2 acquisition, Broselow & Finer examined Japanese (n = 8) and Korean speakers (n = 24) on their production of English onset clusters.¹² Table 3.2 provides a summary of Broselow & Finer (1991).

Table 3.2. Broselow & Finer (1991): summary of the experiment

Broselow & Finer (1991)	
Analytical tool	MSD parameter
Participants	Japanese (n = 8)
	Korean (n = 24)
L2 proficiency	High intermediate
Stimuli	/pj, bj, fj, pr, br, fr/ (onset clusters only)
Sonority scale	stops < fricatives < nasals < liquids < glides
	less marked < more marked
Predictions	a) Cj < Cr (pj < pr)
	b) pC < bC < fC (pr < br < fr)
Results	Predictions were generally supported.

In addition to testing the main prediction of MSD that clusters closer in sonority are harder to learn than clusters containing consonants separated by a wider sonority

¹² The experimental procedure in Broselow & Finer (1991) was as follows. Participants were told that they would learn some English vocabulary and were presented with an audio tape consisting of a total of 34 sentences, each giving a definition of the new 'word' (either real English words or nonsense words) (e.g., 'Puce' [pju:s] is a colour). Participants first heard each defining sentence twice and were asked to repeat the sentence twice. To tease apart production and perception as sources of error, participants were also provided with a sheet giving the word both in English spelling and in IPA transcription.

difference, Broselow & Finer were interested in whether the voiceless-voiced and stop-continuant distinctions are necessary in the sonority hierarchy (p. 39). Thus, they tested language learners' relative mastery of stop + sonorant vs. fricative + sonorant (stop-continuant distinction) and voiceless obstruent + sonorant vs. voiced obstruent + sonorant (voiceless-voiced distinction).

Broselow & Finer (1991: 41) predict that obstruent + glide clusters are less marked than obstruent + liquid clusters, and the obstruent clusters increase in markedness in the order of /pC/ (less marked) < /bC/ < /fC/. Thus, among the clusters tested in the study, /pj/ should be the least difficult cluster type for learners, while /fr/ should be the most difficult cluster.

Table 3.3 presents the results from the Korean participants in this experiment. The last two lines in Table 3.3 present the number of errors in each category. We can see that the number of errors by category corresponds to the predictions made by Broselow & Fine (refer to Table 3.2).

Table 3.3. Errors by cluster type: Korean (Broselow & Finer 1991: 45)

	pj	pr	bj	br	fj	fr
Total errors /n	1/288	1/287	5/288	11/288	13/288	18/286
% of total errors	2	2	10	22	27	37
Epenthesis (CCV → CVCV)	0	0	1	8	0	0
Deletion (CCV → CV)	1	1	4	3	5	6
Total	1	1	5	11	5	6
Errors in manner	0	0	0	0	8 (p)	11 (p)
Initial replacement						1 (b)
Medial replacement	0	0	0	0	0	0
Total	0	0	0	0	8	12
	Cj	Cr	pC	bC	fC	
Total errors (n)	19	30	2	16	31	

The predictions made by MSD were confirmed in Broselow & Finer (1991) in terms of

error rates; however, one of the limitations of MSD seems to lie in the fact that it does not make any predictions about the patterns of error types (refer to Hancin-Bhatt & Bhatt 1997 for a similar criticism for the MSD model).

Note also in Table 3.3 that epenthesis errors occur only when the first segment of a cluster is voiced (e.g., /by/ and /br/). What is more interesting is that the most marked cluster, /fr/, results in no epenthesis error. Instead, the main type of error in this cluster is a feature change (/f/ → [p]). We can see in Broselow & Finer's study that error types are crucially dependent on the cluster type. Vowel epenthesis occurs when the first segment of a cluster is voiced. On the other hand, errors in manner (i.e., /f/ → [p]) occur when the learners' native language (i.e., Korean) does not have a corresponding segment in its sound inventory.

To summarize, closer examination of the actual error types in Korean speakers' production of English onset clusters raises questions about Broselow & Finer's assertion that the predictions based on MDS were supported. Given that manner feature change in fricative + liquid clusters contributed most to the highest error rate in these clusters (thus supporting the predictions of the MSD parameter) and yet manner feature change is not an error specific to consonant cluster acquisition, Broselow & Finer's assertion is questionable.

In this dissertation, I suggest that manner feature changes (/f/ → [p]) should not be counted as an error resulting from consonant cluster markedness. If Broselow & Finer (1991) intended to code feature changes that result from L1 and L2 differences, they also should have coded Korean speakers' production of English /r/ to a flap or /l/ to [r]. There is no justification for doing this while discussing consonant cluster markedness.

3.2.2 Typological Markedness Hypothesis (TMH)

In a direct response to Broselow & Finer (1991), Eckman & Iverson (1993) claim that typological markedness is all that is needed to explain the observed IL patterns in Broselow & Finer's study. Typological markedness is defined by Eckman & Iverson as follows:

Some segments type A is (typologically) marked relative to some other segment type B if the occurrence of A in a language implies the occurrence of B, but the occurrence of B does not necessarily imply the occurrence of A. Thus, fricatives are marked relative to stops (and equivalently, stops are unmarked relative to fricatives) (e.g., Hawaiian). By similar, now familiar reasoning, voiced stops are marked relative to voiceless stops, and voiced fricatives are marked relative to voiceless fricatives (Eckman & Iverson 1993: 240-41).

The typological markedness relationships for obstruents result in the scales shown in Table 3.4.

Table 3.4. Scales based on typological markedness

unmarked	relative to	marked
stops		fricatives
voiceless stops		voiced stops
voiceless fricatives		voiced fricatives

However, typological markedness in itself does not tell us anything about certain sequencing constraints holding within syllables. Hence, Eckman & Iverson make use of Clements' (1990) *Sequential Markedness Principle* (SMP hereafter) to make predictions on consonant cluster markedness.

(3) Sequential markedness principle (Clements 1990: 313)

For any two segments A and B and any given context X_Y, if A is less marked than B, then XAY is less marked than XBY

According to the sequential markedness principle, *pt* has preference over *pk* by virtue of the fact that *t* is simpler than *k*, and *sk* is simpler than *fk* since *s* is the least marked fricative (Clements 1990: 313).

The predictions made by MSD and TMH are very similar as will be shown in Table 3.5. However, there are two areas where TMH makes different empirical claims about the relative difficulty of onset clusters.

First, MSD predicts that /fr/ is harder than /br/ (due to the smaller sonority distance in /fr/ than in /br/). In contrast, Eckman & Iverson note that no marked relations can be established between these two clusters in terms of typological markedness due to voicing differences in stops and fricatives. In other words, the markedness relationship of /br/-/pr/ or /fr/-/pr/ *can* be established based on the markedness of a voiced stop in /br/-/pr/ and the markedness of a fricative in /fr/-/pr/. Yet, the markedness relation between /fr/-/br/ *cannot* be established because /f/ and /b/ involve two different dimensions in terms of voicing and manner features: /f/ = [-voice][+continuant], /b/ = [+voice][-continuant].

Table 3.5 presents the consonant cluster markedness hierarchy proposed by Eckman & Iverson, which is compared with that of Broselow & Finer. Note in Table 3.5 that Broselow & Finer and Eckman & Iverson make different predictions on the markedness of obstruent + glide vs. obstruent + liquid.

Table 3.5. Markedness hierarchy in Broselow & Finer and Eckman & Iverson

	Less marked < more marked
Broselow & Finer (1991: 41)	a) Cj < Cr (pj < pr) b) pC < bC < fC (pr < br < fr)
Eckman & Iverson (1993: 242)	a) Cr < Cj (pr < pj) b) pr < br c) pr < fr

Broselow & Finer propose that obstruent + liquid is more marked than obstruent + glide based on the MSD value difference in these clusters (stop + liquid = 3, stop + glide = 4). On the other hand, Eckman & Iverson make the opposite claim: stop + glide is more marked than stop + liquid. Eckman & Iverson's claim on the relative markedness of stop + glide with respect to stop + liquid is based on Clements' *Sonority Dispersion Principle* (Clements 1990). According to the sonority dispersion principle (SDP hereafter), the markedness of syllable clusters is characterized in terms of how steep the rise in sonority is from the onset to the peak of the syllable and from the peak of the syllable to the coda.

If we now restate the principle of the sonority cycle in terms of demisyllables,¹³ and consider only unmarked demisyllables, we will say that the initial demisyllable maximizes the contrast in sonority among its members, while the final demisyllable minimizes it. (Clements 1990: 303).

The notion of the sonority dispersion principle can be schematized as follows.

¹³ "A syllable is divided into two overlapping parts in which the syllable peak belongs to both; each of these parts is termed a demisyllable" (e.g., [kran] is composed of demisyllable [kran, an]) (Clements 1990: 303).

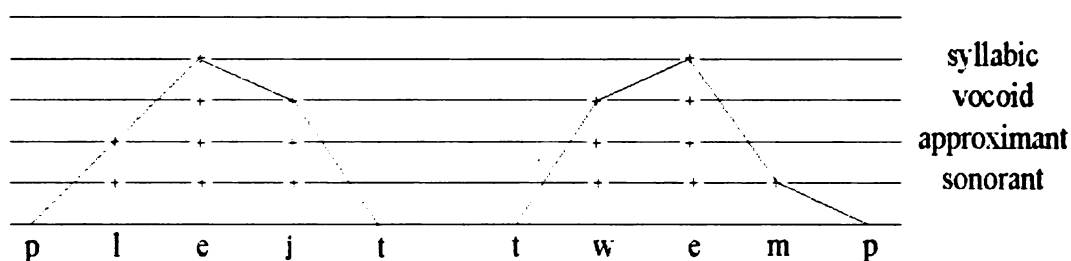


Figure 3.1. Sonority dispersion principle

According to SDP, the initial demisyllable in [plejt] is less marked than that in [twemp] because /ple/ makes a more “sharp and steady rise in sonority” (Clements 1990: 304) than /twe/.¹⁴

Based on the notion of SDP, Clements (1990: 304) propose the following complexity rankings for onset and coda clusters. The complexity ranking in Table 3.6 is calculated based on the sonority scale of obstruent (less sonorous) < nasal < liquid < glide < vowel.

Table 3.6 Complexity ranking (Clements 1990: 304)

(O = obstruent, N = nasal, L = liquid, G = glide, V = vowel)

	less complex < more complex
Initial demisyllable	OLV < ONV, OGV < NLV, NGV < LGV
Final demisyllable	VGL < VLN, VGN < VGO, VNO < VLO

To test the validity of the predictions made by TMH, Eckman & Iverson conducted a production experiment which involved Japanese, Korean, and Cantonese speakers. Table 3.7 presents a summary of this experiment.

¹⁴ I will not present formulas for calculating dispersion in sonority because it is not relevant to the topic of this dissertation. Refer to Clements (1990: 302-311) for the formulas.

Table 3.7 Eckman & Iverson (1993): summary of the experiment

Eckman & Iverson (1993)	
Analytical tool	Typological markedness, SMP, and SDP
Participants	Japanese (n = 4), Korean (n = 4), Cantonese (n = 3)
L2 proficiency	intermediate or high-intermediate
Stimuli	conversational data
	less marked < more marked
Predictions	a) Cr < Cj (pr < pj) b) pr < br c) pr < fr
Results	Predictions were generally supported.

Eckman & Iverson gathered data during short conversations with the participants.¹⁵ L2 learners' production data in Eckman & Iverson (1993) confirmed the consonant cluster markedness hierarchy in Table 3.7. Based on this, Eckman & Iverson (1993) contend that "the correct characterization of this relationship lies in typological markedness rather than in the MSD parameter"(p. 251).

However, it is hard to see much difference between predictions made by TMH and MSD. This is due to the fact that Eckman & Iverson (1993) refer to the notion of sonority whenever it is necessary. We have seen that Eckman & Iverson (1993) adopted the sonority dispersion principle to define the relative markedness between obstruent + liquid and obstruent + glide. In other words, Eckman & Iverson's assertion that typological markedness makes better predictions for L2 data seems to lose ground to some degree given that one of the important findings in their study (relative markedness between obstruent + liquid and obstruent + glide) is not predicted by the typological

¹⁵ Since there was no control over the participants' choice of words, some onsets were not produced at all by some subjects (e.g., /g/ + liquid onsets for one participant), while other onset clusters were produced in abundance (e.g., one participant produced /p/ + liquid onset clusters 93 times). Thus, instead of comparing the production of onset clusters in terms of the number of successes or errors, Eckman & Iverson decided whether or not a participant's interlanguage can be said to have a given cluster. The threshold criterion adopted for this purpose is 80% success on a minimum of four tries.

markedness alone. Instead, they have to refer to the sonority dispersion principle.

I have discussed in this section two previous models of L2 consonant cluster acquisition. Broselow & Finer (1991) predict that L2 learners' cluster acquisition is constrained by the MSD parameter plus L1 sequencing possibilities. On the other hand, Eckman & Iverson (1993) argue that typological markedness relations along with Clements' (1990) SMP and SDP suffice to make correct predictions on L2 consonant cluster acquisition.

I have chosen Broselow & Finer (1991) and Eckman & Iverson (1993) for this review because they provide formal models of L2 consonant cluster acquisition, which are testable with other sets of L2 data. In the following section, I present previous studies that tested these two models and discuss the results from these studies.

3.3. Empirical Data on L2 Acquisition of Consonant Clusters

This section provides published L2 consonant cluster acquisition data to demonstrate how much of the data can be covered by the Minimal Sonority Distance parameter (MSD) and the Typological Markedness Hypothesis (TMH). Even though Broselow & Finer (1991) and Eckman & Iverson (1993) make predictions on a small set of English onset clusters, and they do not claim that their model could be extended to coda, several studies have tested the validity of MSD and TMH on L2 acquisition of coda clusters (e.g., Hancin-Bhatt & Bhatt 1997). Thus, I will first provide the consonant cluster markedness hierarchy generated by MSD and TMH in § 3.3.1. In this way, it will be easier to see whether the published data support MSD and/or TMH.

§ 3.3.2 provides empirical data which *support* the predictions of MSD and/or TMH, whereas § 3.3.3 provides empirical data that *contradict* the predictions of MSD and/or TMH. By the end of this chapter, it will be made clear that neither MSD nor TMH is a good predictor of second language learners' performance on non-native consonant clusters.

3.3.1. Consonant Cluster Markedness in MSD and TMH

Table 3.8 presents the consonant cluster markedness hierarchy generated by the MSD and TMH. Note that obstruent + glide clusters are not included in the table due to the controversy over the status of a glide in languages like Korean and Mandarin Chinese.

Table 3.8. Consonant cluster markedness hierarchy in MSD and TMH
(less marked < more marked; // no markedness relationship)

Minimal sonority distance parameter	
Onset	pl, pr, tr, kl, kr (3) < bl, br, dr, gl, gr (3) < fl, fr, sl (2) < sm, sn (1) < sp, st, sk (-1)
Coda	a) sonorant + C: lp, lt, lk (3) < lb, ld (3) < ls, lz (2) // np, nt, mp (2) < nd (2) < lm , ns (1) < nz (1)
	b) obstruent + C: sp, st, sk (1) < zd (1) < pt, kt (0) < bd, gd (0) < ps, ts, ks (-1) < bz, gz (-1)
Typological markedness hypothesis	
Onset	a) pl, pr, tr, kl, kr < bl, br, dr, gl, gr b) pl, pr, tr, kl, kr < fl, fr, sl
Coda	a) ft, sp, st, sk < pt, kt (Eckman 1991) b) lm < np, nt, mp, ns, nd, nz < lp, lt, lk, lb, ld (SDP only)

Note: The number in the parenthesis indicates the sonority distance in a cluster, and the bold type indicates the cluster where MSD and SDP make a completely opposite prediction.

Table 3.8 illustrates that when based on MSD, voiceless stop + liquid clusters are least marked among onset clusters, while /s/ + stop clusters are most marked (e.g., Hansen

2001; Osburne 1996; Tropic 1987).¹⁶ Among sonorant + C coda clusters, liquid + voiceless stop is least marked, while nasal + voiced strident is most marked.

Table 3.8 also presents predictions of TMH on English consonant cluster markedness. TMH predicts that fricative + stop is less marked than stop + stop in obstruent + obstruent coda clusters. Regarding sonorant + obstruent coda clusters, only the sonority dispersion principle (Clements 1990) is used to generate the consonant cluster markedness hierarchy. This is not only because Eckman & Iverson (1993) did not provide typological markedness for sonorant + obstruent clusters, but also because some studies (e.g., Hansen 2001) tested coda cluster markedness based on the SDP alone.¹⁷ According to the SDP, liquid + nasal clusters are least marked among coda clusters, while liquid + stop clusters are most marked.

The major difference between MSD and TMH occurs in the markedness relationship in coda clusters. Recall that SDP predicts that the steadier the fall of a sonority contour, the better it is as a coda cluster. In contrast, MSD predicts that the bigger the sonority distance between two segments, the better they are as consonant clusters. Accordingly, SDP makes the opposite predictions, in comparison with MSD, on the relative markedness of liquid + stop and liquid + nasal. Note in Table 3.8 that liquid + stop is least marked based on MSD, while it is most marked based on SDP. Similarly, liquid + nasal is least marked based on SDP, while it is one of the most marked coda clusters according to MSD due to the small sonority distance between adjacent segments.

¹⁶ Hansen (2001: 356) notes that “while all three-member codas may be difficult for native speakers of Mandarin, the codas that violate the UCSS appear to be especially difficult.” In Hansen’s study, coda clusters meet UCSS (Universal Canonical Syllable Structure) if the sonority of segments increases as the syllable nucleus is approached.

¹⁷ SDP makes no distinctions between stops vs. fricatives and voiced obstruents vs. voiceless obstruents. Refer to Clements’ proposal in Table 3.6

3.3.2. Empirical Data *Supporting* MSD or TMH

This section presents L2 consonant cluster acquisition data supporting MDS and/or TMH.

Broselow & Finer (1991) and Eckman & Iverson's (1993) study are not included in the table below because the results of these studies were already discussed.

Table 3.9 Empirical data *supporting* MSD or TMH
(vl = voiceless, vd = voiced, O = obstruent, S = sonorant)

Study	Participants	Results
Eckman (1991)	L1: Japanese (n = 4) Korean (n = 4) Cantonese (n = 3) L2: English	implicational relation (A → B: the existence of A implies the existence of B) a) coda: pt, kt → ft, sp, st, sk
Hancin-Bhatt & Bhatt (1997)	L1: Japanese (n = 10) Spanish (n = 10) L2: English	fewer errors < more errors a) coda: liquid + vl stop < liquid + vl fricative < liquid + nasal
Kim (2000)	L1: Korean (n = 16) L2: English	fewer errors < more errors a) onset: vl O + S (pr) < vd O + S (br) b) coda: S + vl O (nt) < vl O + vl O (pt) < vd O + vd O (gd)

Note: O = obstruent, S = sonorant, vl = voiceless, vd = voiced

Eckman (1991) tested the *Typological Markedness Hypothesis* by examining the implicational relation between fricative + stop vs. stop + stop coda clusters. The hypothesis in Eckman's study was that "the universal generalizations that hold for the primary languages hold also for interlanguages" (Eckman 1991: 24).¹⁸ In particular, Eckman proposes the *Fricative-Stop Principle*, which states that no language will have words ending in a stop + stop sequence without also having words ending in a fricative + stop sequence. By the same logic, a language may have words ending in a fricative + stop

¹⁸ Eckman termed this hypothesis the *Interlanguage Structural Conformity Hypothesis* (ISCH).

sequence without having words ending in a stop-stop sequence.¹⁹

The *Fricative-Stop Principle* generates a testable implicational relationship such that the existence of word-final stop + stop sequences implies the existence of fricative + stop sequences: pt#, kt# → ft#, sp#, st#, sk#. Furthermore, when the *Fricative-Stop Principle* is applied to interlanguage phonology, we can predict that L2 learners can have word-final fricative + stop sequences in their grammars even if they do not have stop + stop sequences. The results of Eckman (1991) confirmed that the *Fricative-Stop Principle* holds for L2 learners.

On the other hand, Hancin-Bhatt & Bhatt (1997) tested the predictions of the *Minimal Sonority Distance Parameter* (MSD), and the predictions were generally confirmed in Japanese and Spanish speakers' production of English coda clusters. For instance, liquid + stop is predicted to be the least marked by MSD, and it turned out that the error rate was lower in this cluster.

However, given the type of coda clusters tested in Hancin-Bhatt & Bhatt, it is hard to say whether their study supports predictions of MSD: they did not include a voiced stop in the stimuli. What if they had included a voiced stop in the stimuli (e.g., lb, ld, lz)? Would it be still true that liquid + voiced stop is less marked than liquid + voiced fricative? This question remains unanswered in Hancin-Bhatt & Bhatt (1997).

Lastly, Kim (2000) tested 16 Korean speakers on their production of English consonant clusters. Among various onset and coda clusters tested in his study, the consonant cluster markedness hierarchy confirmed in his study is presented in Table 3.9. For instance, as was predicted by both MSD and TMH, voiceless obstruent + liquid onset

¹⁹ The *Fricative-Stop Principle* is based on Greenberg's (1978) generalizations about initial and final clusters.

clusters induced fewer errors than voiced obstruent + liquid. However, Kim's data is also not immune from conflicting data. These conflicting data are presented in the following section.

To summarize, it seems rather difficult to find studies which confirm the predictions of MSD or TMH in a comprehensive way. For instance, Eckman (1991) tested only one hypothesis (fricative + stop vs. stop + stop). Hancin-Bhatt & Bhatt (1997) did not include a cluster with a voiced obstruent, but gave no justification for this exclusion. Moreover, it will be shown in the following section that there is a fair amount of data that go against the predictions of MSD and TMH.

3.3.3. Empirical Data *Against* MSD and/or TMH

There are five areas where conflicting data on the predictions of MSD and TMH are frequently observed. These five areas are summarized in (4).

(4)

a) Onset: /s/ + stop vs. stop + liquid

MSD predicts that stop + liquid is less marked than /s/ + stop, but Chen (2003), Kim (2000), and Major (1996) revealed the opposite direction of the markedness relationship.

b) Coda: nasal + voiceless stop vs. nasal + voiceless fricative

MSD predicts that nasal + voiceless stop is less marked than nasal + voiceless fricative, but Hansen (2000) demonstrated the opposite direction of the markedness

relationship.

c) Coda: liquid + nasal vs. liquid + stop

SDP (sonority dispersion principle) predicts that liquid + nasal is less marked than liquid + voiceless stop, but Hansen (2000, 2001) observed the opposite direction of the markedness relationship.

d) Coda: stop + fricative vs. fricative + stop

MSD predicts that fricative + stop is less marked ($MSD = 1$) than stop + fricative ($MSD = -1$), but Hansen (2000, 2001), Major (1996), and Wiltshire (2005) demonstrated the opposite direction.

e) Coda: stop + stop vs. stop + fricative

MSD predicts that stop + stop is less marked ($MSD = 0$) than stop + fricative ($MSD = -1$), but the opposite direction is observed in Hansen (2000) and Major (1996).

Table 3.10 presents a summary of the previous studies showing the predictions of MSD and/or TMH do not hold true with their data. Most of the data in Table 3.10 are self-explanatory; therefore, I will focus on illustrating Davidson (2001; 2003) which tested Polish and Czech clusters, respectively. I will also discuss Osburne (1996) and Major's (1996) accounts for the exceptional behavior of some clusters in their study.

Davidson (2001) tested English speakers on the production of Polish-legal onset clusters with sonority distances of 0 and 1 or higher. Sixteen native speakers of English participated in this experiment. The target words used in the study were pseudo-Polish words.

Table 3.10. Empirical data *against* MSD or TMH

Study	Participants	Results
Osburne (1996)	L1: Vietnamese (n = 1) L2: English	less deletion < more deletion a) coda: ts, ndz < kt, pt
Major (1996)	L1: Brazilian Portuguese (n = 4) L2 = English	error probabilities (lower < higher) a) onset: /s/ + stop (0.37) < stop + liquid (0.44) < fricative /s, ʃ/ + liquid (0.70) b) coda: stop + fricative (0.16) < fricative + fricative (0.27) < liquid + stop (0.38) < fricative + stop (0.75) < stop + stop (0.98)
Kim (2000)	L1: Korean (n = 16) L2: English	less epenthesis < more epenthesis onset: /s/ + stop < voiceless stop + sonorant
Hancin-Bhatt (2000)	L1: Thai L2: English	fewer errors < more errors liquid + fricative < liquid + stop fewer errors < more errors
Hansen (2000)	L1: Vietnamese (n = 2) L2: English Longitudinal Study	a) coda: vl stop + vl fricative < vl stop- vl stop b) coda: vd stop+ vd fricative < vd stop- vd stop c) coda: nasal + vl fricative < nasal + vl stop d) coda: liquid + vl fricative < liquid + vl stop
Hansen (2001)	L1: Mandarin Chinese (n = 3) L2: English	fewer errors < more errors a) coda: nasal + obstruent < liquid + nasal
Davidson (2001)	L1: English L2: Polish (n = 16)	fewer errors < more errors a) onset: /zm, zr/ < /kt, kp, pt, gk, tf/ < /vn, vz, dv/
Davidson (2003)	L1: English (n = 20) L2: Czech	fewer errors < more errors a) onset: /sC/ < /fC/ < /zC/ < /vC/
Chen (2003)	L1: Chinese speakers in Taiwan L2: English	fewer errors < more errors a) onset: /s/ + stop < voiceless stop + liquid

The onset clusters in Davidson's study were divided into three different categories according to the sonority distance between them: a) clusters with a sonority distance of 0 (pt, kp, vz), b) clusters with a sonority distance of at least 1 that *do not exist* in English (zm, tf, vn), and c) clusters with a sonority distance of at least one that *do exist* in English

(fr, sm, sn). It is not surprising that participants in Davidson's study demonstrated the best performance with the clusters /fr, sm, sn/ because they do *exist* in English. When it comes to the clusters that do not exist in English, the following error rate was observed.

(5) Error rate onset consonant clusters (Davidson 2001)

Easy: /zm, zr/ < Intermediate: /**kt**, **kp**, **pt**, **gk**, tf/ < Difficult: /vn, **vz**, dv/

Note in (5) that clusters appearing in boldface have a sonority distance of 0, which means that these clusters are most marked based on her groupings. We can see that only one cluster with the sonority distance of 0 is classified as a difficult cluster, while most of the clusters with the sonority distance of 0 belong to the intermediate group. The ordering in (5) highlights the fact that “defining and grouping clusters on the basis of sonority is neither an accurate nor adequate predictor of performance on illegal clusters” (Davidson 2001: 32).

Davidson's (2003) study on English speakers' production of Czech clusters demonstrates similar patterns to her 2001 study. English speakers showed more difficulty with v + obstruent clusters than z + obstruent clusters, which cannot be accounted for in terms of MSD (v + obstruent and z + obstruent clusters have the same MSD value).

Davidson's (2003) study will be discussed in more detail in Chapter 4.

On the study of Vietnamese learners' production of English coda clusters, Osburne (1996) found that coda clusters that violate the Sonority Sequencing Principle (SSP) (ts, ndz) are significantly less likely to be reduced than those which follow it (kt, pt). In addition, it was found that a stop is the main target of deletion, not a strident. Regarding this unexpected pattern, Osburne (1996: 173) notes that many of the clusters which violate the sonority sequencing principle are present because of the addition of an

inflectional ending (e.g., tense or plural marking). Based on this, Osburne (1996) surmises that the clusters with inflectional marking may be more salient to the learner and hence are less likely to be reduced.

In Chapter 4, I will provide an alternative account for the consonant deletion pattern in coda clusters in terms of the perceptibility of a stop and a strident in an obstruent + obstruent sequence. When based on the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM), the observation that there is more deletion in stop + stop coda clusters than in stop + strident clusters is the natural outcome of the *Principle of Perceptual Saliency*, one of the main working principles in PCCAM. The principle of perceptual saliency predicts that a stop, as opposed to a strident, is the main target of deletion in obstruent + obstruent sequences. This is because the perceptual cues of a stop are relatively weak when the stop is not flanked by a vowel. The principle of perceptual saliency is further elaborated in § 4.2.4.

Lastly, Major (1996) tested four adult native speakers of Brazilian Portuguese on their production of English consonant clusters.²⁰ The results show that in onset clusters, stop + liquid clusters contribute more strongly to error than /s/ + stop clusters, which contradicts the predictions of MSD. The predictions of MSD did not hold true in obstruent + obstruent coda clusters either. Recall that the predictions of MSD in these clusters are fricative + stop (1) (less marked) < stop + stop (0) < stop + fricative (-1) [the MSD value is in the parenthesis]. Yet, the actual results in Major demonstrate an opposite direction: stop + fricative (fewer errors) < fricative + stop < stop + stop. A similar pattern is found in Hansen (2000).

²⁰ The clusters tested in the study were /sl, ʃr, sp, st, sk, pr, br, pl, bl, tr, dr, kr, gr, kl, gl/ (onset) and /rp, rb, lp, lb, rt, rd, lt, ld, rk, rg, lk, pt, bd, kt, gd, ps, bz, ts, dz, ks, gz, sp, st, zd, sk, fs, vz/ (coda). VARBRUL analyses were performed to determine probabilities of errors for each cluster.

Regarding the low error probability in /s/ + stop onset clusters and stop + /s/ coda clusters, which is contradictory to the prediction of MSD, Major (1996: 92) finds accounts for the special status of /s/ in the organization of a syllable structure.

From the standpoint of sonority and markedness, it would at first appear that the data do not conform to these implicational hierarchies, since #FS has the lowest error probability yet is more marked than #SL and #FL . . . All #FS clusters in English begin with /s/. Universally, obstruent plus /s/ sequences have the peculiar characteristic of functioning at some level as a single consonant both initially and finally (Selkirk 1984) . . . Therefore, in the present study the lower error probability of #FS may occur because this cluster is acting as a single consonant (Selkirk 1984) or the /s/ is not truly part of the syllable onset (Kaye 1989). Note: F = fricative, S = stop, L = liquid.

The issue of representing /s/ in sC clusters in terms of a syllable template has long been debated, mainly due to the fact that /s/ in sC clusters does not fit smoothly into canonical syllable templates.

For instance, /s/ + stop sequences violate the Sonority Sequencing Principle (Clements 1992; Hooper 1976; Selkirk 1984; Steriade 1982, among others). Recall that the SSP states that in a syllable, the most sonorous segment constitutes the syllable peak, with all other segments organized around it in such a way that relatively more sonorous segments are closer to the peak and relatively less sonorous ones are farthest away from it. Contrary to the proposed organization of a syllable under the SSP, /s/ + stop in onset position demonstrates a sonority reversal: it begins with a relatively more sonorous fricative segment and ends with a relatively less sonorous stop segment.

In addition, sC clusters violate several word-initial English phonotactic constraints:

(1) /st/ and /sl/ clusters violate a phonotactic constraint of homorganic clusters which

bans clusters such as */pw/ */bw/, */fw/, */tɫ/ and */dl/, (2) /s/ + stop clusters also violate a constraint of obstruent + obstruent sequences such as */ft/, (3) A phonotactic constraint of obstruent + nasal sequences which bans */pn/ is violated in /sm/ and /sn/ clusters (Barlow 2001).

Various solutions have been proposed to deal with these exceptional properties of sC clusters. One of the dominating approaches to the properties of sC clusters is to assume that in languages that have sC clusters, /s/ is extrasyllabic (see Davis 1990; Harris 1994; Treiman, Gross, and Cwikel-Glavin 1992). Another view is that /s/ + stop clusters are acting as a single consonant (Selkirk 1984).

Following many other researchers who view /s/ in sC clusters as special, Major conjectures that Brazilian Portuguese learners of English make the lowest errors in /s/ + stop onset clusters due to the special status of /s/ in syllable margin. Major also suspects that the low error probability of /s/ + stop clusters results from positive transfer where vowel devoicing between /s/ and a voiceless stop gives a phonetic sequence of /s/ + stop. In Portuguese, although sC clusters are not allowed, #isC/ is possible. In addition, in casual speech, /i/ is often devoiced and then deleted, giving a phonetic #[sC] sequence: [ˈiskɔla] → [ˈskɔla], 'school' (Major 1996: 88) .

I will not argue against Major's accounts for the low error probability in a /s/ + stop cluster. It may be true that /s/ + stop clusters are special in terms of syllable organization and that vowel devoicing between /s/ and a voiceless stop may have contributed to the low error rates. Yet, it seems that Major's accounts are missing an important question. Let's assume that /s/ is extrasyllabic and the extrasyllabicity of /s/ makes it easier for L2 learners to produce this cluster. But still we do not know what it is

in the extrasyllabicity that makes it easier for L2 learners to produce extrasyllabic segment + CV sequence (e.g., /s/ + stop) than true clusters (e.g., /s/ + liquid). Other questions remain, such as will every extrasyllabic segment + CV sequence be easier for L2 learners? What about a language like Czech that allows a /r/ + obstruent onset cluster, and /r/ is extrasyllabic in the language? (e.g., *rty* "lips", *lhát* "to lie"; the examples are from Scheer 2002). Is this sequence still easy for L2 learners who learn this language due to the extrasyllabicity of /r/?

In § 4. 4, I will show that low error probabilities in /s/ + stop onset clusters can be accounted for without referring to the notion of extrasyllabicity of /s/. More specifically, I will show that phonetic factors associated with /s/ + stop onset clusters (e.g., the articulatory stridency effect, the perceptual stridency effect, and maximal perceptual contrast) efficiently account for the lowest error rate in this sequence.

3.4. Summary of Chapter 3

I began § 3.1 by introducing two major concepts used in the previous models of L2 consonant cluster acquisition, the sonority and the sonority sequencing principle. The main point of this section was that the notions of sonority and the sonority scale are still controversial, even though the sonority sequencing principle has certain explanatory power in accounting for co-occurrence restrictions in syllable margins.

In § 3.2, I reviewed two previous models of L2 consonant cluster acquisition, the Minimal Sonority Distance Parameter (Broselow & Finer 1991) and the Typological Markedness Hypothesis (Eckman & Iverson 1993). I pointed out that these models were not fully tested due to the small number of consonant clusters used in these studies.

Furthermore, I showed in § 3.3. 3 that a large amount of the published data on L2 learners' production of consonant clusters actually supports neither MSD nor TMH. These conflicting data are from L2 learners from a variety of native language backgrounds (e.g., Brazilian Portuguese, English, Korean, Mandarin Chinese, Thai, and Vietnamese) and their production of several target languages (English, Polish, and Czech). Given the findings from this literature review, it can be safely concluded that MSD and TMH are not an adequate predictor of L2 learners' production of non-native consonant clusters.

As an alternative, I will develop a *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM) in Chapter 4. It will be shown that this new model will not only provide a better account for the published data reviewed in this chapter but also explain new data from Korean speakers' production of English consonant clusters, which will be presented in Chapter 6.

CHAPTER 4

PHONETICALLY BASED CONSONANT CLUSTER ACQUISITION MODEL

The main purpose of Chapter 4 is to construct the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM) and provide theoretical and empirical justifications for the consonant cluster markedness hierarchy generated by PCCAM.

The structure of this chapter is as follows. § 4.1 and § 4.2 review theoretical frameworks upon which the phonetic factors in PCCAM are grounded. More specifically, § 4.1 introduces Articulatory Phonology (Browman & Goldstein 1986 et seq.) and demonstrates how the articulatory factor, the degree of gestural overlap, is justified under Articulatory Phonology. The articulatory release effect, the articulatory stridency effect, and the articulatory voicing effect, which are not directly derived from Articulatory Phonology, but are assumed to play an important role in defining consonant cluster markedness, are also discussed in this section.

The following section § 4.2 reviews the basic principles of the P-map (Steriade 2001a, 2001b). According to the P-map hypothesis, a sound change whose modification is less perceptible is more likely to occur than a sound change whose modification is more perceptible. Based on this hypothesis, the perceptual stridency effect, the perceptual voicing effect, and the principle of perceptual saliency are proposed to account for the patterns of vowel insertion and consonant deletion in L2 consonant cluster acquisition. The maximal perceptual contrast parameter is also defined in this section.

With the articulatory and perceptual factors defining consonant cluster markedness in hand, § 4.3 presents the consonant cluster markedness hierarchy generated

by PCCAM. More specifically, this section introduces the mechanism through which PCCAM establishes the relative consonant cluster markedness. The relative consonant cluster markedness in PCCAM is calculated by a three-way computation of consonant cluster markedness: a) markedness of C_1 (the first consonant in a cluster), b) markedness of C_2 (the second consonant in a cluster), and c) markedness of gestural coordination of C_1 and C_2 . Note that central to the consonant cluster markedness hierarchy in PCCAM is the assumption that consonant cluster markedness determines a) language learners' variable degree of production difficulty with non-native consonant sequences and b) learners' repair strategies dealing with certain types of consonant sequences.

The consonant markedness hierarchy generated by PCCAM is justified throughout § 4.4 and § 4.5. These two sections also present relevant data from first and second language acquisition which support the proposed markedness ranking relations among clusters. § 4.6 concludes this chapter by summarizing all the articulatory and perceptual factors defining consonant cluster markedness in PCCAM.

4.1. Articulatory Factors in PCCAM

This section discusses articulatory factors defining consonant cluster markedness in PCCAM. A brief introduction of articulatory factors in PCCAM is given in (1).

(1) Articulatory factors

- a. Degree of gestural overlap: a consonant sequence with less gestural overlap is easier for language learners to acquire than a consonant sequence with more gestural overlap.
- b. Articulatory release effect: a consonant sequence is easier for language learners to acquire when a stop is followed by [+continuant] than when a stop is followed by [-continuant].
- c. Articulatory stridency effect: vowel insertion is less likely to occur after a voiceless strident because the aerodynamic conditions of /s/ remain conducive to the continued turbulence of the strident rather than vocal fold vibration.
- d. Articulatory voicing effect: vowel insertion is more likely to occur after a voiced obstruent due to the articulatory difficulty of maintaining voicing in an obstruent

This dissertation adopts the view of Articulatory Phonology (AP) on phonological acquisition that language learners are supposed to elaborate individual gestures by *differentiating, tuning, and coordinating* those gestures. Moreover, the degree of gestural overlap parameter in PCCAM directly refers to the principle of AP that utterances are modeled as organized patterns of gestures, in which gestural units may overlap in time. Thus, it is necessary to review the basic principles of AP to give a general background on the parameter of the degree of gestural overlap.

To this end, § 4.1.1.1 reviews how sounds and phonological patterns are represented in AP. § 4.1.1.2 illustrates the representation of consonant cluster in AP and discusses how an epenthetic vowel which occurs in consonant clusters is understood in AP. The remainder of § 4.1 is devoted to justifying the four articulatory factors, the

degree of gestural overlap (§ 4.1.2), the articulatory release effect (§ 4.1.3), the articulatory stridency effect (§ 4.1.4), and the articulatory voicing effect (§ 4.1.5).

4.1.1. Theoretical Framework: Articulatory Phonology (AP)

4.1.1.1. Representation of Sounds in Articulatory Phonology

Articulatory Phonology (Browman & Goldstein 1986 et seq.) emerged as an attempt to relate speech activity to more general issues in motor behavior. Drawing upon principles and techniques from general motor behavior such as task dynamics, Articulatory Phonology (AP hereafter) views phonological structures as an organization of the physical actions involved in speaking.

Broadly speaking, AP is a model that takes dynamically-defined articulatory gestures as the basic units of phonological representation. AP is based on the principle that utterances are modeled as organized patterns of gestures, in which gestural units may overlap in time. Gestures are abstract characterizations of articulatory events and include three key dimensions: spatial, temporal, and dynamic dimensions.

Spatial dimensions of gestures derive from the fact that gestures are essentially an instruction for an articulator to achieve a particular constriction in the vocal tract (e.g., an opening of the velum, a closing of the glottis, or a raising of the tongue body). A set of parameters specifies the spatial goals of that constriction: constriction degrees (CD) and constriction locations (CL). These parameters are labeled as vocal tract variables.

Examples of vocal tract variables and articulator sets are listed in Table 4.1.

Notice in Table 4.1 that five articulator sets are specified for gestures: the tongue body, the tongue tip, the lips, the glottis, and the velum. Each articulatory set is further specified with vocal tract variables. For instance, a gesture involving the tongue body (TB) is parametrized by the values of two vocal tract variables: tongue body constriction location (TBCL) and tongue body constriction degree (TBCD).

Table 4.1. Vocal tract variables (Browman & Goldstein 1989: 207, 223)

Articulator sets	Vocal tract variables	
lips	lip aperture	(LA)
tongue tip	tongue tip constriction location	(TTCL)
	tongue tip constriction degree	(TTCD)
	tongue body constriction location	(TBCL)
	tongue body constriction degree	(TBCD)
velum	velic aperture	(VEL)
glottis	glottal aperture	(GLO)
Constriction degree (CD) values (Gafos 2002: 275)		
[closed]	stops	
[critical]	fricatives	
[narrow]	some vowels	
[mid]	approximants and mid vowels	
[wide]	some vowels	
Constriction location (CL) values		
[labial], [dental], [alveolar], [post-alveolar], [palatal], [velar], [uvular], [pharyngeal]		

Table 4.1 also illustrates that the constriction degree (CD) takes on a range of values with five categorical distinctions: [closed], [critical], [narrow], [mid], and [wide]. On the other hand, constriction location (CL) takes the values of [labial], [dental], [alveolar], [postalveolar], [palatal], [velar], [uvular], and [pharyngeal]. Note that constriction degree [CD] corresponds to the manner dimension under feature theory, while constriction location (CL) serves to make the distinctions in terms of place of articulation.

Consider the case of a voiceless velar stop and a voiced velar stop (/k/ vs. /g/).

Both /k/ and /g/ take the same values when it comes to TBCL {velar} and TBCD {closed}. However, these two segments differ in terms of voicing, and voicing is

represented by the values in GLO (glottal aperture). /g/ has GLO {critical}, which represents the status of the glottis while voicing, while /k/ has the GLO {wide}.²¹ Thus the set of gestures for /k/ is TBCL {velar}, TBCD {closed}, GLO {wide}, while the set for /g/ is TBCL {velar}, TBCD {closed}, and GLO {critical}.

To summarize this section, AP represents the contrasts of sounds by the presence or absence of a gesture as well as the degree and location of the constriction. It should be noted that the timing and organization of gestures are also used to represent contrasts of sounds in AP. These aspects of AP (the timing and organization of gestures) are discussed in the following section.

4.1.1.2. Gestural Coordination

Recall that gestures are abstract characterizations of articulatory events, characterized by three dimensions: a) spatial, b) temporal, and c) dynamic dimensions. A spatial dimension of gestures was discussed in the previous section. This section discusses the other two dimensions of gestures, i.e., the temporal and dynamic dimensions.

A temporal dimension of gestures is basically related to the way a gesture unfolds in time. Consider the case of velar stop /k/ with CD = {closed}, CL = {velar}. The CD tract variable changes its value over time from its initial value, which is determined based on the initial tongue body position at the onset of /k/ (i.e., the point in time that the articulator begins to move towards its specific target), to the extent that it eventually attains its target value, {closed}. It is this continuous change in the value of CD that is

²¹ The representation of voicing contrast under AP is based on Gafos (2002: 294) because Browman & Goldstein (1986, et seq.) do not show any glottal gesture for voiced segments.

modeled as a dynamical system.

It is important to note that gestures enter into the temporal relations of overlap with other gestures in speech. That is, the gestures for a given utterance are organized into a large coordinated structure, or *Constellation*, and the pattern of coordinated gestures is represented in a *Gestural Score*. A gestural score has a two-dimensional representation with articulatory tiers on one dimension and temporal information on the other. A gestural score for [spæ̃n] is provided in Figure 4.1.

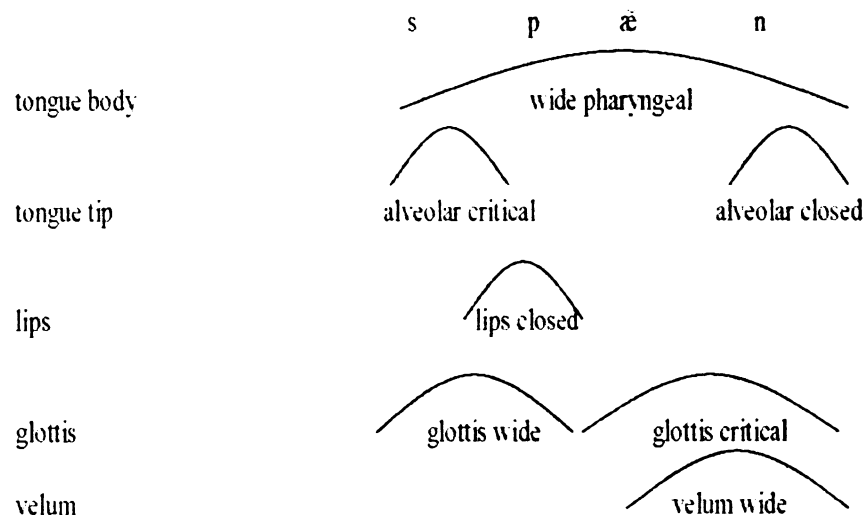


Figure 4.1. Gestural score of English *span* [spæ̃n]

(Browman & Goldstein 1992: 158)

Each curve in Figure 4.1 corresponds to individual gestures. The CD and CL parameters are specified under each gesture. In [spæ̃n], the initial oral gesture is an alveolar closure, and the gesture is parameterized by CL {alveolar} and CD {critical}.

Each curve in Figure 4.1 also represents the time-varying tract variable

trajectories generated by the task dynamic model. Thus, the horizontal extension of each curve represents the interval of time during which that particular gesture is active.

Observe that tongue tip and lip gestures, which represent /s/ and /p/, respectively, overlap each other in time, which demonstrates how consonant cluster articulation is represented in AP. In other words, the gestural overlap of /s/ and /p/ in [spæn] illustrates that the temporal organization is expressed through *coordination relations* between gestures. A coordination relation specifies that some landmark within the temporal structure of one gesture is synchronous with some landmark within the temporal structure of another gesture.

We have seen that gestures are characterizations of the movements of articulators through space and time, and the gestural score shows that gestures are co-articulated with other gestures and that they are co-articulated for a certain period of time. It seems that the notion of gestural overlap is particularly useful in accounting for coarticulation phenomena such as assimilation, reduction, and epenthesis. Consider the case of final /t/ or /d/ deletion in English, a variable process which occurs particularly before initial obstruents (Guy 1980). Browman & Goldstein (1989) view the deletion of /t/ or /d/ as a result of a gestural overlap rather than a deletion of a segment per se. Thus, the production of *perfect memory* /pəˈfekt məməri/ as [pəˈfek məməri] results from the complete overlap of an alveolar gesture with the constrictions of the preceding velar closure and the following labial closure. As additional supporting evidence for this claim, Browman & Goldstein (1990b) provide articulatory data which illustrate that an alveolar gesture is actually existent, even though we cannot hear it. It is observed in the production of *perfect memory* /pəˈfekt məməri/ as [pəˈfek məməri] that the alveolar

closure gesture at the end of *perfect* was produced with much the same magnitude as when the two words were produced in isolation.

To summarize, the deletion of /t/ in [pəˈfɛk mɛməri] can be understood under AP as a result of increased overlap between preceding velar and following labial gestures, which in turn brings about the acoustic effect of hiding the alveolar closure gesture. Other phonological patterns for which the notion of gestural overlap seems to provide particularly insightful accounts are the cases of assimilation and epenthesis. The following section discusses the case of vowel insertion in consonant clusters in detail.

4.1.1.3. Gestural Coordination and an Epenthetic Vowel

AP assumes that the distinction between *beret* and *bray* only lies in the degree of gestural overlap among the initial consonant gestures (bilabial closure and rhotic). That is, in *bray*, the bilabial closure and rhotic gestures are tightly organized with respect to one another, and to the vowel. In contrast, in *beret*, the bilabial closure is set off from the rhotic and vowel gestures, and therefore, no gestural overlap among the consonants occurs (Browman & Goldstein 1992). Thus, the nucleus of a reduced syllable (e.g., schwa) is assumed not to have any explicit vowel gesture associated with it under AP.

Gafos (2002) further develops the issues of gestural coordination and its resulting vocalic segment. In Moroccan Colloquial Arabic, an excrescent schwa (i.e., epenthetic vowel) occurs between consonants in coda clusters. What makes this case even more interesting is that an excrescent schwa does not occur in onset clusters. Gafos (2002) proposes that excrescent schwa occurs only in a coda cluster in Moroccan Colloquial

Arabic basically due to the different coordination relation between onset and coda clusters. The proposal is that in Moroccan Colloquial Arabic, onset clusters have a close transition between two consonants, while coda clusters have an open transition between consonants. The general coordination topology for CCVCC sequences can be described as in Figure 4.2.

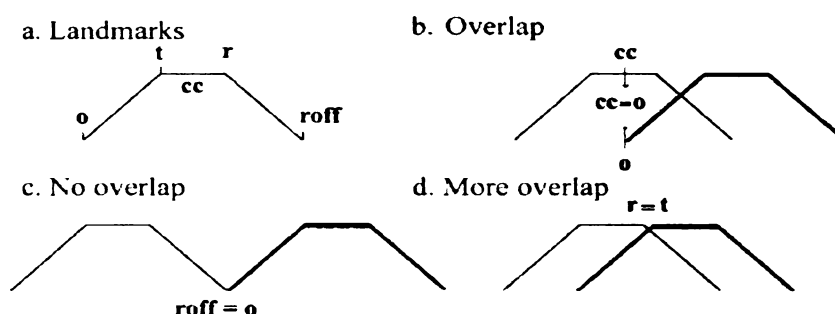


Figure 4.2. Examples of temporal relations

(Notation: 'o' onset, 't' target, 'cc' c-center, 'r' release, 'roff' r offset)

Let us first define the terms used in Figure 4.2. Gafos (2002: 271) used a set of *landmarks*, onset (o), target (t), c-center (cc), release (r), and release offset (roff) to capture the dynamical states of a gesture. The definition of each landmark is presented in (2).

(2) The definition of landmarks (Gafos 2002: 271)

onset: the onset of movement toward the target of the gesture

target: the point in time at which the gesture achieves its target

c-center, the mid-point of the gestural plateau

release: the onset of movement away from the target

release offset: the point in time at which active control of the gesture ends

Figure 4.2 depicts the gestural landmarks (a) and some examples of coordination relations in (b, c, d). We can see that (b), (c), and (d) in Figure 4.2 represent different degree of overlap between gestures. For instance, (b) in Figure 4.2 demonstrates the case where a sequence of two consonants is produced with an intervening acoustic release. Taking an example from Moroccan Colloquial Arabic where the active participle of the verb ‘to write’ [kat³b] is produced with a schwa-like vocalic transition in the coda cluster, Gafos (2002: 271) describes the relation in Figure 4.2 (b) as follows.²²

The *onset* of movement for the lip gesture for /b/ is initiated around the mid-point of the tip-blade gesture for /t/, the *c-center* of /t/ – indicated as ‘cc = o’ As a consequence of this relation, the achievement of the target for the /b/ gesture, lip closure, takes place after the *release* of the /t/ gesture. There is, thus, a period of no constriction in the transition between /t, b/ that is identified as a schwa-like vocalic element.

The coordination relation in Figure 4.2 (d), on the other hand, shows a pattern where the articulatory release of the first gesture coincides with the target of the second gesture, ‘r = t’. Gafos assumes that the coordination in Figure 4.2 (d) may be employed in languages like English where acoustic releases are not evidenced between consonants in a cluster. Lastly, Figure 4.2 (c) with no gestural overlap illustrates the cases where two identical consonants at the end of a word (e.g., /tt/) are produced with an epenthetic schwa [t³t].²³

To summarize these, Gafos (2002) proposes that a transitional schwa in coda

²² Gafos also showed through computational simulations with a model of gestural dynamics that the temporal relation appropriate for the perceptual result of this transition is the one in Figure 4.2 (b).

²³ I will not discuss why Gafos proposes two different gestural coordinations, as in Figure 4.2 (b) and (c), to account for an epenthetic vowel produced in two different consonantal sequences, since it is not relevant to the purpose of this dissertation.

clusters in Moroccan Colloquial Arabic arises from a timing relationship in which the articulatory release of the first gesture does not coincide with the target of the second gesture. This leaves a short period of open vocal tract between the two consonants, which gives rise to the transitional schwa.

At this point, it may well be asked how the representation of consonant clusters (Figure 4.1) and the description of an epenthetic vowel under AP as a result of gestural non-overlap (Figure 4.2) are related to the topic in the present study: Korean speakers' production of English consonant clusters. This question can be answered by discussing how the present study views the acquisition of consonant clusters.

In a broad stroke, this dissertation takes a view that language learners produce words (either L1 or L2), using the available set of vocal gestures and are supposed to elaborate individual gestures by differentiating, tuning, and coordinating those gestures. When it comes to second language acquisition of consonant clusters, this dissertation takes the position that L2 learners' successful production of non-native consonant clusters requires assigning the correct coordination relation to consonant sequences in the target language. An epenthetic vowel, one of the typical errors in L2 learners' non-native consonant cluster production, can then be understood as a consequence of failing to correctly organize the gestures in segments. As shown above, Gafos (2002) addresses how the gestural non-overlap induces an epenthetic vowel.

The idea that an epenthetic vowel occurs as a consequence of failing to correctly organize the gestures in segments has found support in Davidson (2003, 2006), which investigated native English speakers' production of Czech consonant clusters. In contrast to the general assumption that L2 learners employ an epenthetic vowel to repair

segmental sequences that are not allowed in the learners' native language, Davidson (2003, 2006) proposes that the inserted schwa is a result of non-overlapping gestural coordination rather than a phonological repair of the illegal consonant clusters.

The following two sections elaborate more on the implication of Articulatory Phonology on the issue of phonological acquisition. In particular, § 4.1.1.5 discusses Davidson (2003) in more detail to illustrate how L2 learners' production of non-native consonant clusters can be captured under the notion of gestural coordination.

4.1.1.4. Phonological Acquisition under Articulatory Phonology

The main tenet of AP, that phonological structures are organized patterns of gestural units, has important implications for language acquisition. Browman & Goldstein (1989: 202) specifically state that adopting gestures as a basic unit of phonological representation would be particularly useful for a child learning to speak.

The evidence that gestures are pre-linguistic units of action is observed in the babbling stage of infants. For instance, Locke (1983) notes that the consonants such as [h, b, d, g, p, t, k, m, n, j, w, s] constitute the core babbling inventory, and there seems to be little language-dependent variation in the babbling inventory and its developmental sequence. This demonstrates that infants' babbling involves the emergence of simple constriction gestures of independent parts of the vocal tract.

It is also interesting to note that particular consonants that were produced with high frequency in the babbling of a given child also appear with high frequency in that child's early productions (Vihman et al. 1985 cited in Browman & Goldstein 1989).

Browman & Goldstein take this observation as evidence of “a definite relationship between the (non-linguistic) gestures of babbling and in early words” (Browman & Goldstein 1989: 204).

Under AP, phonological acquisition is viewed as differentiating, tuning, and coordinating the pre-linguistic gestures so that those gestures can be used as the units of contrast. That is, the pre-linguistic vocal gestures become linguistically significant when the child begins to produce his first few words. The child begins to produce words, using the available set of vocal gestures and is supposed to elaborate the gross status of individual gestures by differentiating, tuning, and coordinating those gestures.

In particular, differentiation and tuning of gestures are required to develop phonemic contrast. For instance, English contrasts alveolar strident /s/ with alveo-palatal strident /ʃ/. The child in an English-speaking environment is supposed to differentiate these two sounds with different values of constriction location. That is, a single pre-linguistic constriction gesture used in the infants’ babbling must eventually differentiate into a variety of potentially contrastive gestures, tuned with different values of both constriction location and degree.

In addition, the various gestures that constitute a given word must be coordinated appropriately. A child’s attempt to produce a certain word is often unsuccessful due to the incorrect arrangement of individual gestures involved in the word. For instance, Studdert-Kennedy (1987: 76) notes that words of a 15-month old girl such as [mã^ʔ, de^{dn}, hɪn, ^mbō, p^hɪn, ba^h, dauⁿ, buā], which are produced in an attempt to say the word *pen* seem radically different, but they can still be analyzed as the set of gestures: glottal opening, bilabial closure, tongue body lowering, alveolar closure, and velum lowering.

That is, the child seems to know what the relevant gestures are for *pen*, but what she lacks is a proper organization of such gestures.

It seems that AP provides a very useful vehicle that allows a direct connection between analytical investigations of articulatory structure and phonological acquisition, as was demonstrated by the 15-month old girl's production of *pen*. Given that child language learners have to train their pre-linguistic gestures to turn them into the units of contrast in the given languages, the phonological patterns specific to the child language can sometimes be accounted for by looking at which aspect of training (i.e., *differentiating, tuning, and coordinating*) is responsible for those phonological patterns.

Consider the case of cluster reduction in child phonology, which seems to demonstrate that the failure of a proper coordination of gestures is responsible for the erroneous pattern. It is well known that young children tend to omit one of the consonants in a consonant cluster (Bernhardt & Stemberger 1998; Goad & Rose 2004; Ingram 1974, 1989; Locke 1983; Smith 1973). When it comes to the question of which segment is deleted, sometimes the first segment in a string is deleted (e.g., [pun] for *spoon*, [tar] for *star*), and in other times the second segment is deleted (e.g., [bu] for *blue*, [ten] for *train*). More interestingly, /s/ + sonorant clusters are reduced either to the first segment, /s/ or to the second one, a sonorant (e.g., [nou] or [sou] for *snow*).

Pater & Barlow (2003) deal with the variable reduction pattern in /s/ + sonorant clusters in terms of constraint re-ranking in child phonology. In particular, they introduce *Fricative (fricative consonant is not allowed) along with onset sonority constraints (e.g., *Nasal-onsets: Nasal onsets are not allowed). Pater & Barlow (2003: 495) find the motivation of the constraint *Fricative in the well-known “stopping” process in child

phonology. A fricative in child phonology is realized as a stop at a certain stage of development (e.g., “zoo” [du:]).

However, it seems that the postulation of *Fricative in the cluster reduction process is undermined by the existence of contradictory processes in child phonology. Bortolini & Leonard (1991: 8) observed that the consonants that resulted from a process such as spirantization (e.g. /p/ → [f]) in Italian child phonology were seemingly identical to the targets that were replaced by stopping (e.g. /f/ → [p]). Based on this observation, Bortolini & Leonard (1991: 11) conclude that the children seemed to possess the ability to produce a particular consonant, but would sometimes replace it with another due to a coarticulation problem with an adjacent segment.

The spirantization process (e.g., /p/ → [f]) seems rather rare in English child phonology. However, in the case of children learning English, we can find a similar process in the production of affricates (/tʃ, dʒ/). When we solely base our assumption on a constraint like *Fricative, we may expect that an affricate is realized as a stop in child phonology. However, contrary to this prediction, the realization of affricates in child phonology is quite variable. Thus, *change* is realized as [deind], [teind], [seind²⁴], or [seindz] (Smith 1973: 218).²⁴ The observation that an affricate is either realized as a stop or a fricative critically undermines the postulation of *Fricative in accounting for the cluster reduction.

The postulation of the *Fricative constraint for cluster reduction seems to completely lose its empirical ground when we look at the data from D. Ohala (1999). Ohala tested sixteen English-learning children (aged 1;1-3;1) on the production of illegal

²⁴ Refer to Smith (1973: 218) for more examples of this variation.

final consonant clusters such as [-pf] and [-fp] and found that these clusters are dominantly reduced to [f] in child's production (p. 411). This kind of empirical data must be very problematic for Pater & Barlow's analysis where *Fricative serves as one of the major motivations for child cluster reduction.

In light of these findings (Bortolini & Leonard 1991; D. Ohala 1999), language learners' production failure of affricates and consonant clusters seems to lie in the articulatory coordination problem rather than the production difficulty of a certain segment. Given that AP includes gestural coordination as one of the core properties (along with gestural tuning and differentiation) that children need to acquire to move from child to adult phonology, the issues of cluster reduction and complex segment realization seem to receive better accounts under the framework of AP.

In the following section, I will extend the implications of AP to the second language acquisition of consonant clusters.

4.1.1.5. Articulatory Phonology and L2 Acquisition of Consonant Clusters

In the previous section, I discussed how phonological acquisition is viewed under AP and what the implications of this view are on the phonological patterns specific to child language acquisition. In particular, I have shown that the well-known child cluster reduction may be attributable to the coarticulation problem of two adjacent segments rather than to the avoidance of producing of a certain segment (e.g., fricative).

In this section, I will demonstrate that second language acquisition of consonant clusters can also benefit from the general approach of AP which links articulatory

structure and phonological acquisition. To this end, I will review Davidson (2003), which illustrated that epenthetic vowels occurring in second language learners' production of non-native consonant clusters are attributable to the gestural non-overlap of the two adjacent segments.

It has been a general assumption that an epenthetic vowel found in L2 learners' production of non-native consonant clusters is phonological in the sense that it is employed to repair segmental sequences that are not allowed in the learners' native language. In contrast to this general assumption, Davidson (2003) maintains that the inserted schwa is a result of non-overlapping gestural coordination (i.e., transitional schwa) rather than a phonological repair of the illegal consonant clusters.

Based on the fact that English allows no obstruent-obstruent onset clusters other than s + stop clusters, Davidson (2003) hypothesized that native English speakers are *unable* to assign the correct coordination relation to onset clusters like /fp/, /zn/, and /vd/. To test this hypothesis, Davidson had native English speakers produce Czech legal consonant clusters such as /fp/, /zn/, and /vd/. As was predicted, it was found that native English speakers make a fair number of epenthesis errors in fC, zC or vC sequences.

Based on the English speakers' production of phonotactically illegal word-initial *clusters*, Davidson (2003: 61) hypothesized that:

Whereas lexical schwa is likely to have an underlying target, a production task in **which** speakers are given consonant sequences that are prohibited from being **coordinated** in the canonical way for English may in fact lead to a different kind **of schwa**. In this case, it is plausible that speakers attempting to correctly produce **these** sequences may nevertheless fail, but do so by using non-overlapping **coordination** to repair them rather than by inserting a schwa in the phonological **output**.

To support her claim that an epenthetic vowel occurs due to the non-overlapping gestures, Davidson implemented an experiment using ultrasound imaging ²⁵ and compared English speakers' productions of legal words with lexical schwa to their productions of non-native initial clusters. The result of this experiment shows that the epenthetic vowel produced by English speakers when they produce illegal consonant clusters is not consistent with a vowel that has its own gestural target, which consequently supports her claim that the epenthetic vowel is transitional rather than phonological.

It may be clear at this point that Davidson's claim on the nature of an epenthetic vowel has a parallel in Browman & Goldstein (1992) and Gafos (2002). Davidson (2003) further contributes to the topic of gestural coordination and an epenthetic vowel by extending the issue to the case of second language acquisition.

However, the notion of gestural non-overlap has an apparent limitation in accounting for *variable* L2 consonant cluster production data. For example, Davidson (2003: 48) found that vowel insertion occurs in the order of /sC/ (least vowel insertion) < /fC/ < /zC/ < /vC/. That is, vowel insertion is most likely to occur after /v/ and least likely to occur after /s/. How can the different rate of vowel insertion depending on cluster type be explained under the notion of gestural non-overlap?²⁶

Davidson (2003) deals with the variable production difficulty found in her data by incorporating the notion of markedness. Thus, Davidson (2003: 49) attributes the frequent occurrence of vowel insertion after /v/ to the fact that /v/ is marked for both

²⁵ **Ultrasound imaging** is a non-invasive technology that allows for the real-time visualization of tongue motion during speech (Davidson 2003: 58).

²⁶ **That vowel insertion** occurred least frequently after /s/ is not surprising because most of the clusters in /sC/ category were legal English consonant clusters (sm, sn, sf, sp, st sk).

stridency (it is better to have stridency in terms of perception) and voicing (a voiced obstruent is typologically more marked than a voiceless obstruent).

The approach taken by this dissertation regarding Korean speakers' production of English consonant clusters is in line with that of Davidson (2003). That is, the notion of gestural coordination is embraced as an important aspect of consonant cluster acquisition. Target-like consonant cluster production presupposes assigning the correct coordination relations to consonant sequences, and an epenthetic vowel occurs as a result of gestural mistiming.

However, gestural mistiming alone cannot account for L2 consonant cluster acquisition data. What needs to be acknowledged is that variation is a fact of life in language acquisition, and there exists corresponding empirical evidence. First of all, we have seen that frequency of vowel insertion varies depending on the cluster type. Second, it seems that gestural non-overlap accounts do not have much to offer for the variable patterns of consonant deletion, another typical type of error in L2 consonant cluster acquisition.

Given this, it is necessary to identify additional phonetic factors contributing to the patterns of vowel insertion and consonant deletion. To account for native English speakers' less accurate production of /vC/ clusters than /zC/ clusters, Davidson (2003) refers to the perceptual saliency of a strident: /v/ is more marked than /z/ by virtue of the poor perceptual cues of a non-strident fricative which has weak-intensity energy.

Following Davidson's (2003) approach to native English speakers' production of Czech consonant clusters, I identify articulatory and perceptual factors defining consonant cluster markedness and incorporate them in the *Phonetically Based Consonant*

Cluster Acquisition Model (PCCAM). This model is empirically tested through the published L1 and L2 consonant cluster acquisition data and Korean speakers' production of English consonant clusters.

In the following section, I propose articulatory factors defining consonant cluster markedness in PCCAM and justify them on theoretical and empirical grounds.

4.1.2. Degree of Gestural Overlap Parameter

The working hypothesis regarding the degree of gestural overlap and L2 acquisition of consonant clusters is that *a consonant sequence with less gestural overlap is easier for language learners to acquire than a consonant sequence with more gestural overlap*. In what follows, I present theoretical and empirical grounds for this hypothesis.

The hypothesis for the degree of gestural overlap and L2 acquisition of consonant clusters is theoretically based on Articulatory Phonology, particularly on the notion that well-formed consonant sequences will be least overlapped and least variable (Browman & Goldstein 1990b). While providing formal representations of articulatory organization for various phonological phenomena such as consonant deletion, consonant insertion, and assimilation, Browman & Goldstein (1990b: 368-369) propose a view that a possible cluster in a language correlates with the degree of gestural overlap and its variability.

The view we are proposing then suggests that the gestural organization in a language is exquisitely tuned to allow the successive oral gestures in syllable onsets and codas to overlap partially, without obscuring the information in these gestures. This view has been propounded by Mattingly (1981), who argues that "the syllable has more than a phonological or prosodic role; it is the means by

which phonetic influences [cf. gestures] are scheduled so as to maximize parallel transmission" (p. 418). As Mattingly suggests, this organization guarantees, for example, that the constriction and release of an [l] will not occur completely during the bilabial closure in a [pla] utterance. However, for sequences that are not possible syllable onsets (or codas), we hypothesize that the production system does not have the same kind of tight organization available. Thus, for such sequences, variation in degree of overlap is possible even to the point of obscuring the parallel transmission of information (i.e. one gesture may hide another).

According to Browman & Goldstein (1990b), the fact that a certain consonant cluster is possible in a language (e.g., /pl/ for English) presupposes that the successive oral gestures in the sequence overlap with each other only to the extent that the overlap does not obscure the acoustic information in the sequence. In contrast, for a consonant cluster that is not allowed in the language, gestural overlap between segments is quite variable and a large degree of gestural overlap in the sequence often obscures the phonetic information in the consonant sequence (and consequently consonant deletion or assimilation occurs).

Supporting evidence for this proposal is found in the observation that alveolar stops frequently assimilate in alveolar-labial and alveolar-velar sequences (*seven plus* → *seve[mp]lus*), but assimilation (in either direction) is rare in labial-alveolar velar-alveolar sequences. Browman & Goldstein (1990b) interpret this asymmetry in terms of syllable structure differences: "the labial-alveolar and velar-alveolar sequences are all possible syllable codas, and thus would be expected to show a tight phasing organization that prevents complete overlap" (p. 369).

Browman & Goldstein did not provide any hypothesis, to my knowledge, for how the degree of gestural overlap in clusters is correlated with the acquisition of those

clusters. However, it does not seem unreasonable to hypothesize that a less overlapped consonant sequence would be easier for language learners to acquire, since Browman & Goldstein (1990b) essentially proposed that well-formed consonant sequences will be least overlapped and least variable.²⁷

The rationale for the correlation between the degree of gestural overlap and consonant cluster acquisition is as follows. We have discussed in the previous sections that non-overlap of gestures results in the percept of an epenthetic vowel (Browman & Goldstein 1992; Gafos 2002; Davidson 2003). On the other hand, when gestures of two segments are overlapped too much, it brings about the percept of consonant deletion or assimilation (Browman & Goldstein 1989). These observations suggest that assigning a correct coordination to a consonant sequence is a non-trivial issue to language learners.

The question that arises is whether there exists a certain degree of gestural overlap with which language learners may produce clusters more accurately. To put the same question another way, if all other factors are equal, will language learners be better at producing a cluster with less degree of gestural overlap or a cluster with more degree of gestural overlap? The proposal in this dissertation is that language learners may find it easier to produce a consonant sequence which requires a less degree of gestural overlap. This preference may be grounded in perceptual reasons: language learners may be better at producing a less overlapped consonant sequence because it is easier for them to provide acoustic information for the sequence.

In other words, given that one of the primary goals of language learning is to

²⁷ It should be acknowledged that there exists a maximum and minimum degree of gestural overlap that is required for two consonants in a sequence to be perceived as a consonant sequence (rather than CV or VC sequence). Thus, “language learners’ preference for a certain degree of gestural overlap” should occur within this maximum and minimum degree of gestural overlap that is required for consonant sequences.

make oneself understood using the language, there may be a tendency for language learners to produce more accurately a segment or segment sequences that guarantee the acoustic information in the segment or segment sequences is clearly transmitted to hearers. The question arises as to what language learners would do if a certain type of consonant sequence involves a large degree of gestural overlap? Recall that a large degree of gestural overlap often obscures the phonetic information in a consonant sequence (and consequently consonant deletion or assimilation occurs).

There may be two strategies for language learners. They may either delete a consonant in a sequence or assimilate the two consonants, while trying to keep the large degree of gestural overlap (they still do not succeed in producing the sequence with target degree of gestural overlap). Or, language learners provide additional acoustic cues to the segments in the sequence to make sure the clarity of their speech to the hearer. For instance, it is quite conceivable that adult L2 learners may prefer to maintain the clarity of their speech, and thus vowel insertion (i.e., acoustic cues) in consonant sequences is more common, while for child language learners, deletion of a segment is more common because they do not master the proper organization of articulatory gestures. This assumption does not seem unreasonable given that vowel insertion in child phonology is not as common as that in second language phonology (Young-Scholten & Archibald 2000).

In sum, language learners may be better at producing a less overlapped consonant sequence because it is easier for them to provide acoustic information for the sequence. Putting the same idea differently, it will be more difficult for language learners to produce a consonant sequence which requires a large degree of gestural overlap because

they may either have to provide additional acoustic cues to guarantee the clarity of their production or they end up deleting one segment due to the failure of achieving the target degree of gestural overlap.

Note also that the proposal that language learners may be better at producing a less overlapped consonant sequence is analogous to Browman & Goldstein's (1990b) statement that a well-formed sequence is least overlapped and least variable because acoustic information can be maintained well in such an environment.

I have not yet found any studies specifically examining the degree of gestural overlap in clusters and its influence on consonant cluster acquisition. However, it is not difficult to find possible empirical evidence (even though it is indirect) showing how the degree of gestural overlap in the consonant sequences influences consonant cluster acquisition. To demonstrate this, I will first introduce Byrd (1996) who examined the spatial and temporal production of English alveolar and velar stops and alveolar fricatives in sequence using electropalatography (EPG).

Different degrees of gestural overlap in obstruent sequences: Byrd (1996)

Byrd (1996) looked at how the temporal overlap in the articulation of two (or more) gestures varies as a function of the consonants' place and manner, and the placement of syllable boundaries. Sequences of the type [C#C], [#CC], and [CC#] are considered, where # is a word boundary.

Three findings of Byrd (1996) are relevant in the present study. First, *manner effect*: it was found that a stop is more overlapped by an adjacent consonant than is a fricative. An ANOVA including the sequences [d#g] and [s#g] (and the reverse order

[g#d] and [g#s]) demonstrates that a stop-stop sequence exhibits significantly more overlap of C₁ by C₂ than a fricative + stop sequence (or stop + fricative sequence).²⁸

Second, *sequence type effect*: Byrd (1996) looked at the degree of gestural overlap in word-final fricative + stop and stop + fricative sequences to determine how the different order of consonant sequences influences the degree of gestural overlap. It was found in the study that fricative + stop ([sk#]) is gesturally more overlapped than stop + fricative ([ks#]) (Byrd 1996: 231). Third, combining the different degree of gestural overlap by manner (a stop is more overlapped by an adjacent consonant than is a fricative) and by sequence type (fricative + stop is more overlapped than stop + fricative), Byrd (1996: 231) finds the following order of the degree of gestural overlap in word-final obstruent clusters: coproduction *decreases* from [gd#] to [sk#] to [ks#].

The findings in Byrd (1996) regarding the degree of gestural overlap in obstruent + obstruent clusters are summarized in Table 4.2.

Table 4.2. Degree of gestural overlap in obstruent + obstruent clusters (Byrd 1996)

	less overlap < more overlap
manner effect	stop + fricative / fricative + stop < stop + stop
sequence type effect	stop + fricative < fricative + stop
manner effect + sequence type effect	stop + fricative < fricative + stop < stop + stop

When it comes to the question of why a stop is more overlapped by a following consonant than is a fricative (the effect of manner on overlap), Byrd (1996: 233-234) conjectures that there may be a perceptual motivation behind this asymmetry.

²⁸ More specifically, Byrd (1996: 222) notes that “the onset of C₂ relative to C₁ is earlier for [d#g] and the time between onsets is less. There is also a trend for the interval from the release of C₁ to the onset of C₂ to be greater for [s#g] than [d#g]”.

The perceptual product and articulatory constraints are not independent. To put the same idea another way, the back of the tongue mustn't raise too early when [s] precedes [g], as this posture would prevent the appropriate airflow through the alveolar constriction thereby impeding the production of the fricative. Speakers avoid doing this thereby ensuring that the fricative will have the acoustic cues which in turn allow it to be accurately recovered by the listener. (In fact, for an [s] to be identifiable it must have approximately 50 ms of frication (Jongman, 1989). For these [s#g] sequences, the mean time between the first maximum contact for [s] and that for [g], a rather rough approximation of the frication noise duration, is 53 ms, suggesting that they might be overlapped as much as possible without impeding identifiability.

That is, the [s] duration should not be much overlapped by a coarticulated stop due to the perceptual recoverability issues. In contrast, a stop is more lenient about the degree of gestural overlap because for a stop, only an edge needs to be discernible: the abrupt change in amplitude is the perceptually salient cue for a stop.

When it comes to the sequence type effect that fricative + stop ([sk#]) is gesturally more overlapped than stop + fricative ([ks#]), Byrd (1996: 231) speculates that the lesser degree of gestural overlap in stop + fricative may be compatible with a prosodic representation postulating marginal [s]: [s] is not tightly bound to the rest of the syllable. However, it should be noted that Byrd qualified her speculation by saying that “however, there is no evidence for the assumption that [d]’s behave similarly, as would be expected given a theory in which all coronal obstruents are considered to be appendices. The [gd#] sequence has shorter latencies than both [ks#] and [sk#]” (p. 231). In other words, that stop + fricative coda clusters involve less gestural overlap than fricative + stop coda clusters is not compatible with the prosodic representation of /s/ as

an appendix because a coronal in [gd#], which can also be considered as an appendix, involves a greater degree of gestural overlap than both [sk#] and [ks#].

Implications: degree of gestural overlap and consonant cluster acquisition

Recall the hypothesis that I proposed regarding the degree of gestural overlap and consonant cluster acquisition: a consonant sequence with less gestural overlap is easier for language learners to acquire than a consonant sequence with more gestural overlap. Gestural overlap data in Byrd (1996) serve as a good reference to test the degree of gestural overlap parameter. That is, if the degree of gestural overlap parameter is heading in the right direction, we may find that language learners will perform better at producing obstruent + obstruent coda clusters in the order of stop + fricative (least overlapped) < fricative + stop < stop + stop (most overlapped).

Do we have empirical evidence for this? The answer is *Yes*. In this section, I will provide a quick look at what happens when English learning children produce English coda clusters. More detailed empirical evidence for the degree of gestural overlap parameter presented in § 4.5.

In what follows, I present three studies that tested L1 acquisition of coda clusters: Dyson (1988), Olmsted (1971), and Wellman et al. (1931). The summary of coda clusters acquired by age 3 years in these three studies is as follows.

Table 4.3 Summary of clusters acquired by age 3 years

Study	# of participants	Consonant clusters in a child's production
Wellman et al. (1931)	n = 204 (2:0-6:0)	ŋz, ns, mz, mp, ps
Olmsted (1971)	n = 16 (3:1-3:7)	nt, nd, ns, nz, ŋz, ?n, mz, mps, ps, ts, dz, ks, gz, pt
Dyson (1988)	n = 20 (2:0-3:3)	ntf, ns, ps

Note in Table 4.3 that nasal + obstruent sequences are dominant in children's speech by age 3 years. When it comes to the obstruent + obstruent clusters, we can see the presence of stop + fricative and the absence of fricative + stop and stop + stop clusters in children's speech (there is only one occurrence of stop + stop sequences (/pt/) in Olmsted 1971).

Similarly, Templin (1957: 51) tested 480 children (age 3:00-8:00) and observed that /ks/ is acquired at the age of 3:5, while /st, sk/ is acquired at the age of 7:0.²⁹ I take this as supporting evidence for the possible correlation between the degree of gestural overlap and consonant cluster acquisition.

Drawing on the notion that well-formed consonant sequences will be least overlapped and least variable (Browman & Goldstein 1990b), I propose the degree of gestural overlap parameter that *a consonant sequence with less gestural overlap is easier for language learners to acquire than a consonant sequence with more gestural overlap*. The degree of gestural overlap parameter determines the relative ranking relationships in the following consonant clusters.

(3) Relative consonant cluster markedness determined by the degree of gestural overlap parameter (A < B means that A is less marked than B)

²⁹ One may wonder whether the frequency of consonant clusters in the ambient language may have affected the child's production. However, this may not be necessarily the case because there does not seem to be much frequency difference between stop + fricative and fricative + stop sequences. For instance, American English speaking adults' ten most frequently produced word-final clusters reported in three studies are as follows. French et al. (1930): /nt, nd, st, ts, ŋk, ld, rz, ks, kt, rd/; Dewey (1923): /nd, nt, st, rd, ns, nz, ks, kt, ld, rt/; Roberts (1965) /nt, st, nd, rz, nts, rd, ld, rn, kt, ŋk/ (these three studies are cited in Jakielski 1998: 29). Two of the studies (Dewey 1923; French et al.) include both stop + fricative and fricative + stop sequences, while Roberts (1965) includes only a fricative + stop sequence.

Onset clusters

- a. /s/ + liquid < stop + liquid (e.g., sl < pl): /s/ + liquid involves less gestural overlap than stop + liquid, and is thus less marked.

Coda clusters

- a. sonorant + strident < sonorant + stop (e.g., ls < lt): sonorant + strident involves less gestural overlap than sonorant + stop, and is thus less marked.
- b. stop + strident < strident + stop (e.g., ps < sp; bz < zd): stop + strident involves less gestural overlap than strident + stop, and is thus less marked.
- c. strident + stop < stop + stop (e.g., sp < pt; zd < bd): strident + stop involves less gestural overlap than stop + stop, and is thus less marked.

4.1.3. Articulatory Release Effect

The articulatory release effect in PCCAM hypothesizes that *a consonant sequence is easier for language learners to acquire when a stop is followed by [+continuant] than when a stop is followed by [-continuant]*.

First of all, notice that the term, *consonant release*, is used in a rather narrow sense in this dissertation: “consonant release yields a phonetic burst, a perturbed postconsonantal airstream that clarifies voicing and place of articulation contrasts” (Hudson 1995: 655). In what follows, I will justify the articulatory release effect based on Côté (2000).

Hungarian has an optional process of cluster simplification in a sequence of three consonants. What is interesting in the Hungarian case is that not all three consonant sequences are the targets of simplification. Thus, while a stop is the main target of

deletion in a three consonant sequence (fricatives and affricates are never deleted), stop deletion does not occur if a stop is followed by a [+continuant] segment (glides, liquids, and fricatives). Relevant examples are presented in (4).

(4) Hungarian cluster simplification (Côté 2000: 38-39)

a. <i>aktfotó</i>	[ɔktfotó:]	*[ɔkfoto:]	‘nude photograph’
b. <i>hangsor</i>	[hɔŋkʃor]	*[hɔŋʃor]	‘sound sequence’
c. <i>handlé</i>	[hɔndle:]	*[hɔnle:]	‘second-hand dealer’
d. <i>ámbra</i>	[a:mbrɔ]	*[a:mrɔ]	‘ambergris’
e. <i>eszpresszó</i>	[ɛspreso:o:]	*[ɛsres:o:]	‘espresso’
f. <i>centrum</i>	[tʰɛntrum]	*[tʰɛnrum]	‘center’
g. <i>templom</i>	[tɛmplom]	*[tɛmlom]	‘church’

Note in (4) that stops do not delete if followed by [+continuant]: fricatives (4a-b) and liquids (4c-g).

Regarding the pattern that a stop is not deleted when followed by [+continuant], Côté proposes that the role of the [continuancy] value of the following element in stop deletion can be related to the audibility of the release burst. Côté (2000: 138) first identifies a context where a stop is either unreleased or saliently released.

There is a well-known tendency for stops to be unreleased or to lack an audible release in certain contexts. Based on Henderson & Repp (1982), we can usefully distinguish between stops with and without a release that has an observable effect in the acoustic signal. Stops without an acoustically present release actually comprise two distinct types: strictly unreleased and silently-released stops. Articulatorily unreleased stops occur before homorganic nasal or oral stops and utterance-finally. In the first case the constriction is maintained through the

following consonant; utterance-finally it may be delayed. Silently-released stops are found before an oral or nasal stop with a more front articulation. When the closure of the second consonant is made before the release of the first stop, this release has no acoustic effect since the air is trapped behind the front constriction.

Côté (2000) assumes, based on the work of Henderson & Repp (1982), that when a stop is followed by another stop (including a nasal), it is either articulatorily unreleased or silently released. The perceptibility of a stop is crucially dependent on its release burst, and thus unreleased or silently released stops are perceptually very weak.

However, unreleased and silently-released stops, according to Côté (2000), are not found if the stop is followed by a segment that does not involve a complete closure in the oral cavity. [+continuant] specification basically means that there is always an outgoing flow of air, which can carry the effect of the release.

With the stop deletion pattern in Hungarian and the variable acoustic release of a stop depending on the identity of a following segment in hand, Côté (2000) proposes the following generalization: stops that are not followed by a [+continuant] segment want to be adjacent to a vowel, and preferably followed by a vowel. Given that [+continuant] specification of a following segment can provide perceptual cues to the preceding stop, it is preferable to be followed by a [+continuant] segment so that the release of a preceding stop will not be devoid of an acoustic effect.

PCCAM incorporates the basic idea of Côté (2000) on the role of [+continuant] in obstruent sequences and predicts that consonant sequences are easier for language learners to acquire when a stop is followed by [+continuant] than when a stop is followed by [-continuant]. Note that the articulatory release effect only applies to the coda consonant clusters because English does not have onset consonant clusters with stop +

fricative or stop + stop sequences. The articulatory release effect determines the following consonant cluster markedness relationship.

(5) Relative consonant cluster markedness determined by the articulatory release effect

(A < B means that A is less marked than B)

Coda clusters

- a. stop + strident < stop + stop (ps < pt; bz < bd)

4.1.4. Articulatory Stridency Effect

This section discusses the articulatory factor derived from the articulatory configuration of the voiceless strident /s/, which blocks an occurrence of an epenthetic vowel. I will label this the *Articulatory Stridency Effect*.

The empirical support for the articulatory stridency effect is found in the vowel devoicing process. The greater frequency of vowel devoicing was observed after a voiceless strident in Imai (2004) for Japanese and in Jun & Beckman (1993), Jun, Beckman, and Lee (1998), and Jun et al. (1998) for Korean.

Consider the case of Japanese vowel devoicing. In the study involving 42 Japanese speakers from the Tokyo area, Imai (2004) examined the influence of the preceding and following consonant on the frequency of vowel devoicing in Japanese. The results of this study show that *a preceding fricative* (/s/ and /h/ in Imai's study) and *a following stop* (/p, t, k/) are most likely to promote devoicing. The ordering of the

preceding and following consonants which promote devoicing is presented in Table 4.4.

Table 4.4. Influence of an adjacent consonant on Japanese vowel devoicing (Imai 2004: 70)

	Promoter of devoicing < Demoter of devoicing
Preceding consonant	fricative < affricate < stop
Following consonant	stop < affricate < fricative

Table 4.4 shows that vowel devoicing is most likely to occur when the preceding consonant is a fricative and least likely to occur when the preceding consonant is a stop. On the other hand, from the perspective of a following consonant, vowel devoicing is most likely to occur before a stop and least likely to occur before a fricative.

As was noted by Imai (2004), the ordering for preceding and following consonant is a complete reversal, which demonstrates an interesting asymmetry. Also, note that a fricative + stop sequence would be the most favorable environment for vowel devoicing given that a fricative is the promoter of vowel devoicing as a preceding consonant, and a stop is the promoter of vowel devoicing as a following consonant. A similar finding is observed in Korean. Jun et al. (1998) observed that vowel devoicing is most likely to occur when the preceding consonant is a fricative (/s, h/), and the following consonant is a stop.

The question arises as to why vowel devoicing occurs frequently after a voiceless strident, in particular, in a sequence of /s/ + voiceless stop. Regarding the greater frequency of vowel devoicing after a voiceless strident, Beckman (1994) conjectures that the pattern is derived from the articulatory characteristics of /s/. It was reported that the production of /s/ involves an unusually abducted glottis in comparison to voiceless stops due to the necessity of a high rate of airflow (Collier et al. 1978; Hirose & Gay 1972;

Yoshioka, Löfqvist, and Hirose 1982). Given that the vocal folds need to come together to bring about vocal fold vibration for voicing, it is plausible to assume that the wider the glottal opening of an adjacent consonant, the more difficult it is to achieve voicing due to the encroachment of the gesture of the adjacent consonant into the voicing source of the vowel. The proposal that the degree of glottal opening in the preceding consonant is correlated with the frequency of vowel devoicing is also found in Fujimoto (2004).

In addition, articulatory movements into and out of a fricative constriction are often slower than movements into and out of a homorganic stop closure (Kuehn & Moll 1976). According to Beckman (1994: 58), vowel devoicing is most likely to occur after a voiceless strident because the aerodynamic conditions of /s/ remain conducive to the continued turbulence of a strident rather than vocal fold vibration.

I find parallels between vowel devoicing and vowel insertion processes because a) both processes are concerned with the voice source of a vowel and b) both vowel devoicing and vowel insertion processes are crucially affected by the identity of adjacent segments. Following Beckman's proposal that vowel devoicing is most likely to occur after a voiceless strident due to the articulatory and acoustic characteristics of /s/, PCCAM predicts that vowel insertion is less likely to occur after a voiceless strident. This phonetic factor is labeled the *Articulatory Stridency Effect* in the present study.

In addition, PCCAM predicts that vowel insertion is least likely to occur in /s/ + voiceless stop due to the existence of the strongest articulatory stridency effect in this sequence. There are two rationales for this assumption. First, the insertion of an epenthetic vowel in /s/ + voiceless stop would require a rapid change from a very abducted glottis (/s/) to a brief moment of voicing for the epenthetic vowel to, again, an

abducted glottis for the voiceless stop (Byrd 1996: 239), which seems to be an articulatorily challenging task. Second, we have seen that vowel devoicing is most likely to occur in this sequence in Korean and Japanese (Imai 2004: Jun et al. 1998).

When we juxtapose the frequency of vowel devoicing in Japanese and Korean with the frequency of an epenthetic vowel in L2 acquisition of consonant clusters, we find an interesting connection between these two phenomena. For instance, Kim (2000) tested 16 Korean speakers on their production of English consonant clusters. One of the interesting findings of Kim (2000) is that no vowel insertion occurs in /s/ + voiceless stop onset clusters (/sp, st, sk/). Note that this finding is problematic for theoretical frameworks such as Minimal Sonority Distance parameter and Sonority Dispersion Principle because *voiceless obstruent* + *voiceless obstruent* is predicted to be the most difficult cluster type in onset clusters under these frameworks.³⁰ According to MSD, voiceless obstruent + voiceless obstruent should be one of the most difficult clusters (other than voiced obstruent + voiced obstruent) due to the small sonority distance between adjacent segments. In addition, SDP states that, for initial demisyllables, it is optimal to have maximal contrast among members of clusters, but voiceless obstruent + voiceless obstruent sequences demonstrate a sonority plateau. Kim (2000: 88-89) seems to be aware of this contradiction and states that:

I basically accept Broselow & Finer's (1991) Minimal Sonority Distance Parameter-Setting [MSD] and Eckman & Iverson's (1993) prediction based on Clement's (1990) Theorem of Sequential Markedness. However . . . Korean high vowels can be easily devoiced or omitted acoustically between two voiceless

³⁰ Voiceless obstruent + voiceless obstruent is the grouping in Kim (2000) for both onset strident + stop clusters (sp, st, sk) and coda strident + stop (sk), stop + strident (ks), and stop + stop (kt) clusters.

obstruents (Jun 1993; H-S Kim & Jongman 1996). This can work as a positive transfer to production of English onset clusters by Korean speakers. That is, even though a sequence of “voiceless obstruent + voiceless obstruent” has an MSD setting of 1, it can be predicted to be the least difficult cluster type for Koreans to acquire.

Simply put, Kim (2000) attributes the exceptional behavior of /s/ + stop in his study to positive L1 transfer (i.e., vowel devoicing between two voiceless obstruents).

Even though Kim’s conjecture on the transfer of L1 vowel devoicing to L2 cluster production seems to be heading in the right direction, attributing non-occurrence of an epenthetic vowel in a certain context to learners’ first language transfer cannot explain why a similar pattern occurs in child language acquisition (Freitas & Rodrigues 2003) and in the case of L2 learners with other first language background (Chen 2003; Major 1996).

Thus, PCCAM adopts the term articulatory stridency effect (rather than positive L1 transfer) based on the assumption that what brings about the epenthetic vowel blocking effect is not grounded in the first language phonological pattern, but grounded in the phonetic configuration of /s/ and its combinatory effect with a following stop. The consequence of this assumption is that vowel insertion is least likely to occur in /s/ + stop sequences regardless of the L2 learners’ native language.

To summarize this section, I identified the articulatory stridency effect as an articulatory factor defining the relative consonant cluster markedness. The articulatory stridency effect hypothesizes that vowel insertion is less likely to occur after a voiceless strident because the aerodynamic conditions of /s/ remain conducive to the continued turbulence of a strident rather than vocal fold vibration. I have used the vowel devoicing

phenomenon in Japanese and Korean as an initial source of establishing the articulatory stridency effect and have also suggested that there is a connection between the vowel devoicing and the vowel insertion processes. In addition, it is predicted that the articulatory stridency effect exerts the greatest power in /s/ + voiceless stop.

The articulatory stridency effect determines the following consonant cluster markedness relationships.

(6) Relative consonant cluster markedness determined by the articulatory stridency effect

(A < B means that A is less marked than B)

Onset

- a. /s/ + stop < /s/ + nasal or /s/ + liquid (sp < sm, sl)³¹
- b. /s/ + liquid < stop + liquid (sl < pl)

Coda

- a. sonorant + /s/ < sonorant + stop (ls < lp)
- b. voiceless stop + /s/ < stop + stop (ps < pt)

4.1.5. Articulatory Voicing Effect

The articulatory voicing effect in PCCAM hypothesizes that vowel insertion is more likely to occur after a voiced obstruent than after a voiceless obstruent because of the relative difficulty of sustaining voicing in an obstruent. The following section discusses

³¹ It is conceivable that the degree of the articulatory stridency effect may be different between /s/ + nasal and /s/ + liquid clusters. However, relevant data are not available to establish the degree of the articulatory stridency effect in these two sequences at this point. I leave this issue for future study.

the aerodynamics of voicing in obstruents to provide a background for the articulatory voicing effect.

Aerodynamics of voicing in obstruents

Voicing is the acoustic result of vocal fold vibration. In order for voicing to occur there are two basic requirements. First, the vocal cords must have the appropriate degree of tension and the appropriate degree of adduction. In particular, vocal fold vibration is produced when air is pushed from the lungs through *a closed but not tightly compressed glottis*. Second, there must be air flowing through the vocal cords, and transglottal pressure must be maintained in such a way that air pressure is lower above the glottis than below it. In this way, a positive airflow occurs through the glottis from the lungs. When the airflow falls below a certain level, voicing will cease (Catford 1977: 29).

It has been observed that maintaining voicing in stops faces an inherent obstacle. One reason is because the supraglottal airflow is not freely vented to the outside world due to the complete airstream blockage during stop closure. Consequently, the airflow through the vocal cords accumulates in the oral cavity, and the oral air pressure eventually reaches the same level as the subglottal pressure (recall the transglottal requirements for voicing).

Moreover, in order for the vocal fold to vibrate during at least a portion of the closure interval, the volume of the vocal tract needs to be expanded. There are several ways of expanding the volume of the vocal tract. Lowering the larynx contributes to the expansion, and the volume increase can also be achieved by raising the soft palate (Stevens 1998). Yet, during stops, expanding the vocal tract to maintain voicing is not

easily achieved. There is not enough time to manipulate the vocal tract due to the short duration of stops (Hayes & Steraide 2004: 9).

Then what about fricatives? Would fricatives be less affected by the requirements for aerodynamics of voicing because at least some of the air accumulating in the oral cavity escapes, thus tending to keep the oral pressure low? Ohala (1997: 2) notes that this is not the case, and “there is statistically an even greater tendency for fricatives to favor voicelessness than stops (Ohala 1983, 1994): whereas an estimated 24% of the world’s languages have only voiceless stops, some 38% have only voiceless fricatives.” The reason for this, according to Ohala (1997: 2), lies in the aerodynamic requirement to produce friction.

Here one must try to make the pressure drop across the oral constriction as large as possible because that is one of the principal determinants of the air velocity through the narrow constriction which leads to fricative noise. One has no control over the atmospheric pressure, so one must try to keep P_{oral} [oral pressure] as large as possible. Now it is evident why voiced fricatives face an inherent problem: for voicing the P_{oral} should be as low as possible, whereas for friction, it should be as high as possible. Obviously, one can’t satisfy both demands optimally. Some middle ground is possible, of course, but to the extent that a voiced fricative has good voicing, its fricative energy is less and to the extent it has good friction, its voicing may be threatened.

It should be clear that aerodynamic voicing constraints render a serious problem for maintaining voicing in obstruents (both stops and fricatives). With these phonetic observations in hand, PCCAM predicts that an epenthetic vowel is more likely to occur after a voiced obstruent because this helps to maintain voicing in a voiced obstruent, which otherwise could be devoiced in language learners’ production. In other words,

PCCAM predicts that language learners may use an epenthetic vowel as a strategy to maintain voicing in a voiced obstruent.

Supporting evidence for this proposal is found in child language phonology. While discussing various phenomena related to voicing contrast in child phonology, Menn (1983: 18) notes that:

Other rule strategies besides the use of voicing or devoicing rules can be found in children obeying the voicing constraint. We have just mentioned Jacob's occasional deletion of final /g/, but there are much more interesting cases to be found. There are the children who add extra segments in order to render a voicing contrast. It has been claimed that some children add a vowel to the end of a word with a final voiced stop; this brings the sound into the interior of the word where it could be managed. "Bag" might be produced as [bægə] or [bægæ].

Menn (1983) observed that children sometimes add an extra segment (i.e., epenthetic vowel) after a voiced stop to render a voicing contrast. An apparent relationship is found between L2 acquisition and child language acquisition: vowel insertion after a voiced stop renders a final stop, which otherwise could be devoiced, into a voiced stop.

Parker's (1981) discussion of the historical aspect of final devoicing is also informative for the discussion of how the insertion of an epenthetic vowel is related to maintaining voicing in an obstruent. Parker's historical accounts on how final obstruent devoicing occurs are presented in (7). The stages of word-final devoicing proposed by Parker (1981) can be summarized as follows. Stage I: open syllable structure such as CVCV is the most "natural" type (Schane 1972: 207-9), and the unmarked form of a stop is [-voice]. State II: intervocalic consonants become voiced. Stage III: Unstressed vowels

are reduced to [ə], and Stage IV: [ə] in final position is lost.³² In Stage V, final stops are devoiced.

(7) Historical aspect of final devoicing (Parker 1981: 131; Rhee 1998: 124)

Stages		
I	CV CV [-vcd]	→ unmarked stage
II	CV CV [+vcd]	→ intervocalic voicing
III	CV Cə [+ vcd]	→ unstressed vowel reduction
IV	CV C unreleased [+vcd]	→ loss of [ə] causes final voiced stop unreleased
V	CV C [-vcd]	→ devoicing in unreleased position

Here it is important to note the role of stop release in final devoicing. Parker (1981: 132) states that

Since stop releases resemble [ə] acoustically (Malécot 1958: 372), the pressure of shedding final [ə] also forced the remaining stops to be unreleased. This left these stops in an acoustically unstable environment. That is, since the cues for stops are carried by adjacent vowels, then loss of the following [ə] (and the

³² Parker (1981) also notes that the loss of [ə] has occurred in all the Germanic languages (Hooper 1976: 102; King 1976: 21; Lehmann 1962: 161) and in Gallo-Romance (Chen & Wang 1975: 269; Hooper 1976: 104-5). Moreover, it was assumed that this process caused "thousands" of voiced stops to show up in final position in German (Stampe 1969: 453) (the references are cited in Parker 1981).

concomitant loss of release), wiped out half the cues for the post-vocalic voiced stops.

Simply put, the loss of [ə] brings a voiced stop into a final, unreleased position, which results in the weakening of the voiced stop cues. Accordingly, such stops are devoiced.

If Parker's reasoning for the historical stages of final devoicing is correct, we can assume the existence of a backward phonological process (at least for the case of language development). That is, a voiced obstruent would be saved from production failure (which results in devoicing) with the insertion of an epenthetic vowel. Data from child language, L2 acquisition, and historical accounts of final devoicing all indicate that vowel insertion is typically related to maintaining voicing in an obstruent.

Yet, the proposal that an epenthetic vowel occurs as a function of maintaining voicing in an obstruent still seems unsatisfactory because we still do not know how the epenthetic vowel is formed in the first place. In other words, how does a language learner's effort to sustain voicing result in inserting an epenthetic vowel? The next section addresses this question.

Release and an Epenthetic Vowel

While discussing the relationships between synchronic sound patterns and sound change with sources in articulatory variation, Blevins (2004: 146) proposes that final vowel epenthesis occurs as a consequence of gestural strengthening release.

One process, however, which is not often discussed is that . . . the case where a consonant release is strengthened, and as a result, is interpreted by the listener as a vowel. This sound change is often referred to as paragoge or final-vowel

epenthesis, and has occurred independently in the history of many Austronesian languages, including Leti, Dobel, and Kambera (Blevins and Garrett 1998: 542). In Dobel, for example, an excrescent [u] occurs after word-final stops. A similar development has occurred in some Sulawesi languages (Mills 1975a, 1975b) . . .

According to Blevins, some instances of final epenthetic vowels are attributable to the phonologization of audible release. In other words, “hyperarticulation provides the strongly released final consonants, while misperception allows the release of these consonants to be reinterpreted as reduced vowels” (Blevins 2004: 156).

Blevins (2004) supports her interpretation of a final epenthetic vowel (as the case of a phonologization of audible release) by demonstrating that some languages are actually going through the pre-phonologized stage of this development. The Kambera language demonstrates this case. In this language, roots ending in consonants receive a final epenthetic vowel [u], but this vowel is referred to as “weak” and it may disappear altogether in rapid speech (Blevins 2004: 156).³³

Does an epenthetic vowel found in L2 acquisition of consonant clusters occur as a consequence of gestural strengthening release as in the case of a final epenthetic vowel in Austronesian languages? Blevin’s accounts are certainly appealing, yet it seems that there exists a difference between the final epenthetic vowel identified by Blevins (2004) and the epenthetic vowel produced by second language learners.

I propose that an epenthetic vowel occurring when L2 learners produce consonant

³³ The fact that the epenthetic vowel disappears in rapid speech in Kambera is particularly interesting because we can find a similar phenomenon in second language acquisition of consonant clusters. Lin (2003) tested Chinese speakers’ production of English consonant clusters and found that epenthesis/deletion ratio (proportion of epenthesis to deletion) is higher in more formal styles (and thus slower) (e.g., minimal pair, wordlist reading) than in less formal styles (e.g., conversation) (Lin 2003: 457). Refer to § 2.2 for the discussion of this study.

clusters is sometimes formed as a consequence of acoustic release aimed at providing acoustic cues to a segment (rather than misperception of an audible release).³⁴ This explains why vowel insertion is most likely to occur after a voiced stop (Kim 2000; Kwon 2005).

Here is a scenario³⁵ of how an epenthetic vowel occurs in a voiced stop context. Imagine that language learners (L2 learners and children) attempt to produce a voiced stop. Language learners may find that their effort to make their production audible is doubly burdened because a voiced stop has very weak internal and external cues, and it is difficult to maintain voicing in stops due to aerodynamic voicing constraints.³⁶ What would be the language learners' strategy to get around this difficulty? The scenario is that language learners may release a voiced stop to provide some external cues to it, and the acoustic release of a voiced stop reaches to the point where that acoustic release is turned into an epenthetic vowel.

To summarize thus far, two assumptions have been proposed to justify the articulatory voicing effect. First, maintaining voicing in obstruents is a challenging task, and language learners get around this difficulty by inserting an epenthetic vowel after a voiced obstruent. Second, an epenthetic vowel is formed as a consequence of an acoustic release aimed at providing acoustic cues to a segment. A voiced stop has very weak internal and external cues especially when it is not flanked by a vowel. To provide some acoustic cues to a voiced stop, language learners may release a stop to the point where an

³⁴ Note that acoustic release is not the one and only reason for vowel insertion. I have already shown that an epenthetic vowel can also occur as a function of gestural non-overlap. Refer to § 4.1.1.3 and § 4.1.1.5.

³⁵ This is a "scenario" in the sense that I don't attempt to prove it in this dissertation.

³⁶ Voiceless stops in English are in a better position in terms of their perceptibility because aspiration of a voiceless stop may provide some external cues to its perceptibility.

epenthetic vowel is formed. This accounts for why vowel insertion is most likely to occur after a voiced stop.

Before closing this section, it should be noted that this dissertation does not provide a detailed articulatory mechanism of how the acoustic release is acoustically turned into an epenthetic vowel. This is one of the limitations of this dissertation and I leave this issue for future research.

The articulatory voicing effect determines the following consonant cluster markedness relationships.

(8) Relative consonant cluster markedness determined by the articulatory voicing effect

Onset

- a. voiceless stop + liquid < voiced stop + liquid (pr < br)

Coda

- a. sonorant + voiceless strident < sonorant + voiced strident (ns < nz)
- b. sonorant + voiceless stop < sonorant + voiced stop (lt < ld)
- c. voiceless stop + voiceless stop < voiced stop + voiced stop (pt < bd)
- d. voiceless stop + voiceless strident < voiced stop + voiced strident (ps < bz)
- e. voiceless strident + voiceless stop < voiced strident + voiced stop (st < zd)

The articulatory factors defining consonant cluster markedness in PCCAM are summarized in Table 4.5.

Table 4.5. Summary of articulatory factors in PCCAM

less marked < more marked	
Degree of gestural overlap	
Onset	a. /s/ + liquid < stop + liquid (sl < pl)
	a. sonorant + /s/ < sonorant + stop (ls < lt)
Coda	b. stop + strident < strident + stop (ps < sp; bz < zd)
	c. strident + stop < stop + stop (sp < pt; zd < bd)
Articulatory release effect	
Onset	a. stop + strident < stop + stop (ps < pt; bz < bd)
Articulatory stridency effect	
Onset	a. /s/ + stop < /s/ + nasal, /s/ + liquid (sp < sm, sl)
	b. /s/ + liquid < stop + liquid (sl < pl)
Coda	a. sonorant + /s/ < sonorant + stop (ls < lt)
	b. stop + /s/ < stop + stop (ps < pt)
Articulatory voicing effect	
Onset	a. voiceless stop + liquid < voiced stop + liquid (pl < bl)
	a. sonorant + voiceless stop < sonorant + voiced stop (lt < ld)
	b. sonorant + voiceless strident < sonorant + voiced strident (ns < nz)
Coda	c. voiceless stop + voiceless stop < voiced stop + voiced stop (pt < bd)
	d. voiceless stop + voiceless strident < voiced stop + voiced strident (ps < bz)
	e. voiceless strident + voiceless stop < voiced strident + voiced stop (st < zd)

4.2. Perceptual Factors in PCCAM

This section discusses the perceptual factors defining consonant cluster markedness in PCCAM.

(9) Perceptual factors

- a. Perceptual stridency effect: vowel insertion is less likely to occur after the voiceless strident /s/ because an epenthetic vowel after a voiceless strident is perceptually more obtrusive than an epenthetic vowel after a stop (Steriade 2001b).

- b. Perceptual voicing effect: vowel insertion is more likely to occur after a voiced obstruent because an epenthetic vowel after a voiced obstruent is perceptually less obtrusive than an epenthetic vowel after a voiceless obstruent.
- c. Principle of perceptual saliency: the likelihood of vowel insertion and consonant deletion correlates with the perceptual salience of a segment in contrast: the weaker the perceptual cues of a segment, the more likely it is that vowel insertion or consonant deletion occurs. Thus, a stop is the target of vowel insertion and consonant deletion due to its relatively weak perceptual salience (Côté 2000).
- d. Maximal perceptual contrast: language learners prefer a sequence where two segments in an alternating sequence make a maximal perceptual contrast.

Note that the perceptual stridency effect, the perceptual voicing effect, and the principle of perceptual saliency are grounded in the P-map hypothesis (Steriade 2001a, 2001b) that a sound change whose modification is less perceptible is more likely to occur than a sound change whose modification is more perceptible. Thus, I will review the P-map in § 4.2.1 to give a general background of the perceptual principles in PCCAM. Following the review of the P-map, the perceptual stridency effect (§ 4.2.2), the perceptual voicing effect (§ 4.2.3), the principle of perceptual saliency (§ 4.2.4), and maximal perceptual parameter (§ 4.2.5) are justified.

4.2.1. Theoretical Framework: P-map

In this section, I present the P-map (Steriade 2001b) from which I develop the perceptual

factors for the *Phonetically-Based Consonant Cluster Acquisition Model* (PCCAM).

The review of the P-map in this section will be mainly based on Steriade (2001b). The concept of the P-map (P stands for perceptibility) was originally proposed as an attempt to solve the “Too-Many-Solutions” conundrum (Steriade’s term) which arises when the system of constraints and rankings predicts too many solutions for a given phonotactic problem. The case of the obstruent devoicing process demonstrates this problem. A phonotactic constraint on obstruent voicing as in (10) is cross-linguistically well-evidenced.

- (10) A phonotactic constraint (Steriade 2001b: 2)
*[+VOICE]/_: voiced obstruents do not occur at the end of the word.

The crucial observation on which the P-map proposal is based is that even though there are many possible ways of repairing the phonotactic constraint on word-final voicing obstruents (e.g., voiced obstruent devoicing, nasalization, deletion, metathesis, or post-voiced obstruent epenthesis), only devoicing occurs as a reaction to *[+VOICE]/_ violation. The question then arises as to what selectional mechanism works for favoring the ‘devoiced’ output instead of the alternative options for modifying the input.

Steriade (2001b) finds the answer to this question in the P-map, a set of statements about perceived distinctiveness-differences between different contrasts in different contexts. The optimal output, according to Steriade, would be the one that differs minimally from the input when complying with the phonotactic constraints. Following this reasoning, other alternative outputs are avoided because they result in drastic input-output differences.

Two properties of the P-map are critical in understanding its mechanism: *positional effects* and *contrast*. Positional effect is the notion that the distinctiveness of a segment is affected by the syntagmatic context. That is, contrasts in certain phonological contexts show more faithful effects than contrasts in other contexts (e.g., perceptual saliency of voicing contrast may vary depending on the context where it occurs). Contrast refers to the idea that “*a* is more perceptible than *b*” means “*a* is more reliably distinguished from a reference term *x* than *b* is distinguished from *x*” (Steriade 2001b: 15).

Based on the main tenet of the P-map hypothesis discussed above, Steriade (2001b) proposes that devoicing is chosen as an optimal output rather than other alternative outputs, since the pair [tæb– tæp] (devoicing) is perceptually more similar than other possible modifications (e.g., nasalization [tæb– tæm], deletion [tæb– tæ], metathesis [tæb– bæt], or post-voiced obstruent epenthesis [tæb– tæbʰ]). All of the other modifications result in greater input-output dissimilarity than devoicing; therefore, they are systematically avoided (Steriade 2001b: 4).

Yet, there still remains a question of how we know that devoicing of an obstruent differs minimally from the input. It is crucial for the P-map proposal to show the existence of perceived similarity differences associated with different contrasts in different positions. In the same vein, it should also be shown that “perceived degree of similarity differences correlate with choices made in phonological systems between alternative options of modifying an input” (Steriade 2001b: 6).

When it comes to the question of what factors determine the relative similarity of different contrasts, Steriade follows an inductive approach to similarity in which if the pair *z-w* causes more confusion than the pair *x-y*, then *z-w* is more similar than *x-y*:

Simply put, the more confusable the pair is, the more similar it is.

Returning to the issue of similarity between voiced and devoiced obstruent pairs, there is abundant evidence that voicing contrast is perceived as less distinctive than contrast based on obstruency differences. For instance, in Greenberg and Jenkins (1964), participants were asked to list non-sense words associated to nonsense stimuli like [klæb], and the result shows that the most common responses involved voicing changes. Thus, for [klæb], the most commonly mentioned forms were [klæp] (23/46 responses), while the other potential associate, [klæm], which also differs by exactly one feature from the stimulus, was elicited less frequently (11/46 responses). This result clearly shows that difference in nasality ([klæb]- [klæm]) is more distinctive than difference in obstruent voicing ([klæb]-[klæp]).

To summarize, the P-map is a mental representation of the degree of distinctiveness of different contrasts in various positions. The crucial role of the P-map in the sound system is that it performs the function of guiding the speaker in search of the relatively minimal input deviation that solves a phonotactic problem.

It is very important to note that the three perceptual factors defining relative consonant cluster markedness in PCCAM, the perceptual stridency effect, the perceptual voicing effect, and the principle of perceptual saliency, all share the same basic working mechanism with the P-map hypothesis. For instance, the perceptual stridency effect hypothesizes that vowel insertion is less likely to occur after /s/ because an epenthetic vowel after /s/ is perceptually more obtrusive than an epenthetic vowel after a stop. On the other hand, the perceptual voicing effect hypothesizes that vowel insertion is more likely to occur after a voiced obstruent than after a voiceless obstruent because an

epenthetic vowel shares vocalic properties with a voiced obstruent (and therefore an epenthetic vowel is perceptually less obtrusive in a voiced environment). The principle of perceptual saliency (Côté 2000) is also in line with the P-map hypothesis in that the perceptibility of a certain segment plays a crucial role in the process of vowel insertion and consonant deletion.

It should be clear that the main working mechanism of perceptual factors (except the maximal perceptual contrast parameter) in PCCAM is that vowel insertion and consonant deletion are more likely to occur when the output differs minimally from the input of the target language.

The following four sections (§ 4.2.2-§ 4.2.5) justify the four perceptual factors in PCCAM.

4.2.2. Perceptual Stridency Effect

On the theoretical side, the perceptual stridency effect is proposed based on the hypothesized P-map for vowel epenthesis (Steriade 2001b) which shows that vowel insertion in a voiceless strident + obstruent sequence may cause serious dissimilarity between unepenthesized input and epenthesized output.

Table 4.6. P-map for vowel epenthesis (Steriade 2001b: 18)

Zero/vowel	T Ri	S T
Ø/ '	Ø/ '	Ø/ '
Ø/u	Ø/u	Ø/U
Ø/i	Ø /i	Ø/i

Note: bigger letter = less similar to unepenthesized input (' = schwa)

In Table 4.6, “T” stands for an obstruent, “R” stands for a sonorant, and “S” stands for a strident. Notation Ø /u in each cell indicates the perceptual difference between the unepenthesized input (Ø) and vowel epenthesized output (’, u, i). Differences in the size of the notations (e.g., ø u vs. Ø/U) indicate the perceptual similarity difference between input and output: the bigger the size of the letters, the less similar the epenthesized output is to unepenthesized input. For instance, compare the size of Ø/u, Ø/’ and Ø/i under the column of T_Ri and S_T. We can see that the letters under the column of S_T are bigger than those under the column of T_Ri. This is interpreted such that epenthetic vowels /u, i, ’/ in the environment of S_T make the epenthesized output less similar to the unepenthesized input than those in the environment of T_Ri. In short, Table 4.6 shows that vowel insertion in the environment of S_T results in a greater perceptual difference from the unepenthesized input. Therefore, it is less preferred cross-linguistically.

The perceptual stridency effect hypothesizes, based on the P-map for vowel epenthesis, that vowel insertion is less likely to occur after a voiceless strident than after a stop because an epenthetic vowel after a voiceless strident is perceptually more obtrusive than an epenthetic vowel after a stop. This dissertation has placed the hypothesized P-map for vowel insertion on firm empirical ground by implementing the auditory similarity judgment experiment. In this experiment, the degree of similarity to input for each set of epenthesized outputs was tested. In other words, participants in the experiment (native Korean speakers) were asked whether CCV-CvCV (bren-buren) pair is more similar to each other or CCV-CvCV (sken-suken) pair is more similar to each other. The results of this experiment are presented in Chapter 5.

The perceptual stridency effect determines the relative ranking relationships in the

following consonant clusters.

(11) Relative consonant cluster markedness determined by the perceptual stridency effect³⁷

Onset: /s/ + C < stop + C

Coda³⁸: /s/ + C < stop + C

4.2.3. Perceptual Voicing Effect

The principle of the P-map that an alternation that involves a less perceptible change is more likely to take place than an alternation that involves a more perceptible change generates another perceptual factor in PCCAM: the perceptual voicing effect.

The basic idea of the perceptual voicing effect is that vowel insertion is more likely to occur after a voiced obstruent than after a voiceless obstruent because an epenthetic vowel after a voiced obstruent is perceptually less obtrusive than an epenthetic vowel after a voiceless obstruent.

Previous research (Côté 2000; Fleischhacker 2001) and the auditory similarity judgment experiment in Chapter 5 all converge on the observation that an epenthetic vowel after a voiced consonant is perceptually more tolerable. According to Côté (2000:

³⁷ I determine the ranking relationship using the perceptual stridency effect and the perceptual voicing effect only when there are corresponding empirical results supporting the ranking relationships. Thus, I do not use the perceptual stridency effect and the perceptual voicing effect to establish the ranking relationship in sonorant + /s/ and sonorant + stop coda clusters because the auditory similarity judgment experiment in Chapter 5 did not test the perceptibility of an epenthetic vowel in these clusters. This is left for future research.

³⁸ Note that the role of the perceptual stridency effect in coda clusters will be revised in Chapter 5 based on the results of the perceptual similarity judgment experiment.

167), the tendency to insert a vowel after a voiced stop results from the fact that voiced segments share the low frequency energy associated with voicing of vowels; therefore, vowel epenthesis becomes less noticeable in the context of voiced segments.

In the same vein, while discussing the findings of Davidson (2001) that voiced clusters (e.g., /vn, vz, dv/) are more vulnerable to vowel insertion than voiceless clusters (/kt, pt, kp, tʃk, tf/], Fleischhacker (2001: 16) notes that vowel insertion is more likely to occur in a voiced environment because the voiced context is more similar to a vocalic element than a voiceless context, and consequently epenthesis is more tolerable in the former than in the latter.³⁹

The perceptual voicing effect determines the relative ranking relationships in the following consonant clusters.

(12) Relative consonant cluster markedness determined by the perceptual voicing effect

Onset

a. voiceless stop + sonorant < voiced stop + sonorant (pr < br)⁴⁰

Coda

a. voiceless stop + voiceless stop < voiced stop + voiced stop (pt < bd)

b. voiceless stop + voiceless strident < voiced stop + voiced strident (ps < bz)

³⁹ Note also that the auditory similarity judgment experiment in Chapter 5 empirically tested Côté (2000) and Fleischhacker's (2001) proposal (along with the perceptual stridency effect discussed in the previous section) that vowel insertion is more likely to occur after a voiced segment due to the vocalic properties shared by a voiced consonant and a vowel.

⁴⁰ Note that the role of the perceptual voicing effect in onset clusters will be revised in Chapter 5 based on the results of the perceptual similarity judgment experiment.

4.2.4. Principle of Perceptual Saliency (Côté 2000)

In § 4.2.2, I proposed that vowel insertion is less likely to occur after a voiceless strident because an epenthetic vowel after a voiceless strident is perceptually too obtrusive (the perceptual stridency effect). There is another aspect of the stridency effect: vowel insertion is less likely to occur after a strident because a strident can stand out very well in a consonant sequence without a vocalic segment nearby. This aspect of the stridency effect is labeled as the *Principle of Perceptual Saliency* and was originally proposed by Côté (2000).

Côté's (2000) dissertation investigates deletion and epenthesis processes in various languages (e.g., Hungarian, French, Lenakel, Sranan, Cairene Arabic, Iraqi Arabic, etc). In accounting for deletion and epenthesis processes, Côté (2000) develops the *Principle of Perceptual Saliency*, according to which every segment must be sufficiently salient.

According to the principle of perceptual saliency, consonant deletion and vowel insertion apply when a consonant lacks perceptual salience. In other words, consonant deletion and vowel insertion processes are employed when the cues that permit a listener to detect its presence are diminished. When it comes to the vowel insertion process, the principle of perceptual saliency predicts that a consonant whose acoustic cues are relatively weak attracts vowel insertion, since it favors closeness to a vocalic segment to maintain its perceptibility in a string. By the same reasoning, the consonant with relatively strong acoustic cues can perceptually stand out without a vocalic segment nearby.

Consider the case of Hungarian cluster simplification. Hungarian has an optional process of cluster simplification in internal position. All data come from Törkenczy & Siptár (1999, 2000 cited in Côté 2000).

(13) Hungarian cluster simplification (Côté 2000: 38)

	No simplification	Simplification	
a. <i>lambda</i>	[lɔmbdɔ]	[lɔmdɔ]	‘lambda’
b. <i>asztma</i>	[ɔstmɔ]	[ɔsmɔ]	‘asthma’
c. <i>röntgen</i>	[rɔndʒɛn]	[rɔŋʒɛn]	‘X-ray’
d. <i>dombtető</i>	[domptɛtø:]	[domtɛtø:]	‘hilltop’

Data in (13) show that the cluster simplification process applies to sequences of three consonants and always deletes a stop, not a strident or a sonorant. Based on the data in (13), Côté proposes a generalization for the pattern of cluster simplification in Hungarian: *stops, more than other consonants, want to be adjacent to a vowel, and preferably followed by a vowel*. The question is then, “What are the phonetic justifications for the patterns of vowel insertion and consonant deletion in Hungarian?” To answer this question, I will discuss relative perceptual saliency of stops, stridents, and sonorants in the following section.

Relative perceptual saliency of stops, stridents, and sonorants

It is a well-attested fact that a CV transition is perceptually more salient than a VC transition. Ohala (1990) found that listeners rely more on CV transition than on VC transition in deciding what they’ve heard in a VCV context. It is because “there are richer, more reliable cues in the CV transition than in VC transition” (p. 265). In a similar vein,

Kawasaki (1982) showed that CV transitions for consonants can be easily differentiated, whereas VC formant transitions do not show clear transitional differences.

It is important to note that a VC transition is especially unfavorable for the perception of stops. The internal cues of stops are weaker than other obstruents due to the complete blocking of the airflow during its closure duration. The airstream is completely blocked during the closure of stops, and thus the closure part of stops is silent or associated only with low-amplitude vocal fold vibration. Accordingly, stops have very weak internal cues during closure. Then how do we perceive stops if they lack internal cues?

The perceptibility of stops is crucially dependent on their release burst (Wright 1996: 5). However, not all stops take advantage of this release burst because a release burst occurs in prevocalic position, but its occurrence is variable in post-vocalic position. Accordingly, non-prevocalic stops do not reliably benefit from a release burst, and the absence or weakness of the release burst may reduce the salience of stops. In short, the absence or weakness of the burst may severely reduce the salience and perceptibility, and the disadvantage of VC cues against CV ones is amplified in the case of stops as opposed to other consonants.

By contrast, nasals, stridents, and liquids have relatively robust internal cues. For instance, a strident has relatively strong internal cues owing to its frication noise and hissing properties.⁴¹ In addition, the inherent duration of stridents is especially long compared with other consonants [*/s/* = 129 msec, */p/* = 89, */t/* = 77, */k/* = 69, */b/* = 90, */d/* = 83, */g/* = 67, */l/* = 66] (Umeda 1977: 851). The inherent duration is affected by many

⁴¹ Internal cues are produced during the closure part of the consonant (Côté 2000: 133).

factors such as the preceding and following segment and stress, but the long duration of /s/ is nonetheless prominent among the consonants.

Similarly, sonorant consonants such as liquids and nasals have relatively strong internal cues indicated by their formant structures. Accordingly, sonorant consonants do not require additional transitional cues to stand out in a string.

To summarize, consonants that lack the CV transition cues are supposed to be perceptually weaker than consonants in CV transition. Stops suffer more than other consonants from not appearing in prevocalic position due to the fact that stop release burst, on which the perceptibility of stops is crucially dependent, does not necessarily occur in VC context. Thus, the greater tendency of stops to delete, trigger epenthesis or block deletion stems from the weakness of their non-CV cues.

Incorporating Côté's proposal (i.e., the Principle of Perceptual Saliency) that explains the greater propensity of stops to get deleted and to promote vowel insertion into the system, PCCAM predicts that vowel insertion is more likely to occur after a stop than after a strident in L2 learners' production of English consonant clusters. By the same logic, PCCAM predicts that a stop is the main target of deletion in consonant cluster simplification.

It is important to note that the principle of perceptual saliency only applies to *coda clusters* because stops in English onset clusters are all flanked by a sonorant (i.e., no stop + stop onset clusters in English) whose properties are easily extended to that of a vowel. The obstruent-sonorant juncture is acoustically very similar to the obstruent-vowel juncture: both are characterized by the offset of aperiodic noise, and the onset of formant structure (Fleischhacker 2005: 72). Simply stated, a stop in onset clusters is not

the target of deletion because it occurs in CR (sonorant) transition.

The principle of perceptual saliency determines the relative ranking relationships in the following consonant clusters.

(14) Relative consonant cluster markedness determined by the principle of perceptual saliency

Coda

- a. sonorant + strident < sonorant + stop (ls < lp; lz < lb)
- b. stop + strident < strident + stop (ps < sp; bz < zd)
- c. stop + strident < stop + stop (ps < pt; bz < bd)

4.2.5. Maximal Perceptual Contrast

The maximal perceptual contrast parameter in PCCAM hypothesizes that language learners prefer a sequence where two segments in an alternating sequence make a maximal perceptual contrast. The phenomenon is evidenced in a child's preference for a less sonorous onset (Fikkert 1994; Gnanadesikan 2004).

Optimal onset of a syllable

While examining nine Dutch-learning children's speech, Fikkert (1994) observed that stops occurred first as an onset of syllable, followed by nasals and liquids (stops < nasal <

glide, fricative, liquid; Fikkert 1994: 64).⁴² Based on this empirical result, Fikkert (1994: 65) proposes that optimal core syllable in child phonology is represented as a stop + vowel sequence as in Figure 4.3.

It is important to note that Fikkert's proposal for the optimal core syllable in Dutch-learning child phonology is based on the sonority hierarchy. Stops are the least sonorous segment, while vowels are the most sonorous segment. Accordingly, the combination of a stop + vowel sequence makes a maximal contrast in terms of *sonority differences*.

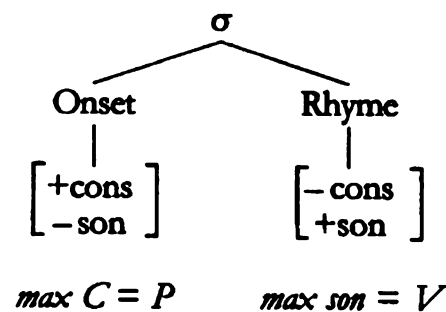


Figure 4.3. Optimal core syllable (Fikkert 1994: 65)⁴³

However, children's preference for stops as an onset of a syllable can also be explained in terms of perception. Voiced vowels are the loudest sounds, while obstruents (in particular voiceless one) involve long durations of silence. Given this, obstruents are preferred as an onset of a syllable because strings containing obstruents and voiced vowels in alternating sequences will result in maximal sequential contrast, which facilitates the perceptual saliency of each element in a string (Blevins 2004: 12).

⁴² Acquisition order among glides, fricatives, and liquids is variable in Fikkert (1994).

⁴³ cons = consonantal, son = sonorant, P = plosive, V = vowel ; "max C = P" means that maximal consonantal segments are plosives, and "max son = V" means that maximal sonorous segments are vowels.

PCCAM assumes that language learners prefer a sequence where two segments in an alternating sequence make a maximal perceptual contrast. Thus, among sC onset consonant clusters where C_1 is constant as /s/, /s/ + stop onset clusters are most optimal in terms of the maximal perceptual contrast parameter (a stop makes a maximal contrast with a following vowel), followed by /s/ + nasal and /s/ + liquid clusters.

The case of /lm/ sequence

Liquid + nasal consonant clusters demonstrate another example where the maximal perceptual contrast plays a role in determining consonant cluster markedness. Compare sonorant + voiceless stop and liquid + nasal coda clusters. We can see that a liquid + nasal coda cluster is worse than sonorant + voiceless stop in terms of the maximal perceptual contrast: all of the three segments (vowel + liquid + nasal) in a liquid + nasal cluster share the same voicing properties.

The proposal that a sequence of similar sounds is disfavored has found support in Kawasaki (1982). While investigating the spectral characteristics of some phoneme sequences, Kawasaki notes that, if two or more sound sequences produced considerably similar trajectories, (i.e., if they were acoustically and auditorily similar), they would be subject to confusion and thus to merger. Consequently, these sequences are cross-linguistically disfavored. Given the account of Kawasaki, we can assume that there can be situations “where a given segment sequence is salient enough but is disfavored due to its merger with another equally salient but acoustically similar sequence” (Ohala 1990: 327-328).

In the spirit of Kawasaki (1982) and Ohala (1990), I propose that the *vowel* +

liquid + nasal sequence (/lm/) demonstrates the kind of situation where each segment, a liquid and a nasal, is acoustically salient enough, but the merger within the sequence makes the cluster disfavored. By contrast, sonorant + voiceless stop clusters are less likely to be subject to this kind of merger because a sonorant and a voiceless stop make a good perceptual contrast with each other. Recall that obstruents (in particular a voiceless one) involve long durations of silence, while a sonorant is one of the loudest sounds. Thus, the strings containing a sonorant and a voiceless obstruent in alternating sequences may result in good perceptual contrast, which facilitates the perceptual saliency of each element in a string.

In addition, it is possible that an epenthetic vowel between a liquid and a nasal does not perceptually stand out, and thus the liquid + nasal sequence induces more vowel insertion (as is the case with perceptual voicing effect). This is certainly plausible, but the auditory similarity judgment experiment in Chapter 5 did not test the perceptibility of an epenthetic vowel in liquid + nasal clusters. Thus, I will not use the perceptibility of an epenthetic vowel in this sequence as a phonetic factor defining the markedness of an /lm/ sequence.

In sum, based on the maximal perceptual contrast parameter that language learners prefer a sequence where two segments in an alternating sequence make a maximal perceptual contrast, PCCAM predicts that L2 learners will insert a vowel to provide additional perceptual cues to the /lm/ sequence or delete one of the segments.

The relative consonant cluster markedness relationships determined by the maximal perceptual contrast parameter are as follows.

(15) Relative consonant cluster markedness determined by the maximal perceptual contrast parameter

Onset: /s/ + stop < /s/ + nasal < /s/ + liquid

Coda: sonorant + voiceless stop < liquid + nasal

The perceptual factors defining consonant cluster markedness in PCCAM are summarized in Table 4.7.

Table 4.7. Summary of perceptual factors in PCCAM

Perceptual stridency effect	
Onset	a. /s/ + C < stop + C (st < pr)
Coda	a. /s/ + C < stop + C (st < pt)
Perceptual voicing effect	
Onset	a. voiceless stop + sonorant < voiced stop + sonorant (pr < br)
Coda	a. voiceless stop + voiceless stop < voiced stop + voiced stop (pt < bd)
	b. voiceless stop + voiceless strident < voiced stop + voiced strident (ps < bz)
Principle of perceptual saliency	
Coda	a. sonorant + strident < sonorant + stop (ls < lp; lz < lb)
	b. stop + strident < strident + stop (ps < sp; bz < zd)
	c. stop + strident < stop + stop (ps < pt; bz < bd)
Maximal perceptual contrast	
Onset	a. /s/ + stop < /s/ + nasal < /s/ + liquid (sp < sm < sl)
Coda	a. sonorant + voiceless stop < liquid + nasal (lt < lm)

4.3. Consonant Cluster Markedness under PCCAM

In § 4.1 and § 4.2, I discussed the theoretical frameworks that constitute the *Phonetically-Based Consonant Cluster Acquisition Model* (PCCAM), Articulatory Phonology and the P-map. More importantly, I identified, under the principles of Articulatory Phonology and the P-map, articulatory and perceptual factors defining consonant cluster markedness.

The eight phonetic factors in Table 4.5 and 4.7 will be used to establish the

pairwise markedness ranking relationships among English consonant clusters (e.g., sp, st, sk < sm and sm < sl). In particular, § 4.4 and § 4.5 are devoted to *empirically* justifying the consonant cluster markedness hierarchy in PCCAM. This is done by referring to published first and second language acquisition data. It will be shown that a fair amount of published data supports the consonant cluster markedness hierarchy in PCCAM.

I will first discuss how the relative consonant cluster markedness is computed in PCCAM.

4.3.1. Three-way Computation of Consonant Cluster Markedness

In the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM), the relative difficulty of producing consonant clusters is calculated through a three-way computation of consonant cluster markedness: a) markedness of C_1 (the first consonant in a cluster), b) markedness of C_2 (the second consonant in a cluster), and c) markedness of the gestural coordination of C_1 and C_2 . Also note that PCCAM defines the markedness of C_1 and C_2 on both articulatory and perceptual grounds. The consonant cluster markedness hierarchy generated by PCCAM is presented in Table 4.8.

Table 4.8 shows that when it comes to onset clusters, /s/ + stop is the least marked while voiced stop + liquid is the most marked. Regarding obstruent + obstruent coda clusters, PCCAM predicts that voiceless stop + voiceless strident is least marked, while voiced stop + voiced stop is most marked.

Table 4.8. Consonant cluster markedness hierarchy in PCCAM ⁴⁴

less marked < more marked; // = no markedness relation	
Onset Clusters	sp, st, sk < sm, sn < sl < pl, pr, tr, kl, kr < bl, br, dr, gl, gr
Coda Clusters: Sonorant + C	ls, ns, < lz, nz // lp, lt, lk, np, nt, mp < lb, ld, nd // lm
Coda Clusters: Obstruent + Obstruent	ps, ts, ks < sp, st, sk < pt, kt // bz, gz < zd < bd, gd

To fully understand the markedness hierarchy presented in Table 4.8, four issues will be discussed: a) consonant markedness hierarchy in PCCAM: language specific or universal?, b) rational for the exclusion of some types of consonant clusters in PCCAM, c) two types of coda consonant clusters, and d) the case where no ranking relationship is established: notation //.

Consonant markedness hierarchy in PCCAM: language specific or universal?

One may wonder whether the consonant cluster markedness hierarchy generated by PCCAM is L2 learners' native language-specific. The answer is *Yes* and *No*.

The consonant cluster markedness hierarchy in Table 4.8 is *universal* in the sense that the phonetic factors identified and defined in this chapter (e.g., the articulatory stridency effect, the articulatory voicing effect) are based on the universal aspects of speech production and speech perception shared by human beings and are thus applicable to any first and second language environment.

⁴⁴ The consonant clusters in Table 4.8 are grouped as follows: voiceless strident + voiceless stop (sp, st, sk); voiceless strident + nasal (sm, sn); voiceless stop + liquid (pl, pr, tr, kl, kr); voiced stop + liquid (bl, br, dr, gl, gr); sonorant + voiceless fricative (ls, ns); sonorant + voiced fricative (lz, nz); sonorant + voiceless stop (lp, lt, lk, np, nt, mp); sonorant + voiced stop (lb, ld, nd); liquid + nasal (lm); voiceless stop + voiceless strident (ps, ts, ks); voiceless strident + voiceless stop (sp, st, sk); voiceless stop + voiceless stop (pt, kt); voiced stop + voiced strident (bz, gz)

But at the same time, PCCAM is specific to the L2 learners' native language in the sense that certain clusters can be more marked than others simply due to the fact that a native sound inventory lacks a specific segment in a cluster. Note that the consonant cluster markedness hierarchy in Table 4.8 is generated based on *Korean* speakers' production of *English* consonant clusters in mind.

To illustrate how the learners' native language plays a role in the consonant cluster markedness hierarchy in PCCAM, consider Korean speakers' and Mandarin Chinese speakers' production of English coda clusters. Korean allows liquids, nasals, fricatives, and stops in the coda position (as a singleton coda), while Mandarin Chinese only allows nasals in the coda. Given this difference in the coda inventories of Korean and Mandarin Chinese, it is natural to predict that Mandarin Chinese learners of English may demonstrate more difficulty with liquid or obstruent coda clusters, while Korean speakers may demonstrate no difference between sonorant + obstruent and obstruent + obstruent sequences when it comes to the markedness of C_1 (the first consonant in a cluster). Thus, to apply the consonant markedness hierarchy in Table 4.8 to Mandarin Chinese learners of English, some of the rankings may need to be rearranged (e.g., the ranking relationship between liquid + obstruent and nasal + obstruent).

However, it should be noted that the rearrangement of markedness hierarchy will be limited to a small number of cases, and most of the markedness hierarchy will hold for L2 learners regardless of their native language. Consider the ranking relationships among obstruent + obstruent clusters: $ps, ts, ks < sp, st, sk < pt, kt // bz, gz < zd < bd, gd$. PCCAM predicts that these ranking relationships will hold for any L2 learners regardless of their native language sound inventory (as long as their native language does not allow

complex margins), because the phonetic factors defining each ranking relationship are assumed to be universal.

It will be shown throughout this chapter that even though the consonant cluster markedness hierarchy in Table 4.8 is constructed with *Korean* speakers' production of *English* consonant clusters in mind, published L1 and L2 consonant cluster acquisition data (e.g., Vietnamese, Brazilian Portuguese, Mandarin Chinese, etc.) actually support most of the ranking relationships in PCCAM. This kind of evidence can be taken as demonstrating the universal aspects of the consonant cluster markedness hierarchy in PCCAM.

Rational for the exclusion of some types of consonant clusters in PCCAM

If any of the consonants in a cluster is not allowed in Korean, the cluster is excluded from the consonant cluster markedness hierarchy. Thus, no English consonant clusters with /f, v, ʃ, θ, ð/ (these segments are not allowed in Korean) are discussed in terms of their relative markedness with other clusters. This exclusion is based on the rational that the errors resulting from the fact that Korean does not have the segments in its sound inventory (/f/ → [p], /v/ → [b], /θ/ → [s], etc) should not be coded as an error resulting from consonant cluster production because those errors can also occur as a singleton onset or coda.

Moreover, including these clusters in the production experiment brings about greater difficulty analyzing and interpreting the data. Consider the case that a Korean speaker inserts an epenthetic vowel in a cluster of /ft/ and also substitutes [p] for /f/ (/ft/ → [put]). When we interpret this data in terms of the consonantal environment in which

vowel insertion occurs, how can we determine which consonant motivates vowel insertion? Is it /f/ or /p/ that brings about vowel insertion? Thus, it was decided to exclude those segments which cause manner feature changes in Korean speakers' production. In this way, it is possible to have more manageable empirical data in terms of data coding and interpretation.

Two types of coda consonant clusters

Note in Table 4.8 that coda clusters are divided into two groups: sonorant + C clusters and obstruent + obstruent clusters. This is done for the systematic comparison between clusters.

Consider the markedness relationship between liquid + voiceless strident and voiceless stop + voiceless strident (ls vs. ps). Note, first of all, C_2 in these clusters is fixed with /s/. Also recall that the consonant cluster markedness in PCCAM is determined by computation of a) markedness of C_1 , b) markedness of C_2 , and c) markedness of gestural coordination of C_1 and C_2 . Since C_2 is fixed with /s/, the evaluation of markedness of C_2 is not relevant. When it comes to the markedness of C_1 and the markedness of gestural coordination of C_1 and C_2 , no phonetic factor in PCCAM determines the relative unmarkedness of C_1 liquid over C_1 obstruent (or vice versa) or the relative unmarkedness of liquid + strident sequences over stop + strident sequences (or vice versa). For this reason, it was decided to divide coda clusters into two groups: sonorant + C vs. obstruent + obstruent.

The case where no ranking relationship is established: notation //

The notation // in Table 4.8 indicates that no markedness relationship can be established between the clusters. There are three of these cases: a) sonorant + voiced strident (/lz, nz/) vs. sonorant + voiceless stop (/lp, lt, lk, np, nt, mp), b) voiceless stop + voiceless stop (/pt, kt/) vs. voiced stop + voiced strident (/bz, gz/), and c) sonorant + voiced stop (/lb, ld, nd) vs. liquid + nasal (/lm/).

The reason why the markedness relationship cannot be established in the first two cases (a and b) is because segments in these clusters differ in two different dimensions. Compare the relationship between sonorant + voiced strident (/lz, nz/) and sonorant + voiceless stop cluster (/lp, lt, lk, np, nt, mp/). We can see that C₁ is fixed with a sonorant in these two clusters and thus the markedness of C₁ is not relevant for evaluation. Looking at the markedness of C₂, we can see that a voiced strident is different from a voiceless stop in two dimensions: voicing and stridency (i.e., /z/ [+voice][+strident] vs. /p/: [-voice][-strident]).

As was shown in § 4.1 and § 4.2, PCCAM crucially refers to the voicing and stridency feature of a segment in a cluster when determining the relative markedness of a cluster. Therefore, when two segments in comparison differ in both voicing and stridency features, a markedness relationship cannot be established.

Returning to the markedness relationship between sonorant + voiced strident (/lz, nz/) and sonorant + voiceless stop cluster (/lp, lt, lk, np, nt, mp/), we can see that the C₂ voiced strident in sonorant + voiced strident is *less* marked than C₂ voiceless stop in sonorant + voiceless stop, because of the stridency feature in a voiced strident. PCCAM predicts, based on the principle of perceptual saliency (Côté 2000), that vowel insertion is less likely to occur after a strident.

In contrast, a voiced strident is *more* marked than a voiceless stop because of the [+ voice] feature in a voiced strident. PCCAM predicts, based on the articulatory voicing effect, that vowel insertion is more likely to occur after a voiced obstruent due to the articulatory difficulty of maintaining voicing in an obstruent. Thus, no markedness relationship is established between sonorant + voiced strident (/lz, nz/) and sonorant + voiceless stop (/lp, lt, lk, np, nt, mp/) because a tie occurs in the computation of the markedness ranking.

The reason that there is no ranking relationship between sonorant + voiced stop and liquid + nasal is rather different than the previous two cases: lack of empirical evidence. It seems quite plausible that an epenthetic vowel between a liquid and a nasal is perceptually less obtrusive than an epenthetic vowel in other sonorant + obstruent clusters, but we do not have empirical evidence to support this proposal. Thus, I will leave the ranking relationship between sonorant + voiced stop and liquid + nasal undetermined. This is certainly one area that merits future research.

We have discussed four issues related to the consonant cluster markedness hierarchy in PCCAM (Table 4.8). The summary of four issues related with the consonant cluster markedness hierarchy in PCCAM is as follows.

- (1) Is the consonant markedness hierarchy in PCCAM language specific or universal?: Given the universality of phonetic factors defining consonant cluster markedness in PCCAM, it is plausible that the markedness hierarchy will hold for any L2 learners learning English consonant clusters. Yet, the consonant cluster markedness may need some revisions based on L2 learners' native language

sound inventory as was demonstrated with Mandarin Chinese learners of English.

- (2) Rational for the exclusion of some types of consonant clusters in PCCAM: If any of the consonants in a cluster is not allowed in Korean, the cluster is excluded from the consonant cluster markedness hierarchy. This is done to make L2 learners' production data more manageable in terms of coding and interpretation.
- (3) Two types of coda consonant clusters: Coda clusters are divided into sonorant + C and obstruent + obstruent clusters for systematic evaluation of consonant cluster markedness.
- (4) Cases where no ranking relationships are established (notation //): no markedness relationship is established between two clusters if a tie occurs in the computation of markedness ranking.

Now that the proposed consonant cluster markedness hierarchy in PCCAM has been established (Table 4.8), the following sections empirically justify each ranking relationship by referring to articulatory and perceptual factors associated with each cluster.

4.4. Relative Markedness of Onset Consonant Clusters

This section discusses articulatory and perceptual factors that define the relative markedness of English onset consonant clusters. Since all the articulatory and perceptual factors have already been defined in § 4.1 and § 4.2, only a brief summary of the phonetic factors for each markedness relationship is provided.

4.4.1. sp, sk, st < sm, sn and sm, sn < sl

This section discusses the relative markedness relationship among sC onset clusters: /s/ + stop, /s/ + nasal, and /s/ + lateral. Because C₁ is fixed with /s/ in these clusters, the relative markedness is determined based on the markedness of C₂ and the markedness of gestural coordination of C₁ and C₂.

Two phonetic factors define the relative markedness relationship among sC onset clusters.

a) Markedness of C₂

- i) Maximal perceptual contrast: a stop is more readily perceived as an onset of a syllable than a nasal. In turn, a nasal is more readily perceived than a lateral as an onset of a syllable. The maximal perceptual contrast parameter determines the relative ranking relationship among sC clusters as follows: /s/ + stop (less marked) < /s/ + nasal < /s/ + lateral.

b) Markedness of gestural coordination of C₁ and C₂

- i) The articulatory stridency effect: vowel insertion is less likely to occur after a voiceless strident because the aerodynamic conditions of /s/ remain conducive to a continued turbulence of a strident rather than vocal fold vibration. The articulatory stridency effect is assumed to be strongest in a sequence of /s/ + stop (and thus vowel insertion is least likely to occur in this sequence). The

articulatory stridency effect parameter generates the following markedness ranking relationship among sC clusters: /s/ + stop (less marked) < /s/ + nasal, /s/ + liquid.

The computation of the two phonetic factors listed above generates the final markedness ranking relationship as follows: sp, st, sk < sm, sn and sm, sn < sl. In the following section, I present empirical evidence from both first and second language acquisition which supports the markedness ranking relationship among sC onset clusters.

4.4.1.1. Empirical Evidence for sp, st, sk < sm, sn and sm, sn < sl

(1) Child language data

The first piece of evidence for the ranking relation between /s/ + stop and /s/ + nasal is found in child language data. Kirk & Demuth (2003) examined the acquisition of English onset and coda clusters in a cross-sectional experimental study (n = 9, aged 2). A clear difference in the number of children who produced /s/ + stop and /s/ + nasal with greater than 70% accuracy was observed in the study. Out of nine children, three children produced /s/ + stop with greater than 70% accuracy, while none produced /s/ + nasal correctly (Kirk & Demuth 2003: 446).

The markedness ranking relation between /s/ + nasal and /s/ + liquid finds support in Farris (2005). In a study examining the inventory of sC clusters of 110 children (3:0-7:4), Farris (2005) observed that of the 32 children who produced sC clusters, over one-

third evidenced a gapped inventory: children who acquired /sw/ and /sn/ did not acquire /sl/. Farris raises the reason of why /sn/ is acquired earlier than /sl/ cluster.

This question cannot be answered based on the notion of sonority. Under the sonority-based account of consonant cluster markedness, the smaller the sonority difference between the two segments in a cluster, the more marked it is as a cluster. Accordingly, /sn/ is more marked than /sl/ because the sonority difference between /s/ and /n/ is 1, while the sonority difference between /s/ and /l/ is 2 (refer to the sonority index in § 3.2.1).

However, the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM) makes different predictions on the relative markedness between /sn/ and /sl/. As was made clear in the previous section, /s/ + liquid is more marked than /s/ + stop and /s/ + nasal clusters both in terms of the perceptual stridency effect and the maximal perceptual contrast parameter. Thus, the gapped inventory observed by Farris (2005) is actually what is predicted by PCCAM.

That /sl/ is more marked than /s/ + stop and /s/ + nasal clusters is also observed in Smit et al. (1990) and Templin (1957). Table 4.9 presents 75% levels of acquisition for word-initial clusters in Smit et al. (1990) and Templin (1957).

Table 4.9. 75% levels of acquisition for word-initial sC clusters

Age of acquisition reported in the Iowa-Nebraska Project (Smit et al. 1990: 55)			Templin (1957 : 51)
Clusters	Females	Males	
/sp/	4:6	5:0	4:0
/st/	4:6	5:0	4:0
/sk/	4:6	6:0	4:0
-----	-----	-----	-----
/sm/	5:6	7:0	4:0
/sn/	5:6	5:0	4:0
-----	-----	-----	-----
/sl/	6:0	7:0	7:0

Note in Table 4.9 that /sl/ is acquired last among the sC clusters tested in Smit et al. (1990) and Templin (1957). The acquisition order from these two studies is reliable because a large number of children were tested. Smit et al. collected speech samples from 1049 children (2:0-9:0), while Templin tested 480 children (3:0-8:0).⁴⁵

However, in contrast to the developmental order of sC onset clusters found in Farris (2005), Kirk & Demuth (2003), Smit et al. (1990), and Templin (1957), there is one study showing opposite patterns of the acquisition order. In Smith's diary study (1973) of normal phonological development, Amahl showed the following progress regarding the acquisition of clusters: Cl/rV (where C is a plosive) appear < sl appear < sm/sn appear < sp/sk and all consonants plus /w/ appear < st < all consonants plus /j/ appear (Smith 1973: 169). That is, Amahl's cluster acquisition order demonstrates the exact opposite direction to the prediction of PCCAM. /sl/ appears first and then /sm, sn/ follow, and /sp, st, sk/ appear last.

Given that Smith (1973) collected data from only one child, this exceptional pattern is not significant enough to question the general pattern of acquisition order in other previous studies. More importantly, there is an issue of whether we can really take the data from Smith (1973) as the order of consonant cluster acquisition. That is, even though Smith herself referred to the data as "all stages of cluster acquisition in A (Smith 1973: 168), she frequently uses the word "appear" in explaining the data. It is not clear whether or not this relies on the first appearance of a given cluster as evidence for the acquisition of the cluster. In other words, it is unclear exactly how frequent appearance of

⁴⁵ The acquisition order of sC clusters by *female* participants in Smit et al. (1990) exactly reflects the predictions of PCCAM on the relative markedness of sC clusters (/s/ + stop < /s/ + nasal < /s/ + liquid). Yet, the acquisition order by *male* participants in Smit et al. (1990) is rather mixed. Note that /sm/ is acquired at the same time as /sl/. It seems that this kind of conflicting data can only be resolved by the future study specifically looking at the developmental order of sC clusters.

a cluster is taken as evidence for the acquisition of the cluster in that study.

The discrepancy found in Smith (1973) and other previous studies regarding the developmental order of sC onset clusters may be derived from the methodological differences in these studies. Stoel-Gammon & Dunn (1985: 33) note that “longitudinal studies revealed substantial differences in rate and order of acquisition of clusters in the types of errors. These differences make it difficult to identify developmental stages between the first attempts at production and final mastery.”

(2) Second language data

The relative markedness of /s/+ liquid over /s/ + stop is also attested in second language acquisition of consonant clusters.

Major (1996) tested four adult native speakers of Brazilian Portuguese on production of English consonant clusters. The onset clusters tested in the study were /sl, ʃr, sp, st, sk, pr, br, pl, bl, tr, dr, kr, gr, kl, gl/. VARBRUL analyses were performed to determine probabilities of errors of each cluster. The results show that in onset clusters, fricative + liquid contributes most strongly to error (0.70), and voiceless strident + voiceless stop is least likely to contribute to error (0.37).⁴⁶ A similar pattern is found in Kim (2000). Kim (2000) tested Korean speakers’ production of English consonant clusters and found that /s/ + stop onset clusters incur no vowel epenthesis errors, while /s/ + sonorant clusters (/s/ + nasal and /s/ + liquid) involve some epenthesis errors.

All in all, the relative markedness relationship among sC clusters seems to be

⁴⁶ Note that Portuguese sound inventory includes /ʃ/.

empirically well supported both in L1 and L2 acquisition data.

4.4.2. sl < pl, pr, tr, kl, kr

This section defines the relative markedness between /s/ + liquid and voiceless stop + liquid. Because C₂ is fixed with a liquid in these clusters, the relative markedness is determined based on the markedness of C₁ and the markedness of the gestural coordination of C₁ and C₂.

Three phonetic factors are employed to define the relative markedness between /s/ + liquid and voiceless stop + liquid.

a) Markedness of C₁

- i) Articulatory stridency effect: vowel insertion is less likely to occur after a voiceless strident because the aerodynamic conditions of /s/ remain conducive to continued turbulence of a strident rather than vocal fold vibration (sl < pl, pr, tr, kl, kr).
- ii) Perceptual stridency effect: vowel insertion is less likely to occur after a voiceless strident because an epenthetic vowel after a voiceless strident is perceptually more obtrusive than an epenthetic vowel after a stop (sl < pl, pr, tr, kl, kr).

b) Markedness of gestural coordination of C₁ and C₂

- i) Degree of gestural overlap: /s/ + liquid is less marked than voiceless stop +

liquid because a fricative is less overlapped by a following consonant than is a stop (sl < pl, pr, tr, kl, kr).

The computation of these three phonetic factors generate the relative markedness in sl < pl, pr, tr, kl, kr.

4.4.2.1. Empirical Evidence for: sl < pl, pr, tr, kl, kr

(1) Child language data

An epenthetic vowel occurs as a strategy to deal with consonant clusters in child phonology along with cluster reduction.⁴⁷ It has been reported that vowel epenthesis occurs in the speech of 2 to 3 year olds (Bortolini & Leonard 1991; Dyson & Paden 1983; McLeod, J., and Reed 2002) as well as in older children (Olmsted 1971).

An interesting pattern of vowel insertion is observed in the production of children learning European Portuguese. Freitas & Rodrigues (2003: 75) reported that Portuguese children insert an epenthetic vowel to break down obstruent + liquid clusters, while /s/ + stop clusters are not separated with an epenthetic vowel (e.g., bruxa /'bruʃɐ/ → [bi'ruʃɐ] 'witch' (Luís: 2: 11.1), estrela /ʃtreɫɐ/ → [ʃti'rew] 'star' (Luis: 2: 5.7)).⁴⁸ The asymmetry of vowel insertion in /s/ + stop and stop + liquid clusters exactly corresponds with the

⁴⁷ Vowel insertion in child phonology is not as common as that in second language phonology. Refer to Young-Scholten & Archibald (2000) for a possible account for this pattern.

⁴⁸ Due to the fact that vowel insertion is not as common as cluster reduction in child phonology, and that there are not many studies specifically examining the site of vowel insertion in child phonology, it is difficult to generalize the vowel insertion pattern found in Freitas & Rodrigues (2003).

predictions of PCCAM on the patterns of vowel insertion: vowel insertion is more likely to occur in stop + liquid than /s/ + stop clusters.

Even though the relative markedness between /s/ + liquid and voiceless stop + liquid clusters is not specifically investigated, some child language studies report the unmarkedness of sC clusters with respect to stop + liquid clusters (Barlow 1997; Freitas & Rodrigues 2003; Gierut 1999). For instance, Barlow (1997: 136) observed that Subject 2 in her study produced sC clusters target-appropriately, while all other clusters were reduced to a singleton. Similarly, Freitas & Rodrigues (2003: 75) observed that children learning European Portuguese acquired sC clusters earlier than obstruent + liquid clusters, despite the fact that Portuguese has both obstruent + liquid and /s/ + obstruent clusters.

Gierut (1999) addresses acquisition order of sC clusters with a treatment study of children exhibiting functional phonological delays (n = 5, aged 3;2-7;8). The assumption in the clinical treatment study is that the treatment with more marked clusters facilitates the acquisition of less marked clusters. Thus, if /s/ + stop clusters are more marked than stop + liquid clusters (as is predicted by the sonority sequencing principle), then treatment with /s/ + stop clusters is expected to facilitate the acquisition of other onset clusters (e.g., stop + liquid).

The results of this study reveal that children who learned /s/ + stop clusters demonstrate patterns of within-class generalization. For instance, Subject 7 in her study was treated with /sp/, and the child acquired only sC sequences such as /sm, sn, sp, st/, and was unable to produce stop + liquid clusters, which is indicative of a within-class generalization (Gierut 1999: 718). Note that the within-class generalization is characteristic of learning patterns associated with treatment on unmarked aspects of

sound systems. In addition, Subject 9, who was also treated with /sp/, exhibited overgeneralization of /s/ in onset position (Gierut 1999: 720). Based on these results, Gierut concludes that /s/ + stop clusters are unmarked among onset clusters.

It should be noted that other empirical data show a slight difference in the relative markedness relationship between /sl/ and stop + liquid clusters than the predictions of PCCAM.

Table 4.10. 75% levels of acquisition for word-initial clusters (Smit et al. 1990)

acquired earlier < acquired later	
tw, kw (3:6) < pl, bl, kl (4:0)* < sp, st, sk, sw, gl, fl, kr , skw (4:6) < sm, sn (5:6) < sl, pr, br, tr, dr, gr, fr, spl (6:0) < θr (7:0) < spr, str, skr	

* Exceptional cases to the predictions of PCCAM are marked in bold.

The data in Table 4.10 illustrate the developmental order of onset clusters found in Smit et al. (1990). We can see that some of the sC clusters (sp, st, sk, sw, sm, sn) are acquired earlier than /pr, br, tr, dr, gr/ clusters, as PCCAM predicts. However, note that /pl, bl, kl/ are acquired earlier than sC onset clusters, which contradicts the prediction of PCCAM.

It seems rather difficult to establish the ranking relationship between /s/ + liquid and stop + liquid on the published child language data, particularly in the study investigating the developmental order of onset clusters (e.g., Smit et al. 1990). One possibility that has been entertained regarding the acquisition of sC onset clusters is that the developmental order of sC onset clusters with respect to other onset clusters may be influenced by the articulatory difficulty typical to child language acquisition: /s/ production is known to be problematic for child language learners (Smit 1993). Thus, for some child language learners, /sl/ may be acquired later than stop + liquid clusters due to the articulatory difficulty of producing /s/.

However, it should be noticed in the child language data when comparing sC clusters with other onset clusters that despite the proven difficulty of producing /s/ in child phonology, some children nevertheless produce sC clusters earlier than other onset clusters (e.g., Barlow 1997; Fikkert 1994; Freitas & Rodrigues 2003; Gierut 1999). It seems that when it comes to the child acquisition of sC vs. stop + liquid clusters, there is tension between keeping the perceptually salient segment /s/ and the articulatory difficulty of producing /s/. This may explain why some children acquire sC clusters earlier than stop + liquid clusters, while the reverse order is found for some other children.

(2) Second language data

On the surface, Major's (1996) study seems to present counter-evidence to the relative markedness between /s/ + liquid and stop + liquid clusters. It was found in the study that fricative + liquid clusters contribute more strongly to error (0.70) than stop + liquid (0.44) in four adult native speakers of Brazilian Portuguese in their production of English consonant clusters. However, this pattern is not problematic at all when we look at the phonotactics of their native language. Brazilian Portuguese allows /pr, br, tr, dr, kr, gr/ clusters. It makes sense then that the learners of this language are better at producing English stop + liquid clusters than fricative + liquid clusters simply because their native language has stop + liquid clusters.

Data from Kim (2000) partially support the relative markedness between /s/ + liquid and voiceless stop + liquid. For the less advanced Korean learners of English group, vowel insertion is more likely to occur in voiceless stop + liquid clusters (mean = .0833)

than in /s/ + sonorant clusters (mean = .0625) (as is predicted by PCCAM). However, for the more advanced Korean learners of English group in Kim's study, the direction reverses; more vowel insertion occurs in /s/ + sonorant clusters (mean = .0625) than voiceless stop + liquid clusters (mean = .0417) (Kim 2000: 138).

Admittedly, the relative markedness between /s/ + liquid and voiceless stop + liquid does not seem to have a strong empirical ground yet. This may be partly because few studies have aimed at establishing the markedness relationship between these two specific clusters. Another possible reason is that the relative markedness between these two clusters is too subtle to make a clear contrast in terms of production difficulty. This idea seems to find support in Kim (2000). In Kim's study, less advanced learners demonstrated the ranking relationship such as /s/ + sonorant (less marked) < stop + liquid, while more advanced learners demonstrated the opposite ranking relationship (stop + liquid < /s/ + sonorant). The production experiment (Korean speakers' production of English consonant clusters), which will be presented in Chapter 6, provides a good testing ground to resolve this kind of conflicting data.

4.4.3. pl, pr, tr, kl, kr < br, bl, dr, gl, gr

This section discusses the relative markedness between voiceless stop + liquid and voiced stop + liquid. The relative markedness between these two clusters is determined solely based on the markedness of C_1 . This is because C_2 is fixed with a liquid in both of the clusters, and also because the markedness of the gestural coordination of C_1 and C_2 cannot be established (i.e., both of the clusters involve a stop + liquid sequence).

a) Markedness of C₁

- i) Articulatory voicing effect: vowel insertion is more likely to occur after a voiced obstruent due to the articulatory difficulty of maintaining voicing in an obstruent.
- ii) Perceptual voicing effect: vowel insertion is more likely to occur after a voiced obstruent because an epenthetic vowel after a voiced obstruent is perceptually less obtrusive than an epenthetic vowel after a voiceless obstruent.⁴⁹

The articulatory and perceptual voicing effects determine that voiceless stop + liquid cluster is less marked than voiced stop + liquid clusters.

4.4.3.1. Empirical Evidence for pl, pr, tr, kl, kr < br, bl, dr, gl, gr

It has been found in many previous studies that vowel insertion is more likely to occur after a voiced obstruent than after a voiceless obstruent: Broselow & Finer (1991), No (1997), Kim (2000), and Davidson (2001, 2003). Since most of these studies have already been reviewed in the previous chapters, I will only introduce results from No (1997) in this section.

No (1997) tested 40 Korean male learners of English (20 elementary school children at the age of 9:00-11:00 and 20 college students) on their production of onset consonant clusters and single codas. No (1997) was particularly interested in the question

⁴⁹ Note that the role of the perceptual voicing effect in defining the relative markedness between voiceless stop + liquid and voiced stop + liquid will be revised in Chapter 5 based on the results of the perceptual similarity judgment experiment.

of whether voicing plays a role in the frequency of vowel insertion. The results reveal a significant voicing effect on the frequency of vowel insertion in codas: $F(1, 38) = 12.36$, $p < .01$. Regarding the role of voicing in the patterns of vowel insertion, No (1997) conjectures that “an attempt to make sounds voiced might result in vowel insertions” (p. 29), which seems to correspond with the notion of the articulatory voicing effect in the present study.

4.4.4. Summary of the Relative Markedness of Onset Consonant Cluster

Table 4.11 summarizes articulatory and perceptual factors defining the onset consonant cluster markedness as in $sp, sk, st < sm, sn < sl < pl, pr, tr, kl, kr < br, bl, dr, gl, gr$.

Table 4.11 also lists the previous studies supporting each relative markedness relationship in onset consonant clusters. If the empirical data listed in the third column do not exactly correspond to the ranking relationship in the first column, additional information is also given in Table 4.11. For instance, Kim’s (2000) study was presented as empirical evidence supporting the markedness ranking relationship between $/s/ + \text{stop}$ and $/s/ + \text{nasal}$ ($sp, st, sk < sm$) in the first line of Table 4.11. However, Kim’s study compared $/s/ + \text{stop}$ with $/s/ + \text{sonorant}$, which includes both $/s/ + \text{nasal}$ and $/s/ + \text{liquid}$. Thus, Kim’s study is relevant for the markedness ranking relationship between $/s/ + \text{stop}$ and $/s/ + \text{nasal}$, yet the grouping of the clusters does not exactly match with the one in the present study. This type of information is given in Table 4.11.

Table 4.11. Relative markedness of onset consonant clusters

Relative markedness relationship	Phonetic factors	Empirical data
sp, st, sk < sm, sn	a. maximal perceptual contrast	Kim (2000): /s/ + stop vs. /s/ + sonorant
	b. articulatory stridency effect	Kirk & Demuth (2003)
		Major (1996): /s/ + stop vs. strident + liquid Smit et al. (1990)
sm, sn < sl	a. maximal perceptual contrast	Farris (2005)
		Smit et al. (1990) only for females
		Templin (1957)
sl < pl, pr, tr, kl, kr	a. articulatory stridency effect	Freitas & Rodrigues (2003): /s/ + obstruent vs. obstruent + liquid
	b. perceptual stridency effect	Gierut (1999): /s/ + stop vs. stop + liquid
	c. degree of gestural overlap	Kim (2000): partially supported (only for less advanced learners)
		Smit et al. (1990): not supported
pl, pr, tr, kl, kr < bl, br, dr, gl, gr	a. articulatory voicing effect	Broselow & Finer (1991)
	b. perceptual voicing effect	Davidson (2001, 2003): voiceless obstruent + voiceless obstruent vs. voiced obstruent + voiced obstruent
		Kim (2000)
		No (1997): for a single coda

4.5. Relative Markedness of Coda Consonant Clusters

This section discusses articulatory and perceptual factors that define relative markedness of English coda consonant clusters. Since all the articulatory and perceptual factors have already been defined in § 4.1 and § 4.2, only a brief summary of the phonetic factors for each markedness relationship is provided.

Note also that liquid + obstruent and nasal + obstruent are combined into one category, sonorant + obstruent. It is assumed in the present study that Korean speakers demonstrate no difference between liquid + obstruent and nasal + obstruent sequences when it comes to the markedness of C₁ (the first consonant in a cluster) because Korean allows both liquids and nasals in the coda position (as a singleton coda). This is the reason that liquid + obstruent and nasal + obstruent are combined into sonorant + obstruent clusters in PCCAM.

4.5.1. ls, ns < lz, nz

This section defines the relative markedness between sonorant + voiceless strident (/ls, ns) and sonorant + voiced strident (/lz, nz/). The relative markedness between these two types of clusters is determined solely based on the markedness of C₂ because a) C₁ is fixed with a sonorant and b) the relative markedness of gestural coordination of C₁ and C₂ cannot be established (i.e., both involve a sonorant + strident sequence).

The phonetic factor defining the relative markedness between sonorant + voiceless strident (/ls, ns) and sonorant + voiced strident (/lz, nz/) is the articulatory voicing effect.

a) Markedness of C₂

Articulatory voicing effect: vowel insertion is more likely to occur after a voiced obstruent due to the articulatory difficulty of maintaining voicing in an obstruent.

4.5.1.1. Empirical Evidence for ls, ns < lz, nz

(1) Second language data

The relative markedness between sonorant + voiceless strident (ls, ns) and sonorant + voiced strident (lz, nz) is supported by the findings of Hansen (2000), who tested two Vietnamese adult learners of English on the production of English single coda and coda clusters.

The Vietnamese learners of English in her study arrived in the United States within one year of the onset of the research. Hansen (2000) collected data on a weekly basis for a duration of ten months to see the development of single codas and coda clusters. Two types of phonological data, from interviews, and word list and passage reading, were gathered at three intervals during the duration of the study. Table 4.12 provides the results from Hansen (2000: 203-206). Correct production of each cluster at three different times is illustrated in Table 4.12. At Time 1, nasal + voiceless fricative clusters are produced 44% correctly, while in Time 2, they are produced 46% correctly. The last column is the total of correct production of each cluster at three different times

(100% x 3 = 300%).⁵⁰ Note that bold type indicates the clusters that are relevant to the markedness ranking relationship under discussion.

Table 4.12. Evidence for ls, ns < lz, nz (Hansen 2000: 203-206)⁵¹

Cluster type	Time 1 (acc %)	Time 2	Time 3	Total ⁵² (300%)
nasal + voiceless fricative	44	46	50	140
nasal + voiceless stop	21	57	40	118
nasal + voiced fricative	40	34	32	106
voiceless stop + voiceless fricative	32	23	37	92
voiced stop + voiced fricative	33	9	41	83
liquid + voiceless fricative	21	22	30	73
liquid + voiced fricative	38	19	12	69
voiceless fricative + voiceless stop	17	27	24	68
nasal + voiced stop	0	17	50	67
liquid + voiceless stop	19	23	16	58
liquid + nasal	11	17	29	57
voiced stop + voiced stop	0	8	17	25
liquid + voiced stop	2	0	16	18
voiced fricative + voiced stop	7	0	7	14
voiceless stop + voiceless stop	3	0	5	8

* Bold type indicates the clusters that are relevant to ls, ns < lz, nz.

We can see in Table 4.12 that Vietnamese learners produced sonorant + voiceless fricative correctly more frequently than sonorant + voiced fricative. For instance, the percentage of correct production in nasal + voiceless fricative is 140%, while nasal + voiced fricative was produced 106% correctly (out of 300% total correct production).

4.5.2. ls, ns < lp, lt, lk, np, nt, mp

⁵⁰ The correct production of English coda clusters by Vietnamese speakers was not averaged over the three different times; it was instead added to equal 300 %. This was done because subtle differences in the correct production depending on the cluster type might have been lost by averaging the percentage of correct production over three times.

⁵¹ This table is reused throughout this section, but only the relevant data will be presented later in the section.

⁵² The total in this column is not from the original study. I added up the accuracy rate in Time 1, Time 2, and Time 3 to show the relative markedness of each cluster.

This section defines the relative markedness between sonorant + voiceless strident (ls, ns) and sonorant + voiceless stop (lp, lt, lk, np, nt, mp). Note that in these two clusters, C_1 is fixed with a sonorant, and thus the relative markedness of C_1 is not relevant for the consonant cluster markedness evaluation. Thus, the relative markedness between sonorant + voiceless strident and sonorant + voiceless stop is determined based on a) the markedness of C_2 and b) the relative markedness of gestural coordination of C_1 and C_2 .

a) Markedness of C_2

- i) Articulatory stridency effect: vowel insertion is less likely to occur after a voiceless strident because the aerodynamic conditions of /s/ remain conducive to the continued turbulence of a strident rather than vocal fold vibration.
- ii) Principle of perceptual saliency: the likelihood of vowel insertion and consonant deletion correlates with the perceptual salience of a segment in contrast; the weaker the perceptual cues of a segment, the more likely it is that vowel insertion or consonant deletion occurs. Thus, a stop is the target of vowel insertion and consonant deletion due to its relatively weak perceptual salience.

b) Markedness of gestural coordination of C_1 and C_2

- i) Degree of gestural overlap: sonorant + voiceless strident is less marked than sonorant + voiceless stop because a strident is less overlapped by an adjacent consonant than is a stop (less gestural overlap = less marked).

4.5.2.1. Empirical Evidence for ls, ns < lp, lt, lk, np, nt, mp

(1) Second language data

The first piece of evidence supporting the relative markedness between sonorant + voiceless strident (ls, ns) and sonorant + voiceless stop (lp, lt, lk, np, nt, m) is found in Hansen (2000). Hansen tested two Vietnamese adult learners of English on the production of English single codas and coda clusters. Table 4.13 presents relevant results from Hansen (2000: 203).

Table 4.13. Evidence for ls, ns < lp, lt, lk, np, nt, mp (Hansen 2000: 203-206)

Cluster type	Time 1 (acc %)	Time 2	Time 3	Total (300%)
nasal + voiceless fricative	44	46	50	140
nasal + voiceless stop	21	57	40	118
liquid + voiceless fricative	21	22	30	73
liquid + voiceless stop	19	23	16	58

Table 4.13 shows that nasal + voiceless fricative are produced correctly more frequently (140%) than nasal + voiceless stop (118%). In the same fashion, liquid + voiceless fricative (73%) are produced correctly more frequently than liquid + voiceless stop (58%).

Recall that one of the phonetic factors defining the relative markedness between sonorant + voiceless strident (ls, ns) and sonorant + voiceless stop (lp, lt) is the *Principle of Perceptual Saliency*. According to the principle of perceptual saliency, stop deletion is more likely to occur in sonorant + voiceless stop (lp, lt) than fricative deletion in sonorant + voiceless fricative (ls, ns) due to the relatively weak perceptibility of a stop compared to a voiceless strident. This prediction is supported in Hansen (2000: 203). At Time 1, the

Vietnamese learners made 14% of deletion errors in liquid + voiceless fricative, while the deletion error rate in liquid + voiceless stop was 41%. The lower rate of deletion in liquid + voiceless fricative is consistent in all three time periods (Time 2: 24% vs. 32%, Time 3: 22% vs. 47%).

Similarly, several previous studies on L2 acquisition of consonant clusters reported that a stop is the main target of deletion in coda clusters. For instance, in Spanish speakers' production of German coda clusters, Tropic (1987: 183-185) observed that, both in a single coda and in coda clusters, the deletion of a final stop is quantitatively dominant. In addition, Hansen notes that a stop is the main target of deletion (e.g., *and* → deletion of /d/) in Mandarin Chinese speakers' production of English codas (Hansen 2001: 356-357).⁵³

4.5.3. lp, lt, lk, np, nt, mp < lb, ld, nd

This section defines the relative markedness between sonorant + voiceless stop (/lp, lt, lk, np, nt, mp/) and sonorant + voiced stop (/lb, ld, nd/). The relative markedness between these two types of clusters is determined solely based on the markedness of C₂ for two reasons: a) C₁ is fixed with a sonorant and b) the relative markedness of the gestural coordination of C₁ and C₂ cannot be established (i.e., both involve a sonorant + stop sequence).

⁵³ That a stop is the main target of deletion is also attested in the Cantonese loanword adaptation process. When English words are nativized in Cantonese, a strident is allowed to occur anywhere regardless of the context (e.g., *tips* /tips/ → [t^hipsi]) whereas a stop gets deleted if not adjacent (in the English word) to a vowel or liquid (e.g., *forecast* /fɔrkæst/ → [fok^hsi]) (Steriade 2001a).

a) Markedness of C₂

- i) Articulatory voicing effect: vowel insertion is more likely to occur after a voiced obstruent due to the articulatory difficulty of maintaining voicing in an obstruent.

The articulatory voicing effect determines the relative markedness between these two clusters as follows: lp, lt, lk, np, nt, mp < lb, ld, nd

4.5.3.1. Empirical Evidence for lp, lt, lk, np, nt, mp < lb, ld, nd

Empirical evidence supporting the relative markedness relationship between sonorant + voiceless stop (lp, lt) and sonorant + voiced stop (lb, ld) is found in Hansen (2000).

Relevant data are presented in Table 4.14.

Table 4.14. Evidence for lp, lt, lk, np, nt, mp < lb, ld, nd (Hansen 2000: 203-206)

Cluster type	Time 1 (acc %)	Time 2	Time 3	Total (300%)
nasal + voiceless stop	21	57	40	118
nasal + voiced stop	0	17	50	67
liquid + voiceless stop	19	23	16	58
liquid + voiced stop	2	0	16	18

Note in Table 4.14 that nasal + voiceless stop (118%) is produced correctly more frequently than nasal + voiced stop (67%). The exact same pattern is observed in liquid + voiceless stop (58%) and liquid + voiced stop (18%).

Mandarin Chinese speakers' production of English codas in Hansen (2001) demonstrates a similar pattern. The data in Hansen (2001) were collected at two different

times with a time span of 6 months, and only naturalistic speech (interview with a researcher) was used as the method of data collection. Specific codas were not elicited or controlled, and therefore some types of coda clusters are missing from the data. Also, participants' data were not pooled by type of clusters (e.g., /lp/ and /lt/ are presented separately rather than as one natural class category, i.e., liquid + voiceless stop).

Table 4.15. Evidence for lp, lt, lk, np, nt, mp < lb, ld, nd (Hansen 2001: 351-352)

Cluster type		Time 1 (acc %)	Time 2	Total (200%) ⁵⁴
liquid + voiced fricative	/lz/	100	100	200
nasal + voiced fricative	/mz/	100	100	200
voiced stop + voiced fricative	/dz/	100	100	200
voiceless stop + voiceless stop	/pt/	100	100	200
liquid + voiceless stop	/lp/	88	100	188
voiceless stop + voiced fricative	/ts/	95	87	182
nasal + voiced fricative	/nz/	94	81	175
voiceless stop + voiced fricative	/ks/	89	83	172
liquid + voiced stop	/ld/	83	75	158
voiced stop + voiced fricative	/gz/	71	75	146
liquid + voiceless stop	/lt/	80	60	140
voiceless stop + voiceless stop	/kt/	40	100	140
nasal + voiceless stop	/nt/	71	67	138
voiceless fricative + voiceless stop	/sk/	83	50	133
voiceless fricative + voiceless stop	/st/	63	68	131
voiced fricative + voiced stop	/zd/	43	56	99
liquid + nasal	/rn/	40	33	73
nasal + voiced stop	/nd/	25	30	55

* Bold type indicates the clusters that are relevant to lp, lt, lk, nt, np, mp < lb, ld, nd.

We can see in Table 4.15 that the prediction of PCCAM on the relative markedness between sonorant + voiceless stop (lp, lt) and sonorant + voiced stop (lb, ld) holds true except for one case. /nt/ (138%) is produced correctly more frequently than /nd/ (55%), and /lp/ (188%) is produced correctly more often than /ld/ (158%), as was predicted by

⁵⁴ The total in this column is not from the original study. I added up the accuracy rate in Time 1 and Time 2 to show the relative markedness of each cluster.

PCCAM. Yet, /ld/ (158%) is produced correctly more frequently than /lt/ (140%), which contradicts the prediction of PCCAM. The problem of this conflicting data will be resolved through the analysis of Korean speakers' production of English consonant clusters presented in Chapter 6.

4.5.4. lp, lt, lk, np, nt, mp < lm

The phonetic factor defining the relative markedness between sonorant + voiceless stop (lp, lt) and liquid + nasal (lm) is the maximal perceptual contrast parameter. In terms of the maximal perceptual contrast parameter, a liquid + nasal coda cluster is worse than liquid + voiceless stop. This is because all three of the segments (vowel + liquid + nasal) in a liquid + nasal cluster share the same sonorant properties, and are thus subject to confusion. By contrast, sonorant + voiceless stop clusters are less likely to be subject to this kind of confusion because a sonorant and a voiceless stop make a good perceptual contrast with each other. Note that obstruents (in particular voiceless ones) involve long durations of silence, while a sonorant is one of the loudest sounds. Thus, the strings containing sonorant and voiceless obstruent in alternating sequences may result in good perceptual contrast, which facilitates the perceptual saliency of each element in the string.

a) Markedness of a C₁ and C₂ sequence

- i) Maximal perceptual contrast: language learners prefer a sequence where two segments in an alternating sequence make a maximal perceptual contrast.

The maximal perceptual contrast determines that sonorant + voiceless stop is less marked than liquid + nasal.

4.5.4.3. Empirical Evidence for lp, lt, lk, np, nt, mp < lm

Supporting evidence for the relative markedness of the cluster /lm/ over sonorant + voiceless stop (lp, lt) is found in Vietnamese speakers' production of English consonant clusters (Hansen 2000) and Mandarin Chinese speakers' production of English consonant clusters (Hansen 2001).

Note that in Table 4.16 Vietnamese learners of English have more difficulty in producing liquid + nasal (57%) than nasal + voiceless stop (118%), even though the difference in correct production rates between liquid + voiceless stop (58%) and liquid + nasal (57%) is negligible.

Table 4.16. Evidence for lp, lt, lk, np, nt, mp < lm (Hansen 2000: 203-206)

Cluster type	Time 1 (acc %)	Time 2	Time 3	Total (300%)
nasal + voiceless stop	21	57	40	118
liquid + voiceless stop	19	23	16	58
liquid + nasal	11	17	29	57

The same pattern is observed in Mandarin Chinese speakers' production of English coda clusters. Even though the /lm/ sequence was not produced by the participants, we can see in Table 4.17 below that the rate of correct production of a similar sequence /rn/ (73%) is far lower than that of sonorant + voiceless stop (e.g., /lp/ = 188%).

Table 4.17. Evidence for lp, lt, lk, np, nt, mp < lm (Hansen 2001: 351-352)

Cluster type		Time 1 (acc %)	Time 2	Total (200%)
liquid + voiceless stop	/lp/	88	100	188
liquid + voiceless stop	/lt/	80	60	140
nasal + voiceless stop	/nt/	71	67	138
liquid + nasal	/rn/	40	33	73

To summarize thus far, we have discussed the relative markedness relationships among sonorant + C clusters. In view of all the evidence presented, it seems safe to say that the predictions of PCCAM on the relative markedness among sonorant + C coda clusters hold true in the published L2 consonant cluster acquisition data.

As a summary of this section, Table 4.18 presents the phonetic factors defining each ranking relationship in sonorant + C coda clusters along with the empirical data supporting the ranking relationships.

Table 4.18. Summary of sonorant + C clusters with empirical data

Relative markedness	Phonetic factors	Empirical data
ls, ns < lz, nz	a. articulatory voicing effect	Hansen (2000)
ls, ns < lp, lt, lk	a. articulatory stridency effect	Hansen (2000)
	b. principle of perceptual saliency	Trope (1987)
	c. degree of gestural overlap	(patterns of consonant deletion)
lp, lt lk < lb, ld, nd	a. articulatory voicing effect	Hansen (2000)
		Hansen (2001)
		(partially supported)
lp, lt, lk < lm	a. maximal perceptual contrast	Hansen (2000)
		Hansen (2001)

Note: Clusters in the first column only serve as the representative case of clusters in each category.

4.5.5. ps, ts, ks < sp, st, sk

This section defines the relative markedness between voiceless stop + voiceless strident (ps, ts, ks) and voiceless strident + voiceless stop (sp, st, sk). The relative markedness between these two types of clusters is determined based on a) the relative markedness of C₁, b) the relative markedness of C₂, and c) the relative markedness of the gestural coordination of C₁ and C₂.

a) Markedness of C₁

- i) Articulatory stridency effect: vowel insertion is less likely to occur after a voiceless strident because the aerodynamic conditions of /s/ remain conducive to the continued turbulence of a strident rather than vocal fold vibration. This parameter determines that /sp, st, sk/ is less marked than /ps, ts, ks/ in terms of the markedness of C₁.
- ii) Perceptual stridency effect: vowel insertion is less likely to occur after the voiceless strident /s/ because an epenthetic vowel after a voiceless strident is perceptually more obtrusive than an epenthetic vowel after a stop. This parameter determines that /sp, st, sk/ is less marked than /ps, ts, ks/ in terms of the markedness of C₁.⁵⁵

b) Markedness of C₂

- i) Articulatory stridency effect: when it come to the markedness of C₂, the

⁵⁵ Note that the role of the perceptual stridency effect in defining the relative markedness between voiceless stop + C and voiceless strident + C coda clusters will be revised in Chapter 5 based on the results of the perceptual similarity judgment experiment.

articulatory stridency effect determines that voiceless stop + voiceless strident (ps, ts, ks) is less marked than voiceless strident + voiceless stop (sp, st, sk).

- ii) Principle of perceptual saliency: PCCAM predicts that stop deletion and vowel insertion is more likely to occur in a sequence of voiceless strident + voiceless stop (sp, st, sk) because the perceptually weak segment, a stop, is not flanked by a vowel in this sequence.

c) Markedness of gestural coordination of C₁ and C₂

- i) Degree of gestural overlap: voiceless stop + voiceless strident (ps, ts, ks) is less marked than voiceless strident + voiceless stop (sp, st, sk) because the gestures in the former sequence are less overlapped than the gestures in the latter sequence.
- ii) Articulatory release effect: a consonant sequence is easier for language learners to acquire when a stop is followed by [+continuant] than when a stop is followed by [-continuant] (ps, ts, ks < pt, kt).

Note that the two types of consonant clusters (sp vs. ps) are mirror images of each other, and therefore, there is no winner in terms of the articulatory and perceptual stridency effects in the two sequences.

Note also that the articulatory release effect in (b, ii) does not directly define the markedness relationship in ps, ts, ks < sp, st, sk, yet it defines the markedness relationship in **ps, ts, ks** < sp, st, sk < **pt, kt** (bold type indicates the markedness relationship defined by the articulatory release effect). The articulatory release effect is

presented here to illustrate the relative unmarkedness of voiceless stop + voiceless strident among obstruent + obstruent coda clusters.

To summarize, the principle of perceptual saliency and the degree of gestural overlap determine that voiceless stop + voiceless strident (ps, ts, ks) is less marked than voiceless strident + voiceless stop (sp, st, sk).

4.5.5.1. Empirical Evidence for ps, ts, ks < sp, st, sk

(1) Child language data

Empirical evidence supporting the ranking relationship between voiceless stop + voiceless fricative (ps, ts, ks) and voiceless fricative + voiceless stop (sp, st, sk) is found in Templin (1957). Templin (1957) established the age of onset and coda cluster acquisition based on the data from 480 children (age 3:0-8:0). Table 4.19 presents the age of acquisition of coda clusters in this study.

Table 4.19. Developmental order in coda clusters (Templin 1957: 51)

Age of acquisition	
/ks/	3:5
/sk/, /st/	7:0
/sp/	8:0

Notice in Table 4.19 that there is a big gap between the age at which voiceless stop + fricative (ks) is acquired and the age at which voiceless fricative + voiceless stop (sp, st, sk) is acquired.

Similar developmental order differences are observed in Dyson (1988). Dyson (1988) tested 20 children at the age of 2:0-3:3 and observed that at the mean age of 3:3,

the word-final clusters found in these children's speech are /ps, ts, ns, ntʃ, ŋk/. Note that two voiceless stop + voiceless fricative clusters (ps, ts) appeared, while none of the voiceless fricative + voiceless stop clusters had appeared yet.

The high frequency of stop + fricative in child speech is also observed in Jakielski (1998). Five infants' production of English consonant clusters was observed in this study (age 0;7-3;6, longitudinal study). Jakielski (1998: 189) notes that for word-initial clusters, the infants produced stop + glide (55%) and fricative + stop (15%) clusters most frequently. For word-final clusters, the five infants produced stop + fricative (45%) and nasal + stop (32%) clusters most frequently. Note that stop + fricative clusters are the most frequent word-final clusters occurring in the infants' production, which illustrates the relative unmarkedness of stop + fricative coda clusters in child speech.

Lastly, Kirk & Demuth (2003) examined the acquisition of English onset and coda clusters in a cross-sectional experimental study (n = 9 aged 2). In this study, there was a noticeable difference in the number of children who produced stop + /s/ and /s/ + stop clusters with greater than 70% accuracy: seven children produced stop + /s/ with greater than 70% accuracy, while there were only two children who produced a /s/ + stop cluster with greater than 70% accuracy. It is important to note that the most common type of error associated with /s/ + stop clusters was metathesis: 23% of children's attempted /s/ + stop clusters resulted in metathesis (e.g., wasp → waps), and six of the nine children in the study made this type of metathesis error (Kirk & Demuth 2003: 446).

Metathesis from /s/ + stop to stop + /s/ sequences may be interpreted as children's preference for less marked consonant sequences. That is, metathesis of /s/ + stop to stop + /s/ changes a more marked consonant sequence to a less marked one in terms of the

degree of gestural overlap parameter and the principle of perceptual saliency.

(2) Second language data

Major (1996) tested four adult native speakers of Brazilian Portuguese on their production of English consonant clusters. VARBRUL analyses were performed to determine the probability of errors for each cluster. The results show that fricative + stop (0.75) contributes more strongly to error than stop + fricative (0.16) (notice the large difference in error probability).

The same patterns are found in Hansen (2000; 2001). Note in Table 4.20 and 4.21 that both Vietnamese and Mandarin Chinese learners of English produced voiceless stop + voiceless fricative (92% for Vietnamese speakers) correctly more frequently than voiceless fricative + voiceless stop (68% for Vietnamese speakers).

Table 4.20. Evidence for ps, ts, ks < sp, st, sk (Hansen 2000: 203-206)

Cluster type	Time 1 (acc %)	Time 2	Time 3	Total (300%)
voiceless stop + voiceless fricative	32	23	37	92
voiceless fricative + voiceless stop	17	27	24	68

Table 4.21. Evidence for ps, ts, ks < sp, st, sk (Hansen 2001: 351-352)

Cluster type		Time 1 (acc %)	Time 2	Total (200%)
voiceless stop + voiceless fricative	/ts/	95	87	182
voiceless stop + voiceless fricative	/ks/	89	83	172
voiceless fricative + voiceless stop	/sk/	83	50	133
voiceless fricative + voiceless stop	/st/	63	68	131

Most recently, Wiltshire (2005) tested speakers of five different Indian L1 dialects

(Angami, Ao, Mizo, Gujarati, and Hindi) and observed that deletions were rare for stop + /s/ (4 instances), compared to fricative-stop clusters (22 instances). Table 4.22 below summarizes these data.

Table 4.22. Evidence for ps, ts, ks < sp, st, sk (Wiltshire 2005: 3)

Cluster (# of token)	Angami	Ao	Mizo	Gujarati	Hindi
Nasal-stop (30)	2 (# of deletion)	1	4	2	0
Lateral-stop (20)	2	2	4	1	1
Lateral-nasal (5)	1	2	3	2	0
Fricative-stop (20)	7	3	9	2	1
stop-/s/ (30)	0	0	2	2	0
stop-stop (20)	9	9	13	2	1

Similarly, Ruffner (2000) tested three Japanese speakers, three Spanish speakers, two Mandarin speakers, and two Indonesian speakers on their production of English coda clusters. It was found in the study that the rate of word-final epenthesis (after C₂) was much lower for C + fricative clusters (24.4%) than for C + stop clusters (88.9%). The finding of Ruffner (2000) seems to support the role of the articulatory stridency effect and the principle of perceptual saliency in PCCAM: both parameters predict a more frequent vowel insertion after a stop than after a strident.

4.5.6. sp, st, sk < pt, kt

This section defines the relative markedness between voiceless strident + voiceless stop (sp, st, sk) and voiceless stop + voiceless stop (pt, kt). The relative markedness between these two types of clusters is determined based on a) the markedness of C₁ and b) the relative markedness of the gestural coordination of C₁ and C₂. The markedness of C₂ is

not relevant for evaluation because it is fixed with a voiceless stop in these two clusters.

a) Markedness of C₁

- i) Articulatory stridency effect: vowel insertion is less likely to occur after a voiceless strident because the aerodynamic conditions of /s/ remain conducive to the continued turbulence of a strident rather than vocal fold vibration (sp, st, sk < pt, kt).
- ii) Perceptual stridency effect: vowel insertion is less likely to occur after the voiceless strident /s/ because an epenthetic vowel after a voiceless strident is perceptually more obtrusive than an epenthetic vowel after a stop. This parameter determines that /sp, st, sk/ is less marked than /pt, kt/ in terms of the markedness of C₁.⁵⁶

b) Markedness of gestural coordination of C₁ and C₂

- i) Degree of gestural overlap: voiceless strident + voiceless stop (sp, st, sk) is less marked than voiceless stop + voiceless stop (pt, kt) because the gestures in the former sequence are less overlapped than the gestures in the latter sequence (sp, st, sk < pt, kt).

The computation of the two aforementioned phonetic parameters determines the relative unmarkedness of voiceless strident + voiceless (sp, st, sk) over voiceless stop + voiceless stop (pt, kt).

⁵⁶ Note that the role of the perceptual stridency effect in defining the relative markedness between voiceless stop + C and voiceless strident + C coda clusters will be revised in Chapter 5 based on the results of the perceptual similarity judgment experiment.

4.5.6.1. Empirical Evidence for sp, st, sk < pt, kt

(1) Second language data

Empirical evidence for the ranking relationship between voiceless fricative + voiceless stop (sp, st, sk) and voiceless stop + voiceless stop (pt, kt) is found in Major (1996). This study showed that stop + stop coda clusters contribute more strongly to error (0.98) than fricative + stop coda clusters (0.75). Vietnamese speakers of English in Hansen (2000) demonstrate the same pattern. Note in Table 4.23 that voiceless fricative + voiceless stop is produced correctly more frequently (68%) than voiceless stop + voiceless stop clusters (8%).

Table 4.23. Evidence for sp, st, sk < pt, kt (Hansen 2000: 203-206)

Cluster type	Time 1(acc %)	Time 2	Time 3	Total (300%)
voiceless fricative + voiceless stop	17	27	24	68
voiceless stop + voiceless stop	3	0	5	8

Contrary to the data in Major (1996) and Hansen (2000), Mandarin Chinese speakers' production data in Hansen (2000) do not support the proposed ranking relationship.

Table 4.24. Evidence for sp, st, sk < pt, kt (Hansen 2001: 351-352)

Cluster type		Time 1 (acc %)	Time 2	Total (200%)
voiceless stop + voiceless stop	/pt/	100	100	200
voiceless stop + voiceless stop	/kt/	40	100	140
voiceless fricative + voiceless stop	/sk/	83	50	133
voiceless fricative + voiceless stop	/st/	63	68	131

Table 4.24 illustrates that Mandarin Chinese speakers did better on voiceless stop +

voiceless stop (pt, kt) than voiceless fricative + voiceless stop clusters (sk, st). The conflicting markedness relationship between these two clusters is also observed in Korean speakers' production of English consonant clusters in the present study. Further discussion of this conflicting data will be presented in Chapter 6.

4.5.7. pt, kt < bd, gd

This section defines the relative markedness between voiceless stop + voiceless stop (/pt, kt/) and voiced stop + voiced stop (/bd, gd/). The relative markedness between these two types of clusters is determined based on a) the relative markedness of C₁ and b) the relative markedness of C₂. The markedness of gestural coordination of C₁ and C₂ is not relevant because both clusters involve a stop + stop sequence.

a) Markedness of C₁

- i) Articulatory voicing effect: vowel insertion is more likely to occur after a voiced obstruent due to the articulatory difficulty of maintaining voicing in an obstruent (pt, kt < bd, gd).
- ii) Perceptual voicing effect: vowel insertion is more likely to occur after a voiced obstruent because an epenthetic vowel after a voiced obstruent is perceptually less obtrusive than an epenthetic vowel after a voiceless obstruent (pt, kt < bd, gd).

b) Markedness of C₂

- i) Articulatory voicing effect: vowel insertion is more likely to occur after a

voiced obstruent due to the articulatory difficulty of maintaining voicing in an obstruent (pt, kt < bd, gd).

We can see that voiceless stop + voiceless stop (pt, kt) is less marked than voiced stop + voiced stop (bd, gd) due to the articulatory and perceptual voicing effect.

Empirical data for this ranking relation are presented in Table 4.26.

4.5.8. bz, gz < zd

This section defines the relative markedness between voiced stop + voiced strident (bz, gz) and voiced strident + voiced stop (zd). The relative markedness between these two types of clusters is determined based on a) the relative markedness of C_2 and b) the relative markedness of gestural coordination of C_1 and C_2 . Three phonetic factors account for the relative markedness relationship between these two types of clusters.

a) Markedness of C_2

- i) Principle of perceptual saliency: PCCAM predicts that stop deletion and vowel insertion is more likely to occur in voiced strident + voiced stop (zd) because the perceptually weak segment, a stop, is not flanked by a vowel in this sequence (bz, gz < zd).

b) Markedness of gestural coordination of C_1 and C_2

- i) Degree of gestural overlap: voiced stop + voiced strident (bz, gz) is less marked than voiced strident + voiced stop (zd) because the gestures in the

former sequence is less overlapped than the gestures in the latter sequence (bz, gz < zd).

- ii) Articulatory release effect: a consonant sequence is easier for language learners to acquire when a stop is followed by [+continuant] than when a stop is followed by [-continuant] (bz, gz < bd, gd)

Note that the articulatory release effect in (b, ii) does not directly define the markedness relationship in bz, gz < zd, yet it is presented here to demonstrate the relative unmarkedness of voiced stop + voiced strident (bz, gz) among obstruent + obstruent clusters in terms of the articulatory release effect (**bz, gz** < zd < **bd, gd**; bold type indicates the ranking relationship determined by the articulatory release effect).

The computation of the aforementioned phonetic factors determines that voiced stop + voiced strident (bz, gz) is less marked than voiced strident + voiced stop (zd). Empirical data supporting this markedness relationship is presented in Table 4.25.

4.5.9. **zd < bd, gd**

This section defines the relative markedness between voiced strident + voiced stop (zd) and voiced stop + voiced stop (bd, gd). The relative markedness between these two types of clusters is determined based on the relative markedness of gestural coordination of C₁ and C₂. The markedness of C₁ is not relevant for evaluation because C₁ in both clusters shares the same voicing feature. Also, the markedness of C₂ is not relevant for evaluation because it is fixed with a voiced stop in both types of clusters.

a) Markedness of gestural coordination of C₁ and C₂

- i) Degree of gestural overlap: voiced strident + voiced stop (zd) is less marked than voiced stop + voiced stop (bd) because the gestures in the former sequence are less overlapped than the gestures in the latter sequence (zd < bd).

4.5.9.1. Empirical Evidence for pt, kt < bz, gz < zd < bd, gd

bz, gz < zd

Table 4.25 presents empirical data from Hansen (2001) supporting the relative markedness between voiced stop + voiced strident (bz, gz) and voiced strident + voiced stop (zd). Notice the big difference in the percentage of accurate production between /dz, gz/ (200% and 146% respectively) and /zd/ (99%).

Table 4.25. Evidence for bz, gz < zd (Hansen 2001: 351-352)

Cluster type		Time 1 (acc %)	Time 2	Total (200%)
voiced stop + voiced fricative	/dz/	100	100	200
voiced stop + voiced fricative	/gz/	71	75	146
voiced fricative + voiced stop	/zd/	43	56	99

bz, gz < bd, gd

Supporting evidence for the relative markedness between voiced stop + voiced strident (bz, gz) and voiced stop + voiced stop (bd, gd) is found in Hansen (2000). Table 4.26 below shows that Vietnamese learners of English performed better on voiced stop + voiced fricative (83%) than voiced stop + voiced stop (25%).

Table 4.26. Evidence for pt, kt < bd, gd and bz, gz < bd, gd (Hansen 2000)

Cluster type	Time 1 (acc %)	Time 2	Time 3	Total (300%)
voiced stop + voiced fricative	33	9	41	83
voiced stop + voiced stop	0	8	17	25
voiceless stop + voiceless stop	3	0	5	8

pt, kt < bd, gd

However, the prediction of PCCAM on the ranking relationship between voiceless stop + voiceless stop (pt, kt) and voiced stop + voiced stop (bd, gd) is not supported in Hansen (2000). Note in Table 4.26 that the correct production of voiced stop + voiced stop (25%) is higher than that of voiceless stop + voiceless stop (8%).

Upon closer examination of the data in Hansen (2000), we can see that a large number of deletions occur in voiceless stop + voiceless stop clusters. Deletion error constitutes 72% of voiceless stop + voiceless stop cluster production for Time 1, 74% for Time 2, and 79% for Time 3. It is unclear why stop deletion occurs more often for a voiceless stop + voiceless stop than for a voiced stop + voiced stop in Hansen (2000). This kind of conflicting data will be resolved through the analysis of Korean speakers' production of English consonant clusters presented in Chapter 6.

4.5.10. Summary of the Relative Markedness of Coda Consonant Cluster

Table 4.27 summarizes the articulatory and perceptual factors defining consonant cluster markedness in coda clusters. Table 4.27 also lists the empirical data relevant to each markedness relationship.

Table 4.27. Summary of coda consonant clusters with empirical data

Relative markedness relationship	Phonetic factors	Empirical data
ls, ns < lz, nz	a. articulatory voicing effect	Hansen (2000)
ls, ns < lp, lt, lk, np, nt, mp	a. articulatory stridency effect b. principle of perceptual saliency d. degree of gestural overlap	Hansen (2000) Tropf (1987) (patterns of consonant deletion)
lp, lt, lk, np, nt, mp < lb, ld, nd	a. articulatory voicing effect	Hansen (2000)
lp, lt, lk, np, nt, mp < lm	a. maximal perceptual contrast	Hansen (2001) (partially supported) Hansen (2000) Hansen (2001)
ps, ts, ks < sp, st, sk	a. principle of perceptual saliency	Dyson (1988)
	b. degree of gestural overlap	Hansen (2000)
	c. articulatory release effect (ps, ts, ks < pt, kt)	Hansen (2001) Jakielski (1998) Kirk & Demuth (2003) Major (1996) Olmsted (1971) Templin (1957) Wellman et al. (1931) Wiltshire (2005)
sp, st, sk < pt, kt	a. articulatory stridency effect b. perceptual stridency effect c. degree of gestural overlap	Major (1996) Hansen (2000) Hansen (2001) not supported
pt, kt < bd, gd	a. articulatory voicing effect b. perceptual voicing effect	Hansen (2000): not supported
bz, gz < zd	a. principle of perceptual saliency b. degree of gestural overlap c. articulatory release effect (bz, gz < bd, gd)	Hansen (2000) (bz, gz < bd, gd)
zd < bd, gd	a. degree of gestural overlap	Hansen (2001)

4.6. Summary of Chapter 4

This chapter was intended to provide theoretical and empirical justifications for the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM). As a theoretical framework, Articulatory Phonology and the P-map were adopted and their working principles were reviewed in § 4.1 and § 4.2.

The main tenet of Articulatory Phonology (AP) that phonological structures are organized patterns of gestural units has important implications for language acquisition. Under AP, phonological acquisition is viewed as differentiating, tuning, and coordinating the pre-linguistic gestures so that those gestures can be used as the units of contrast. Differentiation and tuning of gestures are required to develop phonemic contrast, while gestural coordination is required to combine the gestures for a given segment or segmental sequence.

The degree of gestural overlap parameter was derived from the view of AP that the possible clusters in a language correlate with the degree of gestural overlap and its variability. Referring to the work of Byrd (1996) on the different degree of gestural overlap depending on the consonant sequence, I proposed that a consonantal sequence with less gestural overlap is easier for language learners to acquire than a consonantal sequence with more gestural overlap.

Along with the degree of gestural overlap parameter, I have developed the articulatory release effect and the articulatory voicing effect. The articulatory release effect proposes that a consonant sequence is easier for language learners to acquire when a stop is followed by [+continuant] than when a stop is followed by [-continuant]. On the

other hand, the articulatory voicing effect states that vowel insertion is more likely to occur after a voiced obstruent due to the phonetic difficulty of maintaining voicing in a voiced obstruent.

I have shown that perceptual factors also play an important role in defining consonant cluster markedness in PCCAM. The main tenet of the P-map (Steriade 2001a, 2001b) that a sound change whose modification is less perceptible is more likely to occur than a sound change whose modification is more perceptible was adopted as a guiding principle in identifying the perceptual factors in PCCAM.

The P-map hypothesis makes the following prediction regarding L2 acquisition of consonant clusters. An epenthetic vowel is more likely to occur after a voiced obstruent than after a voiceless obstruent because an epenthetic vowel after a voiced obstruent is perceptually less obtrusive (due to the voicing properties shared by a voiced obstruent and an epenthetic vowel) than an epenthetic vowel after a voiceless obstruent. This phonetic factor is named the perceptual voicing effect in PCCAM. On the other hand, the perceptual stridency effect in PCCAM predicts that vowel insertion is less likely to occur after the voiceless strident /s/ because an epenthetic vowel after a voiceless strident is perceptually more obtrusive than an epenthetic vowel after a stop. Note that both the perceptual stridency effect and perceptual voicing effect crucially refer to the principle of the P-map.

The principle of perceptual saliency proposed by Côté (2000) is also in line with the P-map hypothesis in that the perceptibility of a certain segment plays a crucial role in the process of vowel insertion and consonant deletion. According to the principle of perceptual saliency, the likelihood of vowel insertion and consonant deletion correlates

with the perceptual salience of a segment in contrast; the weaker the perceptual cues of a segment, the more likely it is that vowel insertion or consonant deletion will occur.

A stop is well known for its weak perceptibility, in particular when it is not flanked by a vowel. In contrast, a strident has very strong internal cues characterized by its relatively long duration and its hissing properties. Thus, the principle of perceptual saliency predicts that a stop is the target of deletion and vowel insertion, but not a strident.

The maximal perceptual contrast parameter was employed when defining the ranking relationships in a) sC onset clusters and b) between sonorant + voiceless stop and liquid + nasal. The basic idea of the maximal perceptual contrast parameter is that that language learners prefer a sequence where two segments in an alternating sequence make a maximal perceptual contrast.

I have shown throughout § 4.4 and § 4.5 that most of the ranking relationships in PCCAM are supported in the published child language and second language data. However, there is one area in which the available published data do not support the relative markedness relationships in PCCAM: the relative markedness between /pt, kt/ and /bd, gd/ (pt, kt < bd, gd) was not supported in Hansen (2001). This kind of conflicting data can be resolved through Korean speakers' production of English consonant clusters in Chapter 6.

Note that among the four articulatory factors and the four perceptual factors in PCCAM, the principle of perceptual saliency, the articulatory release effect, and perceptual stridency effect are derived directly from the work of Côté (2000; the principle of perceptual saliency and the articulatory release effect) and Steriade (2001b; the perceptual stridency effect) in that Côté (2000) and Steriade (2001b) make specific

predictions on the patterns of vowel insertion and consonant deletion using the aforementioned phonetic factors.

On the other hand, the remaining five phonetic factors (i.e., the degree of gestural overlap parameter, the articulatory voicing effect, the articulatory stridency effect, the maximal perceptual contrast parameter, and the perceptual voicing effect) have been developed in this dissertation by referring to the published phonetic or language acquisition studies. For instance, the articulatory stridency effect has been developed based on the vowel devoicing processes in Japanese and Korean.

Lastly, it should be noted that the list of articulatory and perceptual factors in PCCAM may not be complete. The PCCAM in the present form only make predictions on the relative markedness of onset and coda clusters. For instance, PCCAM does not make any predictions on the relative markedness between onset and coda clusters, even though previous studies have shown that coda clusters induce more errors than onset clusters (Hancin-Bhatt & Bhatt 1997: 345; Kim 2000: 207). In addition, it is quite conceivable that the frequency of consonant clusters in the target language also influences L2 learners' production. I think the issue of what additional factors should be included in PCCAM can only be determined through the rigorous empirical tests of this model.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 02845 3938

THESIS

4

2007

V.2

1994

1996

1998

LIBRARY
Michigan State
University

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE
SEP 01 2008		
102308		

KOREAN SPEAKERS' PRODUCTION OF ENGLISH CONSONANT CLUSTERS:
ARTICULATORY AND PERCEPTUAL ACCOUNTS

VOLUME II

By

Bo-Young Kwon

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Linguistics and Germanic, Slavic, Asian, and African Languages

2006

5.1.

The

The

con

Mo

sho

con

pro

voi

(1)

⁵⁷Tw

the p

emp

Fikk

testi

for fi

CHAPTER 5

AUDITORY SIMILARITY JUDGMENT

5.1. Introduction

The auditory similarity judgment experiment in this chapter aims to achieve two goals. The first and foremost goal is to test the validity of the two perceptual principles defining consonant cluster markedness in the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM): the perceptual stridency effect and the perceptual voicing effect. It was shown in § 4.2 that these two perceptual factors play an important role in identifying the context where vowel insertion is more likely to occur when second language learners produce non-native consonant clusters.⁵⁷ The perceptual stridency effect and perceptual voicing effect are restated in (1).

(1) Perceptual factors for vowel insertion

- a. Perceptual stridency effect: vowel insertion is less likely to occur after a voiceless strident than after a stop because an epenthetic vowel after a voiceless strident is perceptually more obtrusive than an epenthetic vowel after a stop.
- b. Perceptual voicing effect: vowel insertion is more likely to occur after a voiced obstruent than after a voiceless obstruent because an epenthetic vowel after a

⁵⁷Two other perceptual factors in PCCAM, i.e., the maximal perceptual contrast parameter and the principle of perceptual saliency, are not tested through the perception experiment. I think the empirical evidence for the role of these two factors is quite robust (as we can see in the work of Fikkert 1994 and Côté 2000), yet we can also benefit from the perception experiment specifically testing the role of these two factors in consonant deletion and vowel insertion. I leave this issue for future research.

P-m

show

dete

sonc

4.2

patt

Côte

after

obsti

Mor

voic

asso

the c

voiced obstruent is perceptually less obtrusive than an epenthetic vowel after a voiceless obstruent.

Recall from § 4.2.2 that the perceptual stridency effect was proposed based on the P-map for vowel epenthesis (Steriade 2001b). The P-map for vowel epenthesis basically shows that an epenthetic vowel in a voiceless strident + obstruent cluster is more easily detected by listeners (i.e., more obtrusive) than an epenthetic vowel in obstruent + sonorant clusters. The P-map for vowel epenthesis is presented in Table 5.1. Refer to § 4.2.2 for the discussion of the P-map for vowel epenthesis and its prediction on the patterns of vowel insertion.

Table 5.1. P-map for vowel epenthesis (Steriade 2001b: 18)

Bigger letter = less similar to unepenthesized input (´ = schwa)

Zero/vowel	T Ri	S T
Ø/ ´	ø/ ´	Ø/ ´
Ø/u	Ø/u	Ø/U
Ø/i	Ø /i	Ø/i

On the other hand, the perceptual voicing effect was proposed in § 4.2.3 based on Côté (2000) and Fleischhacker (2001), both of which contend that an epenthetic vowel after a voiced obstruent is more tolerable than an epenthetic vowel after a voiceless obstruent since a voiced consonant shares voicing properties with an epenthetic vowel. More specifically, Côté (2000: 167) notes that the tendency to insert a vowel after a voiced stop results from the fact that voiced segments share the low frequency energy associated with voicing of vowels; therefore, vowel epenthesis becomes less noticeable in the context of voiced segments.

agre

there

and

scru

stric

pres

judg

inse

pres

part

inter

mar

inse

effe

vow

diffi

voic

⁵⁴ Flé

simil

= cor

other

inser

OvC.

adapt

epent

Even though previous studies (Côté 2000; Fleischhacker 2001; Steriade 2001b) agree upon the role of stridency and voicing in determining the site of vowel insertion, there exists an apparent limitation to relying on these previous studies: the role of voicing and stridency on perceptual similarity judgments has not been subject to empirical scrutiny.⁵⁸ In other words, it is necessary to empirically test the role of voicing and stridency features in the patterns of vowel insertion to augment the proposal in the present study (as well as previous studies). More significantly, the auditory similarity judgment experiment in this chapter is directly correlated with the patterns of vowel insertion in Korean speakers' production of English consonant clusters (which will be presented in Chapter 6) because both experiments involve the same population of participants (Korean EFL learners).

The second goal of the auditory similarity judgment experiment is to illustrate the interaction between articulatory and perceptual factors determining consonant cluster markedness in PCCAM. For instance, in § 4.1.5 and § 4.2.3, I have proposed that vowel insertion is more likely to occur after a voiced obstruent due to the articulatory voicing effect and the perceptual voicing effect. The articulatory voicing effect hypothesizes that vowel insertion is more likely to occur after a voiced obstruent due to the articulatory difficulty of maintaining voicing in an obstruent. On the other hand, the perceptual voicing effect accounts for the frequent occurrence of vowel insertion after a voiced

⁵⁸ Fleischhacker (2001) is an exception to this statement. Fleischhacker (2001) tested the similarity between two sets of stimuli, vSC-SC vs. SvC-SC (v = epenthetic vowel, S = strident, C = consonant, O = obstruent, R = sonorant) and found that the pair vSC-SC is more similar to each other. This result can be used as supporting evidence for the hypothesized P-map for vowel insertion in Table 5.1. However, note that she did not directly compare the pair SvC-SC and OvC-OC (her main concern was to test anaptyxis-prothesis asymmetries, common in loanword adaptation and interlanguage phonology) and thus only indirectly supports the P-map for vowel epenthesis.

o
co
th
pe

pe
the
vo
con
the
voi
play

apar
the a
actu
effec
stride

6), as
provi
cluste
Korea

obstruent in terms of the different perceptibility of an epenthetic vowel depending on the context where it occurs: vowel insertion is more likely to occur after a voiced obstruent than after a voiceless obstruent because an epenthetic vowel after a voiced obstruent is perceptually less obtrusive than an epenthetic vowel after a voiceless obstruent.

The question that arises is to what degree the articulatory voicing effect and the perceptual voicing effect interact with each other. Would it be always the case that both the articulatory and perceptual voicing effects contribute to the frequent occurrence of vowel insertion after a voiced obstruent? Or, would it be the case that depending on the context, the degree to which the articulatory voicing and perceptual voicing effects exert their power is different. I think it is conceivable that in some contexts the articulatory voicing effect plays a major role, while in other contexts the perceptual voicing effect plays a major role.

In short, the second goal of the auditory similarity judgment experiment is to tease apart the role of the perceptual voicing and the perceptual stridency effects from that of the articulatory voicing and articulatory stridency effects. It will be shown that there are actually two cases where the articulatory voicing effect and the articulatory stridency effect play a major role rather than the perceptual voicing effect and the perceptual stridency effect. A discussion of these two cases is presented in § 5.3.2.

Note that the perception experiment precedes the production experiment (Chapter 6), as was already mentioned in Chapter 1. This is because the perception experiment provides justifications for the two perceptual factors defining the relative consonant cluster markedness in PCCAM before testing the validity of PCCAM as a whole through Korean speakers' production of English consonant clusters.

jud

5.1

In c

into

dep

voic

gene

+ C.

+ C.

gene

(2) H

⁵⁰ Thre
strident
clusters

The following section presents the hypotheses for the auditory similarity judgment experiment.

5.1.1. Hypotheses

In order to incorporate the perceptual stridency effect and the perceptual voicing effect into the auditory similarity judgment test, the stimuli in the experiment were varied depending on two distinctive features of C_1 : a) a voicing feature of C_1 (voiced vs. voiceless), and b) a manner feature of C_1 (stop vs. strident). Three groups of stimuli were generated for the auditory similarity judgment test: voiceless strident + C vs. voiced stop + C; voiceless strident + C vs. voiceless stop + C; and voiceless stop + C vs. voiced stop + C.⁵⁹ The perceptual stridency effect and the perceptual voicing effect in PCCAM generate the following three hypotheses.

(2) Hypotheses for the auditory similarity judgment experiment

- a) Hypothesis for the perceptual stridency effect # 1: [bCV] vs. [buCV] pair (an epenthetic vowel after a voiced stop) is perceptually more similar to each other than [sCV] vs. [suCV] pair (an epenthetic vowel after a voiceless strident).
- b) Hypothesis for the perceptual stridency effect # 2: [pCV] vs. [puCV] pair (an epenthetic vowel after a voiceless stop) is perceptually more similar to each other than [sCV] vs. [suCV] pair (an epenthetic vowel after a voiceless

⁵⁹ Three groups of stimuli were used instead of four because a voiceless strident + C vs. a voiced strident + C pair was not tested (sC vs. zC). This was done because English has no z-initial onset clusters.

T

in

5.

5.

TI

cl

sy

us

th

19

ap

de

As

strident).

- c) Hypothesis for the perceptual voicing effect: [bCV] vs. [buCV] pair (an epenthetic vowel after a voiced stop) is perceptually more similar to each other than [pCV] vs. [puCV] pair (an epenthetic vowel after a voiceless stop).

These three hypotheses were tested through the auditory similarity judgment experiment involving thirty-eight native Korean speakers.

5.2. Experimental Design

5.2.1. Stimuli

The stimuli for the experiment were nonsense words containing legal English consonant clusters and their associated pairs of epenthesized forms. Nonsense words were used to systematically capture a range of phonetic and phonological features in the stimuli. Also, using nonsense words makes it possible to avoid similarity judgments being affected by the participants' familiarity with real words in the stimuli (see Flege, Takagi, and Mann 1996 for the discussion of the influence of lexical familiarity in L2 perception).

English [u] was chosen as an epenthetic vowel, which is assumed to be the closest approximation of the Korean vowel [i] (Kabak 2003: 72). Note that Korean vowel [i] is a default vowel which is employed to repair consonant clusters in English based loanwords. As a single onset and coda, [n] was used, and /i/, /e/, and /u/ were used as the vowel in a

CC

was

list

to p

a w

rec

mic

for

stu

by

of

net

(3)

Th

20

sim

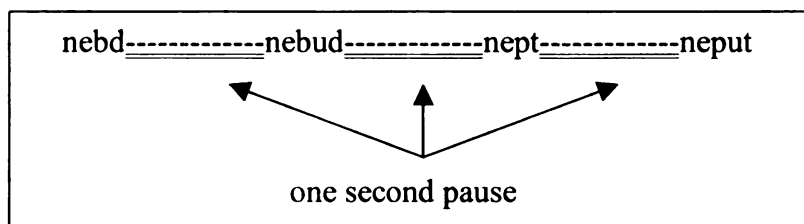
51

CCVC or CVCC sequence.⁶⁰

The stimuli were recorded by a male native speaker of American English who was unaware of the purpose of the experiment. This speaker read a mixed ordered word list composed of nonsense words such as *buREN* and *BREN*. The speaker was instructed to put stress on the capital letter of the word in the wordlist. In this way, the first vowel in a word like *buREN* was realized as high, back, lax vowel, [ʊ].⁶¹ The target words were recorded on a Marantz PMD 201 cassette recorder, using an Audio-Technica ATR35s microphone.

Using Praat (4.3.18), each cluster was paired with its associated epenthesis form (e.g., *nebd-nebud*). Every utterance was followed by a one-second pause. A pilot study showed that a one-second pause was judged to be appropriate for conducting a test by the participants. Then two pairs of sounds were arranged in a sequence for the purpose of similarity comparison among groups (e.g., a voiced pair vs. a voiceless pair = *nebd-nebud* vs. *nept-neput*). The sample template of stimuli is presented in (3).

(3) Sample template of stimuli in the auditory similarity judgment experiment



The order of the two groups in each template was mixed so that the participants'

⁶⁰ Three different vowels were used in the perception experiment to make sure that the perceptual similarity judgment is not biased by the use of one specific vowel.

⁶¹ Thanks to Prof. Preston for providing this idea for the research design.

2

1

v

w

e

st

(4

pa

—

—

—

—

5.2.2

Thirty

were

judgment was not biased by the sequence of the stimuli (e.g., template A: negd-negud vs. nekt-nekut, template B: nikt-nikut vs. nigd vs. nigud).

Summing up the research design, each item in the stimuli contains four words, which are separated by a one-second pause, in a sequence (A-B vs. C-D) and participants were asked whether pair A-B is more similar to each other or pair C-D is more similar to each other. The stimuli by group are given in Table 5.1. The epenthesized form in each stimulus was not presented in Table 5.2 for simplicity's sake. A total 1634 target items (43 pairs x 38 participants) were tested, and five items were not answered by the participants. Thus a total of 1629 were analyzed for the perceptual similarity judgment.

Table 5.2. Stimuli for the auditory similarity judgment test

	voiced stop-strident	voiceless stop-strident	voiced-voiceless
Onset	dren-sten spun-brun brin-stin bren-sken spin-drin smen-glen snun-brun slin-grin	tren-sten spun-prun prin-stin pren-sken spin-trin smen-clen snun-plun slin-crin	dren-pren plin-glin bren-cren clun-drun
Coda	nigd-nisp nisk-nibz nesk-nebz nigd-nist nest-nebd nusp-nugd nubz-nust negd-nesp	nikt-nisp nisk-nips nesk-neps nikt-nist nest-nept nusp-nukt nups-nust nekt-nesp	nikt-nigd nebz-neps nubd-nupt nigz-niks nugd-nukt negd-nekt negz-neks

5.2.2. Procedure

Thirty- eight Korean speakers in Korea participated in the perception experiment. They were attending the same university in Korea at the time of experiment. Many of the

pa

in

hea

(TC

on

unc

pain

unep

that

way

avoi

itself

of the

under

diffic

exper

5.3. R

The re:

C₁: voi

participants in the perception experiment also participated in the production experiment in Chapter 6.

The researcher met with each participant in a quiet room. The participant used headphones (Samsung smh-m150), and the stimuli were played over a desktop computer (TG dreamsys/R3LC, Pentium 4). Participants were instructed to make their judgments on the auditory similarity of two pairs of words, i.e., the similarity between an unepenthesized form and its associated epenthesized form. They were asked to circle the pair that sounded more similar to each other.

The recorded lists were accompanied by an answer sheet presenting only the unepenthesized forms. The epenthesized forms were not presented in the answer sheet so that participants were not aware of the fact that each item was modified in a systematic way by inserting an epenthetic vowel. The type of modification is disguised so as to avoid a situation in which the participants pay too much attention to the epenthetic vowel itself rather than hearing the whole words when making their judgments on the similarity of the pairs. A pretest was given to each participant to check whether the participant understood the test procedure. In the experiment, no participants reported having difficulty in conducting the test. Refer to Appendix A for the answer sheet used in the experiment.

5.3. Results

The results section is composed of two subsections. § 5.3.1 presents the results by type of C₁: voiced stop, voiceless stop, and voiceless strident. The results in this section illustrate

on

th

au

sto

sto

con

in t

coe

5.3

Thi

how

uneq

Fig

C, b

stop

Note

Fig

with

light

only the influence of C_1 in consonant clusters on the auditory similarity judgment. Note that both onset and coda clusters are combined in the results.

Nevertheless, it is conceivable that the identity of C_2 may play a role in the auditory similarity judgment. English onset clusters tested in the experiment include /s/ + stop, /s/ + nasal, /s/ + liquid, and stop + liquid, while coda clusters include stop + /s/, /s/ + stop, stop + stop clusters. Given the various combinations of C_1 and C_2 in English consonant clusters, grouping the results by the feature of C_1 may obscure the effects of C_2 in the auditory similarity judgment. Thus, § 5.3.2 presents the results by type of onset and coda clusters to allow for a more detailed examination of the auditory similarity judgment.

5.3.1. Results by Type of C_1

This section presents auditory similarity judgment results by type of C_1 . It is examined how the manner and voicing feature of C_1 influences Korean speakers' judgment on unpenthesized stimuli and epenthesized stimuli. The results are presented in Figure 5.1. Figure 5.1 contains three groups of bar graphs: a) voiceless strident + C vs. voiced stop + C, b) voiceless strident + C vs. voiceless stop + C, and c) voiceless stop + C vs. voiced stop + C. Two bar graphs in each group make up 100% of the responses for each category. Note also that the number of responses on each category is presented in the table under Figure 5.1.

The first group of bar graphs in Figure 5.1 illustrates the responses on the stimuli with voiceless strident + C vs. voiced stop + C (e.g., nest-nesut vs. nebd-nebud). The light gray bar represents the percentage of the responses on voiceless strident + C and its

associ

respon

another

epenth

respon

to each

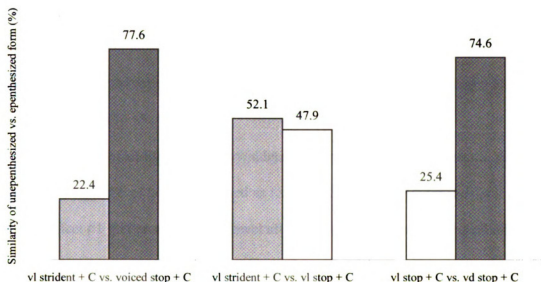
Similarity of unepenthesized vs. epenthesized form (%)

* X-ax

☐	th
☐	th
☐	th
	v
	Gr
vl stri	
vl stri	
vl st	

T

associated epenthesized form (22.4%), and the dark gray bar represents the percentage of responses on voiced stop + C and its associated epenthesized form (77.6%). Stated another way, 22.4% of the participants responded that voiceless strident + C and its epenthesized form are more similar to each other, while 77.6% of the participants responded that voiced stop + C and its associated epenthesized form sound more similar to each other.



* X-axis is abbreviated for simplicity, and the color of the bars represents the following.

☐	the case where voiced stop + C is paired with voiced stop + epenthetic vowel + C
☐	the case where voiceless stop + C is paired with voiceless stop + epenthetic vowel + C
☐	the case where voiceless strident + C is paired with voiceless strident + epenthetic vowel + C

Grouping by type of C ₁	Perceptual similarity judgment			
	A vs. B	A	B	Total (n)
vl strident + C vs. vd stop + C		136 (22.4%)	471 (77.6%)	607
vl strident + C vs. vl stop + C		315 (52.1%)	290 (47.9%)	605
vl stop + C vs. vd stop + C		106 (25.4%)	311 (74.6%)	417
				Significance
				p ≤ .01
				p > .1
				p ≤ .01

Figure 5.1. Results by type of C₁

To see whether the difference in the proportions of the responses is statistically

sig

the

Ko

me

77

sto

gro

voi

C a

epe

expe

strid

obtru

expe

Secor

voice

obstru

25.4°

striden

obtrusi

voicela

⁶² A diff

<http://www>

significant, a difference in proportions test was implemented (refer to the last column of the table under Figure 5.1).⁶² The first group of bar graphs in Figure 5.1 illustrates that Korean speakers dominantly judged that voiced stop + C and its epenthesized form are more similar to each other (22.4%) than voiceless strident + C and its epenthesized form (77.6%). On the other hand, when voiceless strident + C cluster is paired with voiceless stop + C, the similarity judgment is almost evenly distributed as is shown in the second group of bar graphs (52.1% vs. 47.9%). The third group of bar graphs shows that when voiceless stop + C is paired with voiced stop + C, Korean speakers judged voiced stop + C and its epenthesized form as more similar to each other than voiceless stop + C and its epenthesized form (74.6% vs. 25.4%).

The results of auditory similarity judgment by type of C₁ with respect to the experimental hypotheses are summarized as follows. The hypothesis for the perceptual stridency effect #1 that an epenthetic vowel after a voiceless strident is perceptually more obtrusive than an epenthetic vowel after a voiced stop is strongly supported in the experiment (22.4% for voiceless strident + C vs. 77.6% for voiced stop + C, $p \leq .01$). Second, the hypothesis for the perceptual voicing effect that an epenthetic vowel after a voiced obstruent is perceptually less obtrusive than an epenthetic vowel after a voiceless obstruent is also strongly supported in the experiment (74.6% for voiced stop + C vs. 25.4% for voiceless stop + C, $p \leq .01$). However, the hypothesis for the perceptual stridency effect #2 that an epenthetic vowel after a voiceless strident is perceptually more obtrusive than an epenthetic vowel after a voiceless stop is not supported (52.1% for voiceless strident + C vs. 47.9% for a voiceless stop + C cluster, $p > .1$). The question

⁶² A difference in proportions test was implemented using the on-line calculator at <http://www.answersresearch.com/proportions.php>.

that

C is

deta

resu

men

types

C_1 in

stimu

judgr

5.3.2.

5.3.2.1

This se

C_2 on t

clusters

cluster r

pair in F

judgmen

epenthet

epenthet

that arises is why the perceptual stridency effect does not hold when *voiceless strident* + C is paired with *voiceless stop* + C. To answer this question, we need to look at the more detailed auditory similarity judgment test results as the auditory similarity judgment test results by type of C₁ may be obscured by the effects of C₂ in consonant clusters. As was mentioned above, English onset and coda consonant clusters are composed of various types of C₂. Therefore, it is conceivable that the identity of C₂ may interact with that of C₁ in determining the auditory similarity of unpenthesized stimuli and epenthesized stimuli. The following section takes up this issue by discussing the auditory similarity judgment by type of onset and coda clusters.

5.3.2. Results by Type of Onset and Coda Clusters

5.3.2.1. Results by Type of Onset Clusters

This section presents results grouped by type of onset and coda cluster to see the effect of C₂ on the auditory similarity judgment. Figure 5.2 presents results by type of onset clusters. Note that due to space considerations in the figure, only the representative cluster in each cluster category is presented on the X-axis. Thus, in the case of the /st-pr/ pair in Figure 5.2, /s/ + stop (/sp, st, sk/) is paired with voiceless stop + liquid (/pr, tr, kr/).

The results in Figure 5.2 follow the general patterns of the auditory similarity judgment presented in Figure 5.1 (results by type of C₁): Korean speakers judged an epenthetic vowel after a voiceless strident as perceptually more obtrusive than an epenthetic vowel after a stop regardless of cluster type (thus supporting the perceptual

str

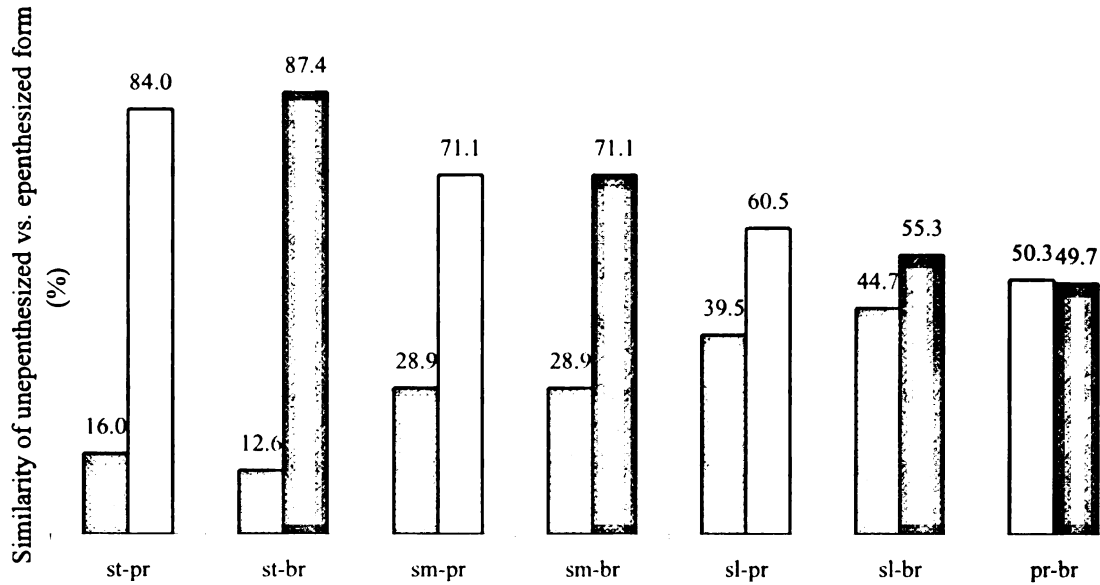
Similarity of unenriched and enriched

* X
☐
☐
☐
☐
Gr

In partic
support t

⁶³ Statist
relatively

stridency effect).



* X-axis is abbreviated for simplicity, and the color of the bars represents the following.

☐	the case where voiced stop + C is paired with voiced stop + epenthetic vowel + C				
☐	the case where voiceless stop + C is paired with voiceless stop + epenthetic vowel + C				
☐	the case where voiceless strident + C is paired with voiceless strident + epenthetic vowel + C				
Grouping by onset clusters		Perceptual similarity judgment			
A vs. B		A	B	Total (n)	Significance
st-pr		30 (16.0%)	158 (84.0%)	188	p≤ .01
st-br		24 (12.6%)	166 (87.4%)	190	p≤ .01
sm-pr		22 (28.9%)	54 (71.1%)	76	p≤ .01
sm-br		22 (28.9%)	54 (71.1%)	76	p≤ .01
sl-pr		15 (39.5%)	23 (60.5%)	38	p≤ .1 ⁶³
sl-br		17 (44.7%)	21 (55.3%)	38	p> .1
pr-br		76 (50.3%)	75 (49.7%)	151	p> .1

Figure 5.2. Results by type of onset clusters

In particular, the first and second groups of the bar graphs (st-br and st-pr) directly support the hypothesized P-map for vowel epenthesis proposed by Steriade (2001b).

⁶³ Statistical significance at $p \leq .1$ level is also used in the perception experiment to allow for a relatively small amount of data in sl-pr and sl-br category ($n = 38$).

The P-map for vowel epenthesis predicts that an epenthetic vowel which breaks down a voiceless strident + obstruent sequence brings about more dissimilarity between unepenthesized stimuli and epenthesized stimuli than an epenthetic vowel in a sequence of obstruent + sonorant (refer to Table 5.1). The first and second groups of bar graphs in Figure 5.2 (st-pr and st-br) illustrate that Korean speakers' judgment on the auditory similarity between unepenthesized stimuli and epenthesized stimuli exactly corresponds with the predictions of the P-map on the perceptibility of an epenthetic vowel. Also note that the difference in the proportions of responses is statistically significant.

In addition, two interesting patterns, believed to demonstrate the effect of C₂ in auditory similarity judgments, are found in Figure 5.2. The first pattern occurs within sC clusters. Interestingly, in sC clusters as the sonority of C₂ in sC clusters⁶⁴ increases (from voiceless stops, nasals, to liquids), the percentage of similarity judgment on sC clusters also increases. In other words, the higher the sonority of C₂ in sC clusters, the more likely the cluster is judged to be similar to its epenthesized output (e.g., the pair /slet-sulet/ is judged as more similar to each other than the pair /smet-sumet/. In turn, the pair /smet-sumet/ is judged as more similar to each other than the pair /sket-suket/). The fact that the different degree of auditory similarity judgment depends on the identity of C₂ in sC clusters clearly demonstrates that the identity of C₂ plays an important role in the auditory similarity judgment.

Thus, the results in Figure 5.2 demonstrate that the perceptual stridency effect can be further elaborated among sC onset clusters. In particular, the following order of the perceptual stridency effect holds: /s/ + stop (an epenthetic vowel is most perceptually obtrusive among sC clusters) < /s/ + nasal < /s/ + liquid (an epenthetic vowel is least

⁶⁴ /s/ is C₁ in this case.

perceptually obtrusive among sC clusters).

The influence of C₂ in auditory similarity judgment is also found when voiceless stop + liquid is paired with voiced stop + liquid. The last pairing of bar graphs in Figure 5.2 indicates that when voiceless stop + liquid is paired with voiced stop + liquid + C, the similarity judgment is almost evenly distributed, which means that no perceptual voicing effect holds in these clusters (50.3% for voiceless stop + liquid vs. 49.7% for voiced stop + liquid).

This finding raises the question why the perceptual voicing effect does not hold between voiceless stop + liquid and voiced stop + liquid. A possible answer to this pattern can be found in Treiman (1983) and Fleischhacker (2005). Treiman (1983) tested native American English speakers' preference for dividing CVC syllables in English, using a novel word-game experiment. In the experiment, Treiman (1983: 61) found that the preferred location for word splitting is at an obstruent-vowel boundary. That is, participants in the experiment divide CVC into C plus VC rather than CV plus C. This indicates that an obstruent-vowel boundary serves as an important perceptual breaking point.

In a similar vein, Fleischhacker (2005: 72-73) proposes a perceptual break theory, which is based on the assumption that the obstruent-vowel juncture serves as a landmark for speech segmentation (as was evidenced in Treiman 1983). She uses the perceptual break theory in accounting for the asymmetric patterns of vowel insertion and consonant deletion in the reduplication and loanword adaptation process. To understand how the perceptual break theory works in Fleischhacker's (2005) analysis, it is important to know the patterns of vowel insertion and consonant deletion she considers.

ana

pro

seg

inse

that

simi

proc

inser

that /

differe

loanw

as *sa*

C₂ in

phenom

of Eng

tasks a

percep

sound

types d

accordi

The patterns of vowel insertion and consonant deletion that Fleischhacker (2005) analyzed in her dissertation are as follows. In reduplication and loanword adaptation processes, obstruent + sonorant onset clusters exceptionally permit deletion of a second segment (or failure to copy the second cluster member in the reduplication process) and insertion of a vowel between the two cluster members. Fleischhacker (2005) proposes that this exceptional behavior of obstruent + sonorant clusters is grounded in a perceptual similarity principle: phonological processes occur more freely when the result of the process sounds quite similar to the original form.

Simply put, the basic idea of Fleischhacker (2005) for the patterns of vowel insertion and consonant deletion in loanword adaptation and reduplication processes is that *pra* sounds more like *pa* or *pira* than *sta* sounds like *sa* or *sita*, and this similarity difference is employed in the loanword adaptation and reduplication processes: in loanword phonology, *pra* is more likely to be adapted as a *pa* or *pira* than *sta* is adapted as *sa* or *sita*.

Fleischhacker (2005) supports her claim about the effect of the identity of C₁ and C₂ in the perceptual similarity between input and modified output by examining linguistic phenomena such as alliteration of Germanic, Middle English, and Old Irish and a corpus of English imperfect puns. She also conducted perception experiments (discrimination tasks and direct similarity judgments). The evidence from linguistic phenomena and perception experiments all converged on the facts that *obstruent* + *sonorant* clusters sound more like their consonant deleted or vowel inserted counterparts than other cluster types do. To summarize, the perceptual break in obstruent + sonorant clusters is exploited, according to Fleischhacker (2005), in vowel insertion and consonant deletion processes:

only obstruent + sonorant clusters allow C₂ deletion and cluster-internal vowel insertion.

Now recall the two interesting patterns found in the auditory similarity judgment when grouped by type of onset clusters (refer to Figure 5.2): a) the effect of C₂ sonority on similarity judgment within sC clusters, and b) no voicing effect when voiceless stop + liquid is paired with voiced stop + liquid.

Here, I add one more pattern; an epenthetic vowel in /s/ + liquid is perceptually more obtrusive than an epenthetic vowel in stop + liquid (refer to the last three groups of bar graphs in Figure 5.2).⁶⁵ Actually, this pattern can be explained by the perceptual stridency effect alone (i.e., vowel insertion after a strident is perceptually more obtrusive than an epenthetic vowel after a stop). However, the perceptual break theory proposed by Fleischhacker (2005) provides an additional phonetic explanation for why an epenthetic vowel is perceptually more obtrusive in /s/ + liquid than in stop + liquid. Thus, I will account for the three patterns referring to the perceptual break theory, which are summarized as follows.

(4) Three patterns of the perceptual similarity judgment to be accounted by the perceptual break theory

- a) As the sonority of the second element in sC clusters increases, the percentage of similarity judgment between an original sC cluster and its epenthesized form also increases (/slet-sulet/ (more similar) < /smet-sumet/ < /stet-sutet/).
- b) When a voiceless strident + liquid cluster is paired with a stop + liquid cluster, a

⁶⁵ The difference in the proportions of responses is not statistically significant in sl (44.7%) -br (55.3%) pair, even though the raw data illustrates the auditory similarity judgment is in the expected direction (i.e., an epenthetic vowel is perceptually more obtrusive in /s/ + liquid). The observation that the difference in the proportions of responses is not statistically significant in this case seems to be due to the relatively small number of total tokens in sl-br pair (n = 38). More observation would be needed to confirm the role of the perceptual stridency effect in this pair.

stop + liquid cluster and its epenthesized forms are more similar to each other than a voiceless strident + liquid cluster and its epenthesized form (pret-puret (more similar) < slet-sulet).

- c) The similarity judgment is almost evenly distributed between a voiceless stop + liquid cluster and a voiced stop + liquid cluster (pret-puret = bret-buret).

These three patterns receive an efficient explanation under the perceptual break theory.

At this point, it is important to note that the perceptual break theory (proposed by Fleischhacker 2005) rests on the assumption that the strength of a perceptual break is enhanced by two factors: a) relatively high-intensity formant structure and b) the presence of a stop closure preceding the onset of formant structure (p. 72).

That a perceptual break is enhanced by relatively high-intensity formant structure accounts for the pattern in (a). Note that C_2 in sC clusters can be a voiceless stop, a nasal, and a liquid, and each C_2 is varied in terms of the intensity formant structure: a liquid has the highest intensity formant structure, followed by a nasal and a voiceless stop.

When we put the different degree of intensity formant structure together with the assumption that a perceptual break is enhanced by relatively high-intensity formant structure, we can see that /s/ + liquid clusters contain a stronger break than /s/ + nasal clusters (due to the fact that liquids have a relatively higher intensity formant structure than nasals). By the same logic, /s/ + nasal contains a stronger perceptual break than /s/ + stop. The effect of C_2 in sC clusters in terms of a perceptual break can also be stated such that the higher the sonority of C_2 in sC clusters, the better the C_2 serves as a perceptual break.

Returning to the pattern (a), it seems that the relative perceptual break difference

resulting from the identity of C₂ nicely explains the perceptual similarity judgment differences among sC clusters. Thus, the perceptual similarity difference between /slet-sulet/ is less likely to be noticed by listeners than the perceptual similarity difference between /smet-sumet/ because liquids provide a better perceptual break (i.e., insertion of a vowel is less obtrusive) than nasals do.

Now we move on to the second pattern found in the auditory similarity judgment test: voiceless stop + liquid was judged to be more similar to its epenthesized form than voiceless strident + liquid is to its epenthesized form. This pattern is accounted for by the second factor (proposed by Fleischhacker 2005) affecting the perceptual break: the presence of a stop closure preceding the onset of formant structure enhances a perceptual break.

The obstruent-vowel juncture is known to be associated with a boost in auditory response relative to an obstruent-obstruent juncture (Wright 2004). This auditory boost effect is maximized “by the presence of stop closure (and thus, silence) preceding vowel onset, and driven down by the presence of frication noise preceding the vowel” (Bladon 1986; Wright 1996 cited in Fleischhacker 2005: 72).

Based on the assumption that the presence of a stop closure preceding the onset of formant structure enhances a perceptual break, we can say that stop + liquid contains a stronger perceptual break than /s/ + liquid due to the presence of a stop closure. This explains the pattern in Korean speakers’ judgment on auditory similarity such that when voiceless strident + liquid is paired with voiceless stop + liquid, voiceless stop + liquid and its epenthesized form are judged to be more similar to each other than voiceless strident + liquid and its epenthesized form (pret-puret (more similar) < slet-sulet). In

other words, the stronger perceptual break in voiceless stop + liquid results in a higher perceptual similarity in voiceless stop + liquid and its epenthesized form.

Lastly, the assumption in the perceptual break theory that a relatively high-intensity formant structure contributes to strengthening a perceptual break provides a possible solution (even though it is only indirect) to the observation that the similarity judgment is almost evenly distributed when voiceless stop + liquid is paired with voiced stop + liquid.

I propose that this may be because the role of a C_2 sonorant receives more weight than that of C_1 obstruent voicing in determining the perceptual similarity between unepenthesized stimuli and epenthesized stimuli. In other words, the relatively high-intensity formant structure of sonorants may neutralize the effect of obstruent voicing when sonorancy and obstruent voicing effect compete with each other as a factor for determining perceptual similarity. Thus, when voiceless stop + liquid is paired with voiced stop + liquid, only the C_2 sonorancy plays a role, not the voicing effect of a C_1 obstruent.

The proposal that voicing effect may be neutralized in the presence of sonorants is also supported by sets of auditory similarity judgments in Figure 5.2 such as /st-pr/ vs. /st-br/; /sm-pr/ vs. /sm-br/; /sl-pr/ vs. /sl-br/. When we assume that voicing of a C_1 obstruent plays a role in these sets, we would expect that voiced stop + liquid and its epenthesized form (i.e., /st-br/) receives a higher similarity judgment than voiceless stop + liquid and its epenthesized form (i.e., /st-pr/). Contrary to this prediction, there is not much difference in the auditory similarity judgment between these two sets of clusters. Furthermore, in the set of /sl-pr/ vs. /sl-br/, /pr/ receives a higher similarity judgment than

/b

ne

Fig

voi

inter

amo

impo

voici

relati

voicin

feature

a cluste

results

present

5.3.2.2.

This se

Recall t

the first

/br/. All this evidence converges on the proposal that the voicing effect induced by C₁ is neutralized in the presence of C₂ sonorants.

To summarize, the auditory similarity judgment in onset clusters (presented in Figure 5.2) supports the perceptual stridency effect that an epenthetic vowel after a voiceless strident is perceptually more obtrusive than an epenthetic vowel after a stop.

In addition, the auditory similarity judgment test results revealed several interesting perceptual similarity judgment differences depending on the context: a) among sC clusters and b) between /s/ + liquid and voiceless stop + liquid. Most importantly, based on the perceptual break theory, I propose that there is no perceptual voicing effect when voiceless stop + liquid is paired with voiced stop + liquid because the relatively high-intensity formant structure of sonorants neutralizes the effect of obstruent voicing in determining the perceptual similarity.

The proposal that the voicing effect is neutralized by the existence of sonorant features in a cluster raises questions such as “What happens when there is no sonorant in a cluster? Is the voicing effect attested then?” The auditory similarity judgment test results on coda clusters provide answers to these questions. The following section presents the auditory similarity judgment results by type of coda clusters.

5.3.2.2. Results by Type of Coda Clusters

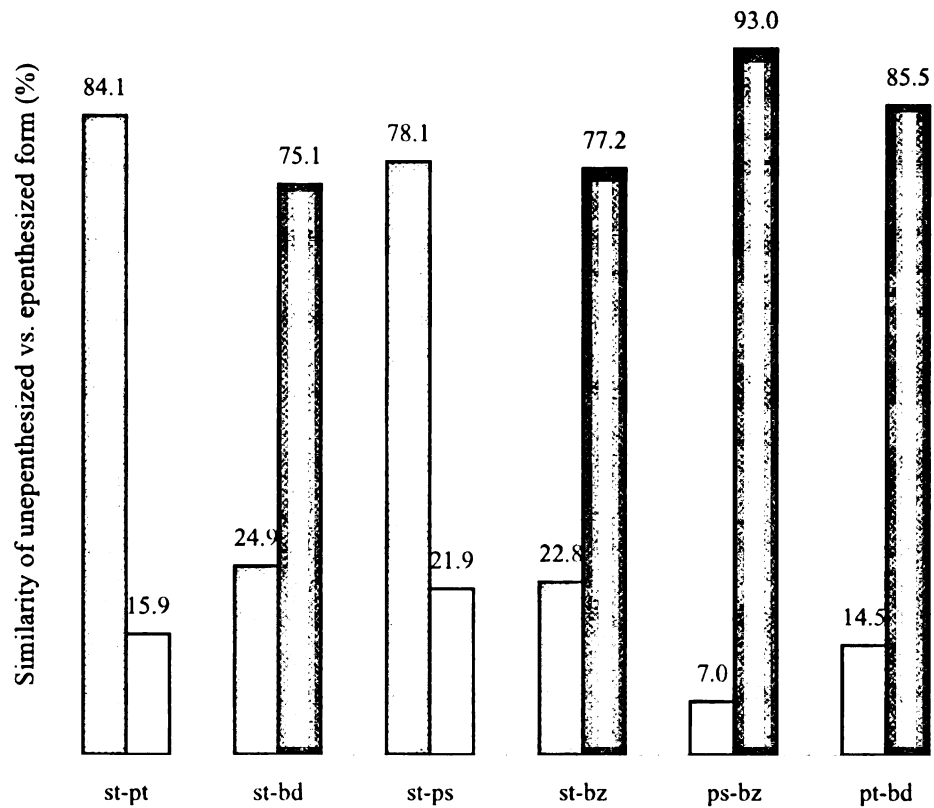
This section presents the auditory similarity judgment results by type of coda clusters. Recall that the notations on the x-axis are the representative case of each cluster. Thus, the first group of bar graphs shows the case where voiceless strident + voiceless stop is

paired with voiceless stop + voiceless stop (/st-pt/).

The results in Figure 5.3 illustrate that when voiceless strident + voiceless stop is paired with voiceless stop + voiceless stop (/st-pt/), Korean speakers judged that voiceless strident + voiceless stop and its epenthesized form are more similar to each other (84.1% for /st/ vs. 15.9% for /pt/). On the other hand, when voiceless strident + voiceless stop is paired with voiced stop + voiced stop (/st-bd/), Korean speakers judged that voiced stop + voiced stop and its epenthesized form are more similar to each other (24.9% for /st/ vs. 75.1% for /bd/). The differences in the proportions of responses are statistically significant in these cases.

When we relate the results in Figure 5.3 to the experimental hypotheses, we can say that the perceptual voicing effect is strongly supported. Note that for pairs of ps-bz and pt-bd, Korean speakers dominantly judged voiced obstruent + voiced obstruent pairs and their epenthesized forms as more similar to each other than their voiceless counterparts (e.g., 7.0% for /ps/ vs. 93.0% for /bz/).

By contrast, the perceptual stridency effect is only partially supported. That is, for pairs of st-bd and st-bz, Korean speakers judged voiced stop + C and its epenthesized form as more similar to each other than its strident counterparts (e.g., 24.9% for /st/ vs. 75.1% for /bd/). However, for pairs of st-pt and st-ps, Korean speakers judged voiceless strident + stop and its epenthesized form (/st/) as more similar to each other than its voiceless stop counterparts (e.g., 84.1% for /st/ vs. 15.9% for /pt/), which contradicts the prediction of the perceptual stridency effect.



* X-axis is abbreviated for simplicity, and the color of the bars represents the following.

☐	the case where voiced stop + C is paired with voiced stop + epenthetic vowel + C			
☐	the case where voiceless stop + C is paired with voiceless stop + epenthetic vowel + C			
☐	the case where voiceless strident + C is paired with voiceless strident + epenthetic vowel + C			
Grouping by coda clusters				
Perceptual similarity judgment				
A vs. B	A	B	Total (n)	Significance
st-pt	159 (84.1%)	30 (15.9%)	189	p≤ .01
st-bd	47 (24.9%)	142 (75.1%)	189	p≤ .01
st-ps	89 (78.1%)	25 (21.9%)	114	p≤ .01
st-bz	26 (22.8%)	88 (77.2%)	114	p≤ .01
ps-bz	8 (7.0%)	106 (93.0%)	114	p≤ .01
pt-bd	22 (14.5%)	130 (85.5%)	152	p≤ .01

Figure 5.3. Results by type of coda clusters

How is this conflicting pattern accounted for while keeping the basic principles of the perceptual stridency effect and perceptual voicing effect? For this purpose, I

intro

relea

saying

dete

voic

eper

voic

cha

pair

tha

tha

asy

eff

sin

pla

me

th

ef

ve

pe

introduce the perceptual release effect. It is very important to note that the perceptual release effect only applies to obstruent + obstruent coda clusters because it goes without saying that consonants of syllable onsets are of necessity released.

The perceptual release effect considers both the identity of C_1 and C_2 in determining the perceptual similarity and specifically states that an epenthetic vowel in voiceless strident + voiceless stop sequences is perceptually less obtrusive than an epenthetic vowel in voiceless stop + voiceless stop sequences because the release of a voiceless strident is similar to the property of a vowel: both a vowel and a strident are characterized by continuous airflow through mouth (thus both are [+continuant]).

This leads to a question about the case where voiceless strident + voiceless stop is paired with voiced stop + voiced stop. Recall that for this pair, Korean speakers judged that voiced stop + voiced stop and its epenthesized form are more similar to each other than voiceless strident + voiceless stop and its epenthesized form. Regarding this asymmetry, I propose that the perceptual voicing effect may override perceptual release effect when these two factors compete with each other in determining perceptual similarity. Thus, for the pair of nest-nesut vs. nept-neput, the perceptual release effect plays a role in determining perceptual similarity and the nest-nesut pair is perceptually more similar than the nept-neput pair. However, for the pair nesk-nesuk vs. nebd-nebud, the perceptual release effect in /s/ + stop sequence competes with the perceptual voicing effect in voiced stop + voiced stop sequence. In this kind of situation, the perceptual voicing effect overrides the perceptual release effect. Accordingly, the pair nebd-nebud is perceptually more similar to each other than the pair nesk-nesuk.

5.4. 3

As a
judgr
achie

(5) G
a

b,

When
the per
no per
liquid
paired
effect b
percept

elabora
stridenc

5.4. Summary of Chapter 5

As a summary of this chapter, let us discuss how the two aims of the auditory similarity judgment experiment (which I presented at the beginning of this chapter) have been achieved. The two goals were as follows.

(5) Goals of the auditory similarity judgment experiment

- a) to test the validity of the two perceptual principles defining consonant cluster markedness in the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM): the perceptual stridency effect and the perceptual voicing effect.
- b) to illustrate the interaction between articulatory and perceptual factors that determine consonant cluster markedness in PCCAM.

When it comes to the first goal, it has been shown that the perceptual stridency effect and the perceptual voicing effect hold true in all the cluster types except in two contexts: a) no perceptual voicing effect when voiceless stop + liquid is paired with voiced stop + liquid and b) no perceptual stridency effect when voiceless strident + voiceless stop is paired with voiceless stop + voiceless obstruent. I have explained no perceptual voicing effect by referring to the perceptual break theory. In addition, I have accounted for no perceptual stridency effect by postulating the perceptual release effect.

It has also been shown that the perceptual stridency effect can be further elaborated among sC onset clusters. In particular, the following order of the perceptual stridency effect holds: /s/ + stop (an epenthetic vowel is most perceptually obtrusive

amc

obtr

This

strid

perc

more

stop

perce

plays

phon

liquic

relati

relati

effect

to the

simila

factor

cluster

< /s/ +

among

among sC clusters) < /s/ + nasal < /s/ + liquid (an epenthetic vowel is least perceptually obtrusive among sC clusters).

The second goal is actually closely related with the first goal of the experiment. This is because the observation that the perceptual voicing effect and the perceptual stridency effect does not hold true in some contexts can be interpreted as evidence that perceptual factors do not play a role in these contexts. That is, if vowel insertion occurs more frequently after a voiced stop in voiced stop + liquid clusters than after a voiceless stop in voiceless stop + liquid clusters, this is not because the epenthetic vowel is perceptually less obtrusive after a voiced stop, but because the articulatory voicing effect plays a role. Stated another way, the perceptual voicing effect cannot be used as a phonetic factor determining the relative markedness relationship between voiced stop + liquid and voiceless stop + liquid.

Thus, the results of the auditory similarity judgment test are reflected in the relative consonant cluster markedness relationships in Table 5.3. Table 5.3 lists the relative consonant cluster markedness relationships which refer to the perceptual voicing effect or perceptual stridency effect as a phonetic factor. We can see that when it comes to the ranking relationships in sl < pl, pr, tr, kl, kr, and pt, kt < bd, gd, the auditory similarity judgment experiment in this chapter directly supports the role of the perceptual factors. Note also that the perceptual stridency effect is further divided among sC onset clusters (i.e., /s/ + stop (an epenthetic vowel is perceptually most obtrusive) < /s/ + nasal < /s/ + liquid), and this observation is reflected in the relative markedness relationships among sC onset clusters.

However, the auditory similarity judgment experiment did not support the

existence of the perceptual stridency or the perceptual voicing effect in pl, pr, tr, kl, kr < bl, br, dr, gl, gr and sp, st, sk < pt, kt. These results are reflected in the phonetic factors defining each markedness relationship in the second column of the table. Thus, the relative markedness relationship in pl, pr, tr, kl, kr < bl, br, dr, gl, gr and sp, st, sk < pt, kt can be understood such that articulatory factors play a major role in defining the relative markedness in these clusters, rather than perceptual factors.

Table 5.3. Consonant cluster markedness relationships vs. auditory similarity judgment

Relative markedness relationship	Phonetic factors	Auditory similarity judgment
Onset		
sp < sm	a. maximal perceptual contrast b. articulatory stridency effect c. perceptual stridency effect (addition)	st-pr (16.0% vs. 84.0%) sm-pr (28.9% vs. 71.1%)
sm < sl	a. maximal perceptual contrast b. perceptual stridency effect (addition)	sm-pr (28.9% vs. 71.1%) sl-pr (39.5% vs. 60.5%)
sl < pl, pr, tr, kl, kr	a. articulatory stridency effect b. perceptual stridency effect c. degree of gestural overlap	sl (39.5%) vs. pr (60.5%)
pl, pr, tr, kl, kr < bl, br, dr, gl, gr	a. articulatory voicing effect (no perceptual voicing effect)	pr (50.3%) vs. br (49.7%)
Coda		
sp, st, sk < pt, kt	a. articulatory stridency effect b. degree of gestural overlap (no perceptual stridency effect)	st (84.1%) vs. pt (15.9%)
pt, kt < bd, gd	a. articulatory voicing effect b. perceptual voicing effect	pt (14.5%) vs. bd (85.5%)

Note: The exceptional cases to the perceptual stridency effect or perceptual voicing effect are highlighted in the table.

CHAPTER 6

KOREAN SPEAKERS' PRODUCTION OF ENGLISH CONSONANT CLUSTERS

This chapter presents a production experiment involving 30 Korean EFL (English as a Foreign Language) students. The production experiment is designed to test the validity of the consonant cluster markedness hierarchy generated by the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM).

Even though much of the empirical data from child language and second language acquisition supports the predictions of PCCAM on the consonant cluster markedness hierarchy (refer to § 4.4 and § 4.5), there are several limitations when relying only on the published data to test the validity of PCCAM.

First, many previous studies used only a subset of English consonant clusters (e.g., Broselow & Finer; Eckman & Iverson 1993). Thus, the predictions of PCCAM can only be partially tested by using previously published data alone. Second, the categories of consonant clusters vary among studies. For instance, PCCAM divides sC consonant clusters into three groups: /s/ + stop (/sp, st, sk/), /s/ + nasal (/sm, sn/) and /s/ + liquid clusters (/sl/), while some studies divide them into /s/ + stop (/sp, st, sk/) and /s/ + sonorant clusters (/sm, sn, sl/) (e.g., Kim 2000). Third, some previous studies only provide an overall error rate for each cluster without discussing what kinds of errors occur in each cluster (e.g., Major 1996). Given that PCCAM makes specific predictions on the site of vowel insertion and consonant deletion, it is difficult to fully confirm the predictions of PCCAM based on the error rate presented in the previous studies because it is possible that other types of errors (e.g., feature change) contribute to the overall error

rate.

chap

vowe

conso

conso

predi

gener

struct

desig

More

conso

exper

§ 6.4 c

6.1. K

This se

presen

and al

only se

allowe

rate.

To improve upon the limitations of using the data from the previous studies, this chapter presents a production experiment specifically designed to test the patterns of vowel insertion and consonant deletion in Korean speakers' production of English consonant clusters. It will be shown that not only the patterns of vowel insertion and consonant deletion but also the relative consonant cluster markedness relationship predicted by the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM) in general holds true in Korean speakers' production of English consonant clusters.

Chapter 6 is organized as follows. § 6.1 presents Korean consonant and syllable structure as a background for the following sections. § 6.2 presents the experimental design for the production experiment. § 6.3 presents results of the production experiment. More specifically, § 6.3.1 and § 6.3.2 present the patterns of vowel insertion and consonant deletion, respectively. § 6.3.3 discusses the results of the production experiment in view of the consonant cluster markedness hierarchy generated by PCCAM. § 6.4 concludes this chapter by summarizing the consonant cluster production results.

6.1. Korean Consonant and Syllable Structure

This section discusses the Korean consonant system and syllable structure. Table 6.1 presents the Korean consonant chart. Korean has 19 consonants in its sound inventory, and all the consonants except /ŋ/ can appear in syllable initial position. On the other hand, only seven consonants (unreleased stops [p̚, t̚, k̚], nasals [m, n, ŋ], and lateral [l]) are allowed to appear in syllable-final position (Lee & Ramsey 2000: 66).

Table 6.1 Korean consonant chart (Sohn 1999: 153)

		Bilabial	Alveo- dental	Palatal	Velar	Glottal
Stop	Lax	p	t	c	k	
	Aspirated	p ^h	t ^h	c ^h	k ^h	
	Tensed	p'	t'	c'	k'	
Fricative	Aspirated		s			h
	Tensed		s'			
Nasal		m	n		ŋ	
Liquid	Lateral or Flap		l			

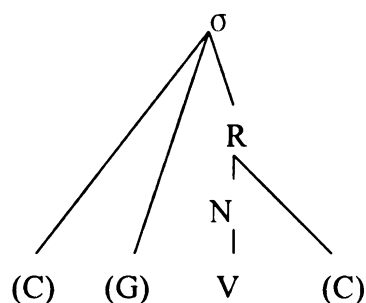
Due to the restrictions on the types of consonants that can appear in syllable-final position, coda-neutralization processes occur in Korean: a) in syllable-final position bilabial, alveolar, and velar obstruents are neutralized to [p', t', k'] respectively and b) palatal and glottal consonants /c, c', c^h, h/ are merged into [t'].

Korean has three types of stops (aspirated, lax, and tensed) as phonemes, and all Korean stops, affricates, and fricatives are voiceless. Voiced stops occur as allophones between sonorants in Korean.

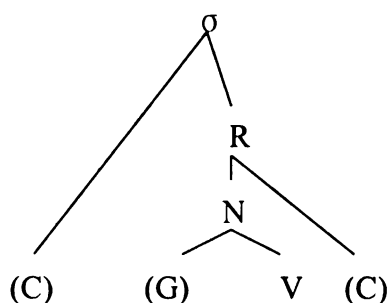
There has been general agreement that Korean syllable structure is composed of [(C)(G)V(C)]_σ (where G is a glide). However, several different views exist on the status of a glide in a syllable. Some researchers treat Korean glides as part of the nucleus (e.g., Kim & Kim 1991; Kim 1990; Sohn 1987), while others regard them as part of an onset cluster (e.g., Ahn 1985; Lee 1991; Lee 1993). The structural representations for these two positions are shown in (1a) and (1b), respectively.

(1) Korean syllable structure (a: glide as a part of onset, b: glide as a part of nucleus)

(a)



(b)



(R = rhyme, N = nucleus, C = consonant, V = vowel G = glide)

Whether Korean syllable structure is represented as (1a) or (1b) is important for those L2 researchers whose analyses are based on sonority hierarchy or typological markedness, since different positions make different predictions on Korean speakers' production of L2 consonant clusters. For instance, if it is assumed that Korean allows a consonant + glide cluster as a possible onset cluster, the prediction would be that Korean speakers will have less difficulty with an L2 consonant + glide cluster.

Chapter 3 examined this controversy while discussing Broselow & Finer (1991) and Eckman & Iverson (1993). Broselow & Finer (1991), who use the syllable representation of (1a) (glide as a part of onset), predict that a voiceless C + glide cluster would be the least difficult English cluster for Korean speakers. Eckman & Iverson (1993) use the syllable representation (1b). Therefore, they predict that a voiceless stop + liquid cluster is easier for Korean speakers than a voiceless stop + glide cluster (see Tak 1997: 26-56 for in-depth discussion of the status of glides in Korean). Due to the controversy over the status of a glide in Korean phonology, PCCAM does not make any predictions on the relative markedness of C + glide clusters.

absen

coda

conse

word

6.2.1

6.2.1

Thirty

time c

Englis

major

from 2

exclud

possib

themse

was co

To summarize, Korean syllables are simpler than those of English due to the absence of consonant clusters. There are no consonant clusters either in syllable onset or coda in Korean (assuming that a glide is part of nucleus). In contrast, English has consonant clusters both in onset and coda position, and inflectional suffixes at the end of words create complex codas with three consonants in a row (e.g., strengths).

6.2. Experimental Design

6.2.1. Participants

Thirty Korean EFL students in Korea participated in the production experiment. At the time of the experiment, all participants were taking the same courses provided by the English department, and so most of the participants were English language and literature majors or minors. They were in the second or third year of college, and their ages ranged from 21 to 27.

Participants who lived in English-speaking countries more than three months were excluded from the study to keep the amount of English exposure as homogeneous as possible. In the questionnaire about their English proficiency, participants marked themselves as “low middle” or “middle.” More detailed information on the participants was collected through a questionnaire (Appendix B).

6.2.2. Materials

All two-member onset and coda consonant clusters in English are tested in the production experiment.⁶⁶ However, for the purpose of this dissertation, if any of the consonants in a cluster are not allowed in Korean, the cluster was excluded from the data coding. Thus, no English consonant clusters with /f, v, ʃ, θ, ð/ (these segments are not allowed in Korean) were coded in the present study. This was done based on the justification that the errors resulting from the fact that Korean does not have the segments in its sound inventory (/f/ → [p], /v/ → [b], /θ/ → [s], etc) should not be coded as an error resulting from consonant cluster production: those errors can also occur in a singleton onset or coda (see § 4.3.1 for difficulty of data coding and interpretation resulting from the inclusion of these segments in the production experiment). Refer to Appendix C for the target words and fillers used in the present study.

The types of onset and coda clusters analyzed for this dissertation are presented in Table 6.2. The structures of the words are CCVC and CVCC for target words, and VCCV for fillers. A filler was inserted after every three words. Thus, in the production experiment, each participant was asked to read 350 words (262 target words + 88 fillers).

The target words analyzed for this dissertation are 164 tokens for each participant. The reason that only 164 words (out of 262 target items) were analyzed for this dissertation is that English consonant clusters with /f, v, ʃ, θ, ð/ (these segments are not allowed in Korean) were not included in the present study.

⁶⁶ /lʒ/ and /r/ + C clusters (e.g., /rp/, /rt/, /rd/, /rn/) were not included in the production experiment by accident.

words

of the

phonos

that an

instan

2001).

sheer a

feature

that K

detecte

"Sever
were ke

Table 6.2. Target consonant clusters in the production experiment

Cluster Type	
Onset Clusters	a. strident + voiceless stop: /sp, st, sk/
	b. voiceless strident + nasal: /sm, sn/
	c. voiceless strident + liquid: /sl/
	d. voiceless stop + liquid: /pr, pl, tr, kr, kl/
	e. voiced stop + liquid: /br, bl, dr, gr, gl/
Coda Clusters	a. voiceless stop + voiceless strident: /ps, ts, ks/
	b. voiceless strident + voiceless stop: /sp, st, sk/
	c. voiceless stop + voiceless stop: /pt, kt/
	d. voiced stop + voiced strident: /bz, gz/
	e. voiced strident + voiced stop: /zd/
	f. voiced stop + voiced stop: /bd, gd/
	g. sonorant + voiceless strident: /ls, ns/
	h. sonorant + voiced strident: /nz/
	i. sonorant + voiceless stop: /lp, lt, lk, mp, nt/
	j. sonorant + voiced stop: /lb, ld, nd/
	k. liquid + nasal: /lm/

Note also that in the current study, nonsense words were used instead of English words to ensure that errors based on the knowledge of the words were avoided.⁶⁷ It is one of the assumptions in this production experiment that if we are to investigate L2 learners' phonological competence, then we need to isolate a learner's competence in phonology so that any contribution from their competence in the other components of language, for instance knowledge in lexicon, does not affect the phonological competence (Tench 2001). For this reason, the experiment in this dissertation used nonsense words as stimuli.

Even though there is disadvantage of using nonsense words, for instance, the sheer artificiality of the event, we can capture a range of phonetic and phonological features in a systematic way by using nonsense words. It is also assumed in this study that Korean speakers' intuition about the English consonant clusters is more easily detected when most of the words are unknown words.

⁶⁷ Several of the nonsense words actually exist in English (e.g., tent, slot). However, these words were kept in the word list to maintain the consistency of the word list.

Nonsense words are constructed using /t/ and /n/ as a singleton onset and coda, and /e/ and /o/ are used as a nucleus. Mid vowels were chosen instead of high vowels because a high vowel such as /i/ and /u/ may trigger a similar type of vowel insertion in consonant clusters to maintain vowel harmony (e.g., Kim (2000) reported that Korean speakers used /i/, /u/, and /i/ to resolve consonant clusters in English).⁶⁸

6.2.3. Procedure

In the practice session, participants were given a chance to read over the wordlists before being recorded so that they could familiarize themselves with the nonsense words. Each participant was instructed to read the word as it was spelled and was not allowed to ask questions on the pronunciation of a specific word. Participants were encouraged to read the nonsense words as if they were real English words.⁶⁹

Each word was presented to the participants on index cards one at a time, but a mixed order was kept to make sure there was no order effect. The participants read each word embedded in a carrier phrase “I am saying___” The carrier phrase was used to provide a consistent phonological environment before each target word. The experiment was recorded on a Marantz PMD 201 cassette recorder, using an Audio-Technica ATR35s microphone.

⁶⁸ Vowel harmony was rigidly observed in old Korean, but rules have significantly weakened in modern Korean. However, vowel harmony still continues to play an important role in modern Korean in onomatopoeic and mimetic words.

⁶⁹ It was made clear in the practice session that the nonsense words were made up based on real English words so that participants could read the words as if they were real English words.

6.2.4. Measures

The recordings of Korean speakers' production of English consonant clusters were digitized through USB pre, and waveforms and spectrograms were generated by Praat (version 4.2.04). The following section discusses measures for identifying vowel insertion and consonant deletion in the waveforms and spectrograms.

Vowel insertion

An epenthetic vowel was identified according to three criteria. First, vocalic complex patterns in the waveform were identified. Then, it was checked whether the vocalic complex patterns in the waveform were accompanied by a F1 and F2 formant in the spectrogram. Lastly, the portion of a presumed epenthetic vowel was examined in an expanded view. If there were periodic wave shapes, the token was finalized as having an epenthetic vowel. This procedure is described from Figure 6.1 to Figure 6.4.

Figure 6.1 shows the waveform and spectrogram of /dret/ with an epenthetic vowel between /d/ and /r/. The vocalic complex patterns in the waveforms and corresponding F1 and F2 formants in the spectrogram are indicated by a circle.

Figure 6.2 is the expanded waveform of the vocalic complex patterns in Figure 6.1. The periodic movement of the waveform makes it clear that the token contains an epenthetic vowel.

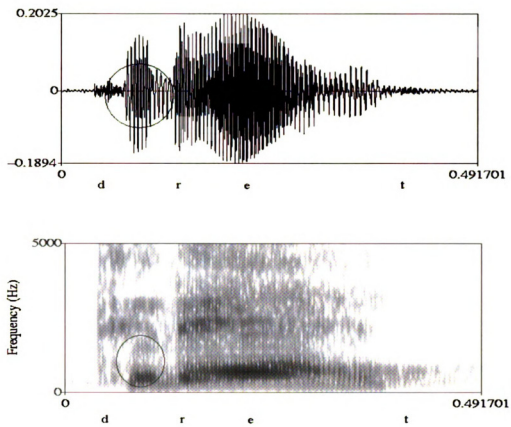


Figure 6.1. Waveform and spectrogram of /dret/ with an epenthetic vowel

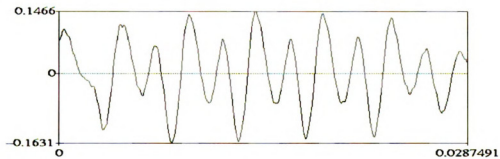


Figure 6.2. Extended waveform of an epenthetic vowel in Figure 6.1

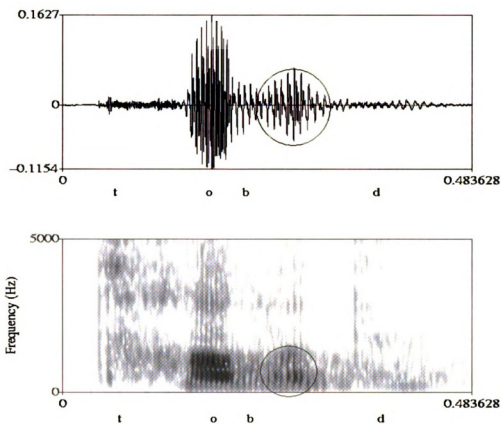


Figure 6.3. Waveform and spectrogram of /tobd/ with an epenthetic vowel

Figure 6.3 presents the waveforms and spectrograms of the target word /tobd/ with an epenthetic vowel after C_1 , while Figure 6. 4 presents the waveforms and spectrograms of the target word /tesk/ with an epenthetic vowel after C_2 . The epenthetic vowels are indicated by circles in the waveforms and spectrograms.

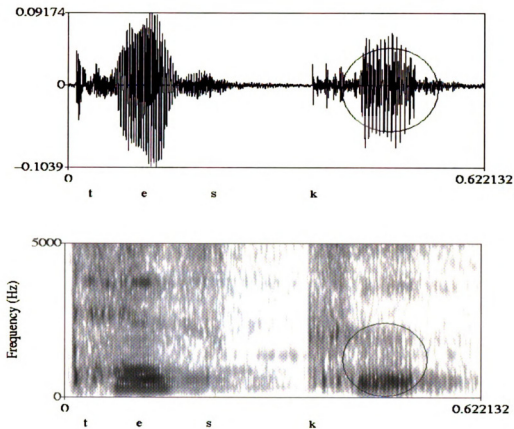


Figure 6.4. Waveform and spectrogram of /tesk/ with an epenthetic vowel

Consonant deletion

Consonant deletion is identified when there is no audible sound and no visible signs of a corresponding consonant in the spectrogram.⁷⁰ Figure 6.5 presents a waveform and spectrogram of the target word /tesp/ with final stop deletion.

⁷⁰ It is conceivable that even though a segment is not audible (and thus interpreted as deletion), the articulatory gestures of a segment may still exist in the speaker's production (as Browman & Goldstein 1989 views final /t, d/ deletion in English; refer to § 4.1.1 for the discussion of /t, d/ deletion under Articulatory Phonology). There is still controversy over what should be considered as segment deletion. However, in the present study, if there is no audible indication of a segment, it was coded as segment deletion.

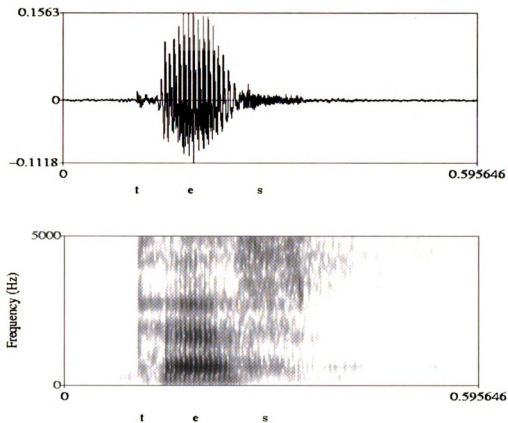


Figure 6.5. Waveform and spectrogram of /tesp/ with final stop deletion

Non-applicable tokens

Non-applicable tokens mainly indicate the cases where the participants produced target words with errors other than vowel insertion, consonant deletion, and metathesis. The cases of non-applicable tokens are listed as follows.

- a. use of a non-targeted consonant (e.g., /negz/ → [neɪz])
- b. consonant insertion (/nobz/ → [nɒlbz])
- c. missing data

Missing data are the cases where the target words were lost during the recording or digitizing process. There are 9 instances of missing tokens out of 4920 target words.

Note that changing /l/ to a flap (/blet/ → [bret]) or /r/ to [l] are counted as target-like productions in order to allow for the phonetic differences between English liquids (/l/ and /r/) and Korean liquids (/l/ and flap). By the same logic, devoicing of voiced obstruents is counted as target-like production since voiced obstruents are not allowed in Korean.

Table 6.3 presents the number of valid and invalid tokens in the production experiment.

Table 6.3. Number of tokens by cluster type

Cluster type	Representative	Valid	Non-applicable	Total
Onset				
voiceless strident + voiceless stop	sp	359	1	360
voiceless strident + nasal	sm	237	3	240
voiceless strident + liquid	sl	120	0	120
voiceless stop + liquid	pr	596	4	600
voiced stop + liquid	br	597	3	600
Coda				
voiced stop + voiced stop	bd	227	13	240
voiced stop + voiced strident	bz	228	12	240
voiceless stop + voiceless stop	pt	235	5	240
voiceless stop + voiceless strident	ps	348	12	360
voiceless strident + voiceless stop	sp	354	6	360
voiced strident + voiced stop	zd	114	6	120
liquid + voiced stop	ld	238	2	240
liquid + voiceless stop	lt	360	0	360
liquid + voiceless strident	ls	120	0	120
liquid + nasal	lm	110	10	120
nasal + voiced stop	nd	115	5	120
nasal + voiceless stop	nt	230	10	240
nasal + voiced strident	nz	114	6	120
nasal + voiceless strident	ns	116	4	120
Total		4818	102	4920

The first column in Table 6.3 presents the type of onset and coda clusters analyzed in the present study. The second column in the table presents a representative cluster in each cluster type. Thus, /br/ indicates voiced stop + liquid clusters which include /br, bl, dr, gr, gl/. This is done to save some space in the figures and tables presented in the following sections. Table 6.3 shows that among the total 4920 tokens, 102 tokens were classified as non-applicable. Thus, 4818 tokens were analyzed for the patterns of vowel insertion and consonant deletion.

6.3. Results

This section provides results from the production experiment. The results section is composed of three parts. § 6.3.1 presents the patterns of vowel insertion depending on the identity of C_1 or C_2 (C_1 = the first segment of a cluster, C_2 = the second segment of a cluster). § 6.3.2 discusses patterns of consonant deletion depending on the cluster type. It will be shown that the predictions of the *Phonetically Based Consonant Acquisition Model* (PCCAM) regarding the pattern of vowel insertion and consonant deletion hold true in Korean speakers' production of English consonant clusters. Lastly, § 6.3.3 discusses the determination of onset and coda consonant cluster markedness based on the number of vowel insertions, consonant deletions, and metathesis involved in each consonant cluster.

6.3.1. Patterns of Vowel Insertion

6.3.1.1. Vowel Insertion in Onset Consonant Clusters

This section presents the patterns of vowel insertion in Korean speakers' production of English consonant clusters. Table 6.4 illustrates the number of vowel insertions by type of C_1 in onset clusters. To see whether the difference in the number of vowel insertions depending on the identity of C_1 is statistically significant, a pairwise Pearson chi-square test was implemented. Chi-square results were computed using SPSS (version 12.0), and significance at $p \leq .05$ was assumed. Statistically significant pairs are indicated by * at $p \leq .05$ level, ** at $p \leq .01$ level, and *** at $p \leq .001$ level.

Table 6.4 Epenthesis by type of C_1 in onset clusters

Type of C_1 (onset)	Epenthesis (n)	Total (n)	Epenthesis/Total (%)
voiced stop	76	597	12.7
voiceless stop	19	596	3.2
voiceless strident	11	716	1.5
Total (n)	106	1909	5.6
Pairwise comparison	Pearson Chi-Square (χ^2)		p-value
voiceless stop < voiced stop	37.055		$p = .000^{***}$
voiceless strident < voiceless stop	3.971		$p = .046^*$

The first column in Table 6.4 lists three different types of C_1 in onset clusters, and the second column lists the number of vowel insertions after each type of C_1 . The third column indicates the total number of tokens in each C_1 category, and the last column indicates the percentage of epenthesis in each C_1 category. Table 6.4 illustrates that 76 instances (out of 597 tokens) of vowel insertion occurred after voiced stops, while 11 instances (out of 716 tokens) of vowel insertion occurred after voiceless stridents. We can see that an epenthetic vowel occurred most often after voiced stops (12.7%) and occurred least often after voiceless stridents (1.5%). Pairwise chi-square tests revealed

that differences in the number of vowel insertions by type of C_1 are statistically significant.

Thus, the predictions of PCCAM are supported in the patterns of vowel insertion in onset consonant clusters. Vowel insertion is most likely to occur after voiced stops due to the articulatory voicing effect and is least likely to occur after voiceless stridents due to articulatory and perceptual stridency effects.

6.3.1.2. Vowel Insertion in Coda Consonant Clusters

Note that two types of epenthetic vowel occur in coda clusters. Epenthesis by type of C_1 is the case where an epenthetic vowel occurs after C_1 ($VCC \rightarrow VCvC$). Epenthesis by type of C_2 is the case where an epenthetic vowel occurs after C_2 in coda clusters ($VCC \rightarrow VCCv$). Table 6.5 illustrates the patterns of vowel insertion in coda clusters by type of C_1 .

Table 6.5. Epenthesis by type of C_1 in coda clusters

Type of C_1 (coda)	Epenthesis (n)	Total (n)	Epenthesis/Total (%)
voiced stop	80	455	17.6
voiceless stop	5	583	0.9
voiced strident	1	114	0.9
voiceless strident	0	354	0.0
liquid	31	828	3.7
nasal	0	575	0.0
Total (n)	117	2909	4.0
Pairwise Comparison	Pearson Chi-Square (χ^2)		p-value
voiceless stop < voiced stop	95.079		p = .000***
voiceless strident < voiceless stop	3.052		p = .081 / p = .163 (FET) ⁷¹
voiceless strident < voiced strident	3.112		p = .078 / p = .244 (FET)

⁷¹ The P-value from the Chi-square is reliable when no more than 20% of the expected counts are less than 5, and all individual expected counts are 1 or greater. However, there is no lower bound on the amount of data that is needed for Fisher's Exact Test. Thus the P-value from Fisher's exact test (FET) is reported along with the P-value from Chi-square test for the variables with more than 20% of the expected counts are less than 5.

The patterns of vowel insertion after C_1 in coda clusters are consistent with those in onset clusters. Vowel insertion is most likely to occur after voiced stops and least likely to occur after voiceless stridents. The pairwise chi-square tests reveal that differences in the number of vowel insertions in each category are statistically significant only between voiced stops and voiceless stops.⁷² That is, the raw data in Table 6.5 illustrates differences in the number of vowel insertions in the expected direction (e.g., voiceless strident (0.0%) < voiceless stop (0.9%); voiceless strident (0.0%) < voiced strident (0.9%), but these differences are not statistically significant.

The observation that the differences in the number of vowel insertions after C_1 are not statistically significant for some cases seems to be due to the small number of vowel insertions after C_1 . Note that there were only five instances of vowel insertion after voiceless stops and one instance after voiced stridents. Moreover, Table 6.6 below shows that vowel insertion after C_2 demonstrates the same pattern as that after C_1 , and the statistical significance is robust in this case. This allows us to interpret the statistically non-significant results in Table 6.5 as a consequence of the low rate of vowel insertion after C_1 .

Lastly, liquid + nasal (/lm/) involves a relatively frequent vowel insertion. It seems that liquid + nasal clusters cause the most difficulties for Korean speakers' production: Korean speakers either insert a vowel or delete one of the consonants in liquid + nasal clusters.

Let us move onto the patterns of vowel insertion after C_2 in coda clusters. Table

⁷² A statistical analysis for vowel insertion after nasals and liquids was not conducted because vowel insertion after nasals and liquids is associated with a single coda cluster, liquid + nasal. Thus no comparable pair is established.

6.6 illustrates the patterns of vowel insertion by type of C₂ in coda clusters.

Table 6.6. Epenthesis by type of C₂ in coda clusters

Type of Coda Cluster C ₂	Epenthesis (n)	Total (n)	Epenthesis/Total (%)
voiced stop	158	694	22.8
voiceless stop	141	1179	12.0
voiced strident	61	342	17.8
voiceless strident	50	584	8.6
nasal	3	110	2.7
Total (n)	413	2909	14.2
Pairwise Comparison	Pearson Chi-Square (χ^2)		p-value
voiceless stop < voiced stop	38.034		p = .000***
voiceless strident < voiceless stop	4.667		p = .031*
voiceless strident < voiced strident	17.586		p = .000***

When we compare the epenthesis rate in Table 6.5 and Table 6.6, we can notice that vowel insertion occurred more often after C₂ than after C₁. 117 instances of vowel insertion occurred after C₁ (VCC → VCvC) out of 2909 tokens, while 413 instances of vowel insertion occurred after C₂ (VCC → VCCv).

The observation that vowel insertion occurs more often after C₂ than after C₁ in coda clusters is consistent with the proposal in this dissertation that an epenthetic vowel is formed as a consequence of an acoustic release aimed at providing acoustic cues to a segment (refer to § 4.1.5).

Note that English requires an obligatory *non-release* of a consonant when it occurs after [-cons] and before [-syll] (e.g., infect, rapt; the relevant consonant is italicized), while consonant release is preferred when it occurs after [+cons] and before a pause (e.g., He's wasp; why not ask?) (Selkirk 1982: 374-375).⁷³ For instance, for stop + stop coda clusters such as /pt/, the acoustic release of /p/ is not allowed, while the acoustic release of /t/ is preferred when it occurs before a pause. When combining the

⁷³ Consonant cluster production before a pause is relevant here because in the production experiment, Korean speakers used a carrier phrases "I am saying _____".

two observations: a) acoustic release is attributable to vowel insertion and b) C_2 release is preferred before a pause, the more frequent occurrences of vowel insertion after C_2 are accounted for successfully.

We can see in Table 6.6 that among obstruents, vowel insertion is most likely to occur after voiced stops and least likely to occur after voiceless stridents. This pattern is consistent with the patterns of vowel insertion in other environments (epenthesis after C_1 in onset and coda clusters). As was noticed in the patterns of vowel insertion after C_1 , vowel insertion after nasals results from liquid + nasal sequences, the most difficult type of coda clusters for Korean speakers. Pairwise chi-square tests revealed that differences in the number of vowel insertions in each category are statistically significant.

In summary, the predictions of PCCAM on the patterns of vowel insertion and consonant deletion are supported in Korean speakers' production of English clusters: Among obstruents, vowel insertion is most likely to occur after voiced stops and least likely to occur after voiceless stridents. Among sonorants, there was no vowel insertion after nasals or liquids unless it was a liquid + nasal cluster: liquid + nasal involves vowel insertion both after liquids and after nasals (vowel insertion after liquids is dominant).

6.3.2. Patterns of Consonant Deletion

This section discusses the patterns of consonant deletion in coda clusters. When it comes to onset clusters, there were only three tokens of liquid deletion (two instances in voiced stop + liquid, and one in voiceless stop + liquid). The number of liquid deletions in onset clusters is too small to deserve any analysis. Table 6.7 illustrates the patterns of

consonant deletion in coda clusters.

Table 6.7. Consonant deletion by type of coda clusters

Cluster type	Stop	Strident	Liquid	Nasal	Cluster type Total (n)
bd	10	NA	NA	NA	227
bz	3	1	NA	NA	228
pt	5	NA	NA	NA	235
ps	2	3	NA	NA	348
sp	27	1	NA	NA	354
zd	27	5	NA	NA	114
ld	26	NA	0	NA	238
lt	10	NA	11	NA	360
ls	NA	1	0	NA	120
lm	NA	NA	39	9	110
nd	3	NA	NA	0	115
nt	6	NA	NA	5	230
nz	NA	0	NA	0	114
ns	NA	1	NA	0	116
Deletion (n)	119	12	50	14	2909

Note: NA = non-applicable

Referring to the last line of Table 6.7, which presents the total number of deletions by consonant type, we can see that stop deletion is dominant (119 instances of stop deletion out of 2909 tokens). In contrast, there were only 12 instances of strident deletion.

PCCAM predicts that stops are the target of deletion, in particular by the principle of perceptual saliency (Côté 2000; refer to § 4.2.4 for the principle of perceptual saliency).

Also note that in Table 6.7 most liquid and nasal deletions occurred in liquid + nasal sequences, which indicates the relative markedness of liquid + nasal coda clusters.

6.3.3. Consonant Cluster Markedness Hierarchy

This section discusses the relative markedness of onset and coda clusters. Table 6.8

presents a summary of the relative consonant cluster markedness and phonetic factors associated with each ranking relationship so that it serves as background for the discussion in the following section.

Table 6.8. Consonant cluster markedness hierarchy in PCCAM

Relative markedness relationship	Phonetic factors
Onset (less marked < more marked)	
sp, st, sk < sm, sn	a. maximal perceptual contrast b. articulatory stridency effect c. perceptual stridency effect
sm, sn < sl	a. maximal perceptual contrast b. perceptual stridency effect
sl < pl, pr, tr, kl, kr	a. articulatory stridency effect b. perceptual stridency effect c. degree of gestural overlap
pl, pr, tr, kl, kr < bl, br, dr, gl, gr	a. articulatory voicing effect
Coda	
sonorant + C	
ls, ns < lz, nz	a. articulatory voicing effect
ls, ns < lp, lt, lk, np, nt, mp	a. articulatory stridency effect b. principle of perceptual saliency c. degree of gestural overlap
lp, lt, lk, np, nt, mp < lb, ld, nd	a. articulatory voicing effect
lp, lt, lk, np, nt, mp < lm	a. maximal perceptual contrast
obstruent + obstruent	
ps, ts, ks < sp, st, sk	a. principle of perceptual saliency b. degree of gestural overlap c. articulatory release effect (ps, ts, ks < pt, kt)
sp, st, sk < pt, kt	a. articulatory stridency effect b. degree of gestural overlap
pt, kt < bd, gd	a. articulatory voicing effect b. perceptual voicing effect
bz, gz < zd	a. principle of perceptual saliency b. degree of gestural overlap c. articulatory release effect (bz, gz, zd < bd, gd)
zd < bd, gd	a. degree of gestural overlap

Recall that coda clusters are divided into two groups to ensure a systematic comparison between clusters: sonorant + C and obstruent + obstruent. Among obstruent +

obstruent clusters, it was predicted that voiceless stop + voiceless strident is least marked, while voiced stop + voiced stop is most marked. Among sonorant + C clusters, sonorant + voiceless strident is least marked and sonorant + voiced stop and liquid + nasal are most marked (no ranking relationship between sonorant + voiced stop and liquid + nasal).

The following presents whether the predictions of PCCAM on the relative markedness of English onset and coda clusters hold true for Korean speakers' production of English consonant clusters.

6.3.3.1. Relative Markedness of Onset Consonant Clusters

Table 6.9 illustrates the relative markedness of English onset consonant clusters. The first line of Table 6.9 presents the type of English onset clusters tested in the production experiment. The second line illustrates the total number of errors among the total valid tokens. We can see that only two errors occurred in /s/ + stop among 359 valid tokens.

Table 6.9. Errors in onset clusters

Cluster type	sp	sm	sl	pr	br
Total Errors /n	2/359	1/237	8/120	20/596	78/597
% error	0.6	0.4	6.7	3.4	13.1
Epenthesis(CCV → CvCV)	2	1	8	19	76
Deletion (CCV → CV)	0	0	0	1	2
Pairwise comparison	Pearson Chi-Square (χ^2)		p-value		
sp < sm	0.052		p = .819 / p = 1.000 (FET)		
sm < sl	12.642		p = .000*** / p = .001*** (FET)		
sl < pr	2.914		p = .088 / p = .116 (FET)		
pr < br	37.293		p = .000***		

Note: n = valid tokens in each category

The third line of Table 6.9 illustrates the percentage of errors in each category; thus, the

error rate amounts to 0.6% for /s/ + stop. The fourth and fifth lines provide the number of vowel insertions and consonant deletions by cluster type, respectively. We can see that vowel insertion occurred most frequently in voiced stop + liquid cluster (76 instances out of 597 tokens) among onset clusters.

/s/ + stop (less marked) < /s/ + nasal (more marked)

Table 6.9 shows that /s/ + nasal induced the least errors (0.4%) followed by /s/ + stop (0.6%). This may give the impression that the prediction of PCCAM does not hold for the markedness ranking relationship between voiceless strident + voiceless stop and voiceless strident + nasal: PCCAM predicts /sp, st, sk/ is less marked than /sm, sn/.

However, it should be noted that the number of errors that occurred in these two clusters is very low (only two errors in /s/ + stop and one error in /s/ + nasal), and a pairwise chi-square test reveals that the error rate difference in /s/ + stop and /s/ + nasal is not statistically significant ($\chi^2 = 0.052$, $df = 1$, $p = .819$ / $p = 1.000$ (FET)). The error rate in /s/ + stop and /s/ + nasal basically indicate that there is no ranking relationship established between these two clusters for the Korean speakers tested in the present study.

I would like to suggest that the statistically non-significant ranking relationship between /s/ + stop and /s/ + nasal reflects the dynamic nature of language acquisition. That is, the Korean speakers who participated in the present study have been learning English at least for six years (this information was gathered through questionnaire); therefore, these learners may already have passed the stage where the relative markedness of /s/ + nasal clusters imposes more difficulty in their production. This account does not seem unreasonable especially given that it is not difficult to find empirical data

demonstrating the ranking relationship between /s/ + stop and /s/ + nasal in child language acquisition (e.g., Kirk & Demuth 2003). I think we need to test second language learners at their early initial state to establish the ranking relationship between /s/ + stop and /s/ + nasal.

/s/ + nasal < /s/ + liquid

When it comes to the error rate between /s/ + nasal and /s/ + liquid, a quite large difference was found: /s/ + nasal (0.4%) and /s/ + liquid (6.7%). This difference is statistically significant ($\chi^2 = 12.642$, $df = 1$, $p = .000***$ / $p = .001***$ (FET)).

/s/ + liquid < voiceless stop + liquid

That the error rate in /s/ + liquid (6.7%) is higher than the error rate in voiceless stop + liquid (3.4%) actually goes against the predictions of PCCAM, even though the difference in error rate in these two clusters is not statistically significant ($\chi^2 = 2.914$, $df = 1$, $p = .088$ / $p = .116$ (FET)).

voiceless stop + liquid < voiced stop + liquid

As was predicted by PCCAM, voiced stop + liquid induced the most errors in onset clusters, and the difference in the error rate between voiced stop + liquid and voiceless stop + liquid is statistically significant ($\chi^2 = 37.293$, $df = 1$, $p = .000***$).

In summary, the predictions of PCCAM on the relative markedness of onset consonant clusters were supported except for the ranking relationships between a) /s/ + stop vs. /s/ + nasal and b) /s/ + liquid and voiceless stop + liquid. However, it should be

recalled that the error rate differences in these clusters (/s/ + stop vs. /s/ + nasal and /s/ + liquid and voiceless stop + liquid) are not statistically significant, which seem to reflect the dynamic nature of language acquisition.

6.3.3.2. Relative Markedness of Coda Consonant Clusters

This section discusses the relative markedness of coda consonant clusters. Table 6.10 and Table 6.11 present error rates in obstruent + obstruent and sonorant + C clusters, respectively. Since vowel insertion occurred either after C₁ or after C₂ in coda clusters, epenthesis errors are divided into two groups: VCC → VCvC and VCC → VCCv (lower case v indicates an epenthetic vowel).

Deletion errors are divided into four groups based on the type of consonant deleted: stop deletion, strident deletion, liquid deletion, and nasal deletion. The fifth line of Table 6.10 presents metathesis that occurred in obstruent + obstruent clusters. The metathesis found in the production experiment either changed strident + stop to stop + strident, or vice versa.

Note in Table 6.10 that among obstruent + obstruent clusters, voiced stop + voiced stop (55.1%) induced the most errors, while voiceless stop + voiceless strident (12.4%) induced the least errors.

Table 6.10. Errors in coda clusters (obstruent + obstruent)

		bd	bz	pt	ps	sp	zd
Total errors /n		125/227	89/228	41/235	43/348	97/354	61/114
% error		55.1	39.0	17.4	12.4	27.4	53.5
Epenthesis	VCC → VCvC	43	37	3	2	0	1
	VCC → VCCv	72	48	33	33	60	26
Deletion	VCC → VC						
	stop	10	3	5	2	27	27
	strident	NA	1	NA	3	1	5
	liquid	NA	NA	NA	NA	NA	NA
	nasal	NA	NA	NA	NA	NA	NA
Metathesis		0	0	0	3	9	2
Cluster type		Pearson Chi-Square (χ^2)		P-value			
ps < sp		24.879		p = .000***			
sp < pt		7.801		p = .005**			
pt < bd		70.983		p = .000***			
bz < zd		6.466		p = .011*			
zd < bd		0.074		p = .785			

Note: NA = non-applicable

The metathesis that occurred in obstruent + obstruent clusters raises a very interesting pattern. Fourteen instances of metathesis occurred in obstruent + obstruent clusters, and 11 out of 14 changed strident + stop to stop + strident: /sp/ → [ps], /zd/ → [dz] (and accordingly, three instances of metathesis changed stop + strident to strident + stop: /ps/ → [sp]). Note also that the same pattern of metathesis is attested in child language acquisition. Kirk & Demuth (2003: 446) found that 23% of children's attempted /s/ + stop clusters resulted in metathesis (e.g., wasp → waps), and six of the nine children in their study produced this type of metathesis.

That the metathesis error mainly renders /s/ + stop to stop + /s/ in child speech is consistent with cross-linguistic patterns of metathesis (Hume 2002). Regarding the directional asymmetry of metathesis, Hume (2002) proposes that metathesis of /s/ + stop to stop + /s/ seems to result in improved acoustic/auditory cues in the output: a stop flanked by a vowel is perceptually more salient. Note that Humes' account of the patterns

0

(

(

f

p

of metathesis is parallel with the principle of perceptual saliency proposed by Côté (2000): a stop wants to be flanked by a vowel to maintain its perceptibility.

Metathesis from /s/ + stop to stop + /s/ sequences may also be interpreted as children's preference for less overlapped consonant sequences, as was proposed by the degree of gestural overlap parameter in PCCAM. Note that stop + /s/ sequences are less overlapped than /s/ + stop sequences. Then, metathesis of /s/ + stop to stop + /s/ changes a more marked consonant sequence to a less marked one in terms of the degree of gestural overlap parameter.

Let us move on to the error rates in sonorant + C clusters. Table 6.11 presents the errors in these clusters.

Table 6.11. Errors in coda clusters (sonorant + C)

		ld, nd	lt, nt	ls, ns	nz	lm
Total errors /n		89/353	80/590	19/236	13/114	82/110
% error		25.2	13.6	8.1	11.4	74.5
Epenthesis	VCC → VCvC	0	0	0	0	31
	VCC → VCCv	60	48	17	13	3
Deletion	VCC → VC					
	stop deletion	29	16	NA	NA	NA
	strident deletion	NA	NA	2	0	NA
	liquid deletion	0	11	0	0	39
	nasal deletion	0	5	0	0	9
Cluster type		Pearson Chi-Square (χ^2)		P-value		
ls, ns < lt, nt		4.849		p = .028*		
lt, nt < ld, nd		20.389		p = .000***		
ls, ns < nz		1.040		p = .308		
lt, nt < lm		193.869		p = .000***		

Note: NA = non-applicable

Table 6.11 illustrates that liquid + nasal induced the most errors (74.5%), which is followed by sonorant + voiced stop (25.2%). Table 6.12 presents the predictions of PCCAM on the relative markedness of coda clusters together with the error rates in

Korean speakers' production.

Table 6.12 Coda cluster markedness hierarchy in PCCAM vs. Korean speakers' production

less marked < more marked; = markedness is the same; // no markedness relationship	
Coda	ls, ns (8.1%) < nz (11.4%) // lt, nt (13.6%) < ld, nd (25.2%) // lm
Sonorant + C	(74.5%)
Coda	ps (12.4%) < sp (27.4%) < pt (17.4%) // bz (39%) < zd (53.5%) <
Obstruent + Obstruent	bd (55.1%)

Note: Bold type indicates where raw data do not support the prediction of PCCAM.

Comparing the coda consonant markedness hierarchy in PCCAM with the error rates, we can see that the raw error rates support the predictions of PCCAM except for one case: the ranking relationship between /s/ + voiceless stop and voiceless stop + voiceless stop. PCCAM predicts that voiceless stop + voiceless stop cluster is more marked than /s/ + voiceless stop, yet it turned out that /s/ + voiceless stop (27.4%) induced more errors than voiceless stop + voiceless stop (17.4%). How can we account for this conflicting pattern?

Actually, previous studies also reported conflicting results on the relative markedness between voiceless stop + voiceless stop (/pt, kt/) and voiceless strident + voiceless stop (/sp, st, sk/). Hansen (2000) found that Vietnamese learners of English produced voiceless strident + voiceless stop accurately more often than voiceless stop + voiceless stop which is consistent with the predictions of PCCAM on the markedness relationship between these two types of clusters.

In contrast, Hansen's (2001) study reported that Mandarin speakers of English demonstrate an opposite pattern: voiceless stop + voiceless stop is produced accurately

more frequently than voiceless strident + voiceless stop. At this point I cannot provide any account for why /s/ + stop coda clusters induced more errors than voiceless stop + voiceless stop clusters in Korean speakers' production. It seems apparent that the phonetic factors identified in PCCAM alone cannot explain this pattern. I leave this issue for future research.

When it comes to the statistical significance of error rate differences in each cluster, 7 instances turned out to be statistically significant among 9 instances of pairwise chi-square comparisons in coda clusters. There were two cases where the raw data in Table 6.10 and Table 6.11 showed differences in the expected direction, but the differences are not statistically significant: a) $zd < bd$ and b) $ls, ns < nz$. The error rate differences in these pairs are consistent with the predictions of PCCAM, but more observations would be needed to provide confirmation since error rate differences in these clusters are not statistically significant.

Table 6.13 summarizes the results in Korean speakers' production of English consonant clusters.

Table 6.13. Summary: Korean speakers' production of English consonant clusters

Relative markedness relationship	Phonetic factors	Production results	Significance
Onset			
sp, st, sk < sm, sn	a. maximal perceptual contrast b. articulatory stridency effect c. perceptual stridency effect	sp, st, sk (0.6%) < sm, sn (0.4%)	p = .819
sm, sn < sl	a. maximal perceptual contrast b. perceptual stridency effect	sm, sn (0.4%) < sl (6.7%)	p = .000***
sl < pl, pr, tr, kl, kr	a. articulatory stridency effect b. perceptual stridency effect c. degree of gestural overlap	sl (6.7%) < pl, pr, tr, kl, kr (3.4%)	p = .088
pl, pr, tr, kl, kr < bl, br, dr, gl, gr	a. articulatory voicing effect	pl, pr, tr, kl, kr (3.4%) < bl, br, dr, gl, gr (13.1%)	p = .000***
Coda			
ls, ns < lz, nz	a. articulatory voicing effect	ls, ns (8. %) < nz (11.4%)	p = .308
ls, ns < lp, lt, lk, np, nt, mp	a. articulatory stridency effect b. principle of perceptual saliency c. degree of gestural overlap	ls, ns (8.1%) < lp, lt, lk, np, nt, mp (13.6%)	p = .028*
lp, lt, lk, np, nt, mp < lb, ld, nd	a. articulatory voicing effect	lp, lt, lk, np, nt, mp (13.6%) < lb, ld, nd (25.2%)	p = .000***
lp, lt, lk, np, nt, mp < lm	a. maximal perceptual contrast a. principle of perceptual saliency b. degree of gestural overlap c. articulatory release effect	lp, lt, lk, np, nt, mp (13.6%) < lm (74.5%) ps, ts, ks (12.4%) < sp, st, sk (27.4%)	p = .000*** p = .000***
ps, ts, ks < sp, st, sk	a. articulatory stridency effect b. degree of gestural overlap	sp, st, sk (27.4%) < pt, kt (17.4%)	p = .005**
sp, st, sk < pt, kt	a. articulatory stridency effect b. degree of gestural overlap	pt, kt (17.4%) < bd, gd (55.1%)	p = .000***
pt, kt < bd, gd	a. articulatory voicing effect	bz, gz (39%) < zd (53.5%)	p = .011*
bz, gz < zd	a. principle of perceptual saliency b. degree of gestural overlap c. articulatory release effect	zd (53.5%) < bd, gd (55.1%)	p = .785
zd < bd, gd	a. degree of gestural overlap		

The first column of Table 6.13 presents the relative markedness relationships in PCCAM, and the second column lists the phonetic factors defining each markedness relationship. For instance, we can see that the relative markedness relationship between voiceless stop + voiceless stop and voiced stop + voiced stop (pt, kt < bd, gd) is defined by the articulatory voicing effect and perceptual voicing effect. Even though the present study did not test the weighting of each phonetic factor in defining consonant cluster markedness, it is certainly conceivable that when L2 learners produce a certain type of consonant clusters, particular factors play a more important role than other factors associated with the consonant sequence (or vice versa).

It also possible that factors such as frequency of English consonant clusters and syllable position (i.e., onset vs. coda clusters) influence Korean speakers' production of English consonant clusters. To understand the complex array of various factors associated with consonant sequences and the systematicity of vowel insertion and consonant deletion in Korean speakers' production, it is necessary to use multivariate analysis that refers to statistical procedures involving more than one variable (e.g., logistic regression, multiple linear regression, and factor analysis). Of course, we need to use quantifiable variables to implement multivariate analysis. I believe that multivariate analysis provides deeper understanding of L2 learners' performance on non-native consonant clusters. I leave this for future research.

The third column presents the results from Korean speakers' production of English consonant clusters. The shaded areas indicate the three cases (out of 13) where the predictions of PCCAM were not supported in Korean speakers' production. The last column lists the statistical significance in each ranking relation. The shaded areas indicate

that the difference in error rates is in the expected direction, yet the difference is not statistically significant.

Note in Table 6.13 that among the three cases where the predictions of PCCAM seem not to be supported, the differences in error rates were not statistically significant for two cases (i.e., /s/ + stop < /s/ + nasal and /s/ + liquid < voiceless stop + liquid). I think it is rather rash to say that PCCAM fails to predict correct ranking relationships in these clusters because there were no statistical significances in error rates in these clusters. It seems that these statistically non-significant ranking relationships reflect the dynamic nature of language acquisition. That is, the statistically non-significant cases demonstrate that Korean speakers in the present study passed the stage where the relative markedness of certain types of consonant clusters imposes more difficulty in their production. For instance, the error rate difference in Korean speakers' production of /s/ + liquid and voiceless stop + liquid is not statistically significant because Korean speakers in the present study have already passed the stage where the relative markedness of voiceless stop + liquid clusters impose more difficulty in their production.

6.4. Summary of Chapter 6

In this chapter, the validity of the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM) was examined in Korean speakers' production of English consonant clusters.

§ 6.3.1 and § 6.3.2 discussed the patterns of vowel insertion and consonant deletion. As was predicted by articulatory and perceptual voicing effects and articulatory

and perceptual stridency effects in PCCAM, vowel insertion is most likely to occur after voiced stops and least likely to occur after voiceless stridents. When it comes to consonant deletion, stops are the major target of deletion.

§ 6.3.3 and § 6.3.4 discussed overall onset and coda consonant cluster markedness based on the number of vowel insertions, consonant deletion, and metathesis. Out of 13 relative markedness relationships in PCCAM, 8 ranking relationships support the relative markedness hierarchy in PCCAM, and the error rate differences in each ranking relationship are statistically significant. There were two cases where the raw data illustrate error rate differences in the expected direction, but the differences are not statistically significant (sonorant + voiceless strident vs. sonorant + voiced strident, and voiced strident + voiced stop vs. voiced stop + voiced stop).

Lastly, there were three cases where the predictions of PCCAM were not supported (/s/ + stop vs. /s/ + nasal, /s/ + liquid vs. voiceless stop + liquid, and voiceless strident + voiceless stop vs. voiceless stop + voiceless stop). Yet, it should be noted that the error rate differences between /s/ + stop vs. /s/ + nasal and /s/ + liquid vs. voiceless stop + liquid are not statistically significant. Thus, I think the real problematic case to the prediction of PCCAM is the relative markedness relationship between /s/ + voiceless stop and voiceless stop + voiceless stop. The error rate difference in these clusters is opposite to the prediction of PCCAM, and the error rate difference is also statistically significant. It may be that PCCAM needs to incorporate more phonetic factors to explain the special status of voiceless strident + voiceless stop coda clusters.

Th

Ch

In

PC

aut

the

clu

Ch

the

fact

PC

L2

obs

CHAPTER 7

FORMAL ANALYSIS

The previous two chapters examined the validity of the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM) through the perception and production experiments. In particular, Chapter 5 provided empirical evidence for the two perceptual factors in PCCAM, the perceptual stridency effect and the perceptual voicing effect, by testing the auditory similarity between unepenthesized stimuli and epenthesized stimuli. Chapter 6 then tested the predictions of PCCAM on the relative markedness of English consonant clusters by examining Korean speakers' production of English consonant clusters.

In this chapter, I turn to the question of how the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM) is formalized into a grammar. More specifically, the question to be dealt with in this chapter is “How are the articulatory and perceptual factors in PCCAM formalized into an interlanguage grammar?”⁷⁴

There are several difficulties in implementing this task. First of all, recall that PCCAM crucially uses phonetic (articulatory and perceptual) factors in constructing an L2 consonant cluster acquisition model. The difficulty of translating phonetic observations into phonological constraints (or rules) is well expressed by Hayes (1999: 5).

What is crucial here (and recognized in earlier work) is that a research result in phonetics is not the same thing as a phonological constraint. To go from one to the other is to bridge a large gap . . . In many cases, the phonetic research that explains the phonological pattern has been done very well and is quite

⁷⁴ The term, Interlanguage, will be defined in § 7.1.

convincing; it is the question of how to incorporate it into a formal phonology that is difficult.

To incorporate phonetic evidence into a formal phonological theory, the formal phonological theory should allow us to do this. As will be shown in § 7.2.3 in the case of nasal place assimilation, a rule based approach (e.g., the Sound Pattern of English by Chomsky & Halle 1968) fails to provide a framework where rules that embody phonetic explanation are formalized. Accordingly, a rule-based approach generally lacks a deeper explanation of a phonological phenomenon at hand. Another difficulty with translating phonetic observations to phonological constraints is that *phonology* is characteristically categorical and far less variable, while *phonetics* involves gradient and variable phenomena.

The second issue arises from the type of data that this dissertation deals with: language acquisition data. Modeling language learning (in this case, phonological learning) is a challenging task, and an array of questions has to be answered: what kind of information is accessible to the learner from the environment, what is the initial state of language learning, what kinds of language-specific principles must be learned, and so on.

The situation becomes even more complicated in second language learning. For second language learning, not only the properties of universal principles, but also the properties of learners' native language and the properties of a target language need to be taken into account.

In this dissertation, I will not deal with all the questions related with phonological learning. However, in this chapter I will show that the intrinsic structure of Optimality Theory (Prince & Smolensky 1993) makes it eminently suitable for the two

aforementioned issues: a) translating phonetic information into phonological constraints and b) modeling phonological acquisition.

For instance, when it comes to the issue of translating phonetic evidence to phonological constraints, OT opens up a new way to incorporate phonetically-based constraints into a grammar. OT separates markedness constraints that embody a particular form of phonetic difficulty from the issue of how and whether the difficulty is avoided. This basic mechanism of OT provides an efficient way for phonetic observation to be encoded as a form of ranked, conflicting constraints in the system. I will demonstrate this in § 7.2.3 with the case of the nasal place assimilation and word-final obstruent devoicing process.

Furthermore, the two basic assumptions in OT that constraints are universal, and that particular rankings of these universal constraints determine language-specific phonological patterns have an immediate consequence for acquisition: they restrict the learners' search space. According to OT, children acquire the phonology of an ambient language by establishing the ranking of the *predetermined* set of universal constraints. The view of language acquisition under OT frees language learners from the burden of deducing some sets of language-specific phonological constraints from the ambient language, and consequently, OT provides simpler accounts for phonological acquisition.

The issues of translating phonetic observation and modeling phonological acquisition will be further elaborated in § 7.2. The discussion of these two issues will serve as a background for the formal analyses implemented in this chapter.

This chapter is organized as follows. § 7.1 discusses the term, *Interlanguage*, which is widely used in second language acquisition research to describe the unique

patterns of the second language learners' production of a target language. § 7.2 presents the main tenets of Optimality Theory and how phonological acquisition is understood under OT. § 7.3 provides specific constraints which embody phonetic factors in PCCAM. § 7.4 provides formal analyses of the patterns of vowel insertion and consonant deletion found in Korean speakers' production of English consonant clusters using the constraints proposed in § 7.3.

7.1. Interlanguage Phonology

Several early L2 researchers proposed that second language learners have their own linguistic system (e.g., Corder 1971; Nemser 1971; Selinker 1969, 1972).⁷⁵

For instance, Corder (1971) labeled the grammar of second language learners as *idiosyncratic dialect* in the sense that “it is regular, systematic, meaningful, i.e. it has a grammar, and is, in principle, describable in terms of a set of rules, some sub-set of which is a sub-set of the rules of the target social dialect” (p. 151). On the other hand, Nemser (1971) described the structure employed by the L2 learner attempting to utilize the target language as an “approximative and deviant linguistic system” (p. 116). Selinker (1969, 1972) defines *Interlanguage* as follows.

Since we can observe that these two sets of utterances [utterances by target language learners and those by second language learners] are not identical, then in the making of constructs relevant to a theory of second-language learning, one would be completely justified in hypothesizing, perhaps even compelled to

⁷⁵ The three researchers use different labels for the grammar of a second language learner, and they motivated their proposals on theoretical grounds rather than through empirical studies.

hypothesize, the existence of a separate linguistic system based on the observable output which results from a learners' attempted production of a TL form. This linguistic system we will call 'interlanguage' (IL). One of the main points of this paper is the assumption that predictions of behavioral events in a theory of second language learning should be primarily concerned with the linguistic shapes of the utterances produced in ILs (Selinker 1972: 214).

As can be observed from the above statement from Selinker (1972), the basic motivation for hypothesizing that L2 learners construct their own grammar is that non-native speech often demonstrates systematic patterns which are not directly attributable either to a target language (L2) or to a learners' native language (L1).

Given the concept of IL, the question that arises is whether the basic mechanisms governing IL are similar in important ways to those of a first language. In this dissertation, I propose that the universal principles governing the grammar of the first language also govern the grammar of IL. That is, there exist universal principles that shape any human language, and IL, as an independent linguistic system of a learner, is governed by these universal principles.

Taking an example from the consonant cluster acquisition, this dissertation proposes that the principles governing the acquisition of L2 phonotactics are grounded in the universal aspects of speech production and perception shared by human beings. More specifically, this dissertation identifies a) articulatory difficulty inherent in a certain segment and segmental sequences and b) perceptual similarity differences between the modified form and its unmodified counterpart as two major forces shaping phonological constraints in consonant cluster acquisition. It is assumed that the same phonological constraints apply to the grammar of IL.

This dissertation has empirically tested the existence of universal principles

governing the acquisition of L2 phonotactics by constructing the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM). It has also tested PCCAM through Korean speakers' production of English consonant clusters. Recall from § 4.4 and § 4.5 that most of the predictions made by PCCAM hold true in previously published L1 and L2 acquisition data.

The remaining task is to formalize these universal principles into linguistic constraints under Optimality Theory. This task is undertaken in the following sections.

7.2. Optimality Theory

7.2.1. Overview

Basic tenets of OT

Optimality Theory (Prince & Smolensky 1993) views human languages as a ranked system of conflicting universal constraints. All constraints in OT are determined by Universal Grammar, but the rankings of the constraints are language specific. Also, constraints are violable, but the violation must be minimal: violation of a higher-ranked constraint incurs more cost than a lower-ranked constraint. Therefore, the candidate with least costly violation is chosen as an optimal output. The process of output candidate generation and evaluation is presented schematically in Figure 7.1.

Figure 7.1 illustrates that, for a given input, an infinite set of output candidates is generated by the grammar (GEN). The issue of which structural change is optimal is determined by the relative ranking of the constraints, in this case $C_a \gg C_b \gg C_n$. Each of

the constraints in the system requires some aspects of grammatical output forms, and the constraint may be either satisfied or violated by an output form. Simply put, under the OT system, the grammar evaluates the output candidates with ranked constraints and selects the optimal candidate. The optimal candidate is the actual output.

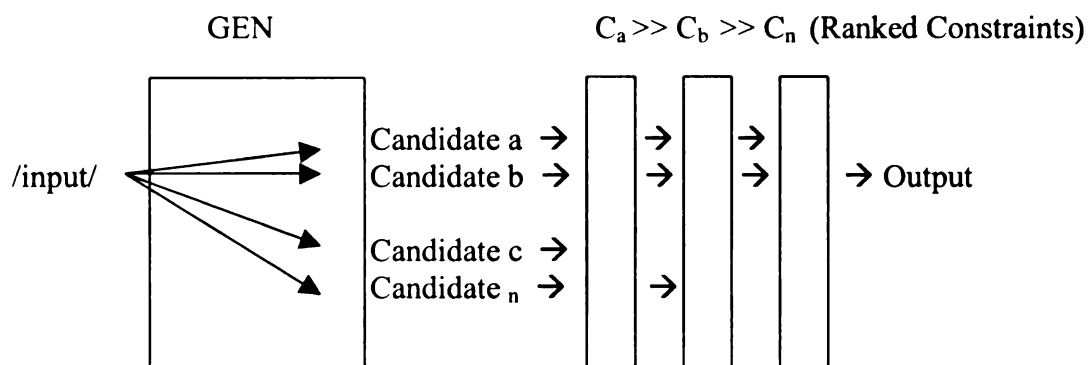


Figure 7.1. Candidate generation and evaluation

An evaluation of output candidates by a set of ranked constraints is displayed in a *Tableau*. For instance, tableau (1) shows four hypothetical output candidates (candidate A, B, C, and D) in competition, and their relative well-formedness is measured by four ranked constraints (constraint 1, 2, 3, and 4). The first vertical column lists four output candidates, and the first horizontal row lists four constraints. Note that constraints are listed in a descending ranking from left to right. Since constraints are listed in a descending ranking, we can interpret Tableau (1) such that constraint 1 outranks constraint 2, constraint 3, and constraint 4, while constraint 2 outranks constraint 3 and constraint 4. The constraint ranking hierarchy is also expressed in the form of constraint 1 >> constraint 2 >> constraint 3, constraint 4.

(1) Tableau for constraint 1 >> constraint 2 >> constraint 3 >> constraint 4

/input/	constraint 1	constraint 2	constraint 3	constraint 4
☞ candidate <i>A</i>			*	****
candidate <i>B</i>	*!			
candidate <i>C</i>		*!		
candidate <i>D</i>			**!	

The asterisk * in the tableau indicates that the constraint heading the column is violated. The number of asterisks expresses how many times the violation occurs. In Tableau (1), candidate B violates constraint 1 once, and candidate D violates constraint 3 twice. The exclamation mark ! indicates that the candidate makes a critical violation of a constraint so that it cannot be an optimal candidate anymore. The columns after the exclamation mark are shaded in gray. The optimal candidate is marked by the pointing index finger ☞.

The dotted line between constraint 3 and 4 indicates that ranking between them cannot be determined with the available data.

Interaction between markedness and faithfulness constraints

When it comes to the constraints involved in the selectional mechanism, OT recognizes two types of constraints: *Markedness constraints* and *Faithfulness constraints*. An optimal candidate is determined by the relative ranking of Markedness constraints with respect to Faithfulness constraints. Interactions between Markedness and Faithfulness constraints also model the extent to which marked structures of certain kinds are allowed in a language.

Markedness constraints require that output forms meet criteria of structural well-formedness, and these well-formedness constraints are built upon the notion of markedness.

Unmarked values are cross-linguistically preferred and basic in all grammars, while marked values are cross-linguistically avoided and used by grammars only to create contrast (Kager 1999: 3). For instance, voiceless obstruents are cross-linguistically preferred to voiced obstruents, and oral vowels are preferred to nasal vowels. When it comes to syllable structures, simple onsets and codas are preferred to complex onsets and codas.

Based on this markedness relation in syllable margins, two markedness constraints are generated: *COMPLEX^{ons} and *COMPLEX^{cod}

(2) Markedness constraints (Kager 1999: 97)

a) *COMPLEX^{ons} ('Onsets are simple')

b) *COMPLEX^{cod} ('Codas are simple')

Along with the markedness constraints, OT recognizes faithfulness constraints. Faithfulness constraints basically require that surface forms match specific properties of other forms, for instance, their lexical input. The current dominating view of faithfulness is Correspondence Theory (McCarthy 1995, McCarthy and Prince 1995).

Correspondence is a relation between two structures such as input-output (or base-reduplicant), where candidate outputs are subject to evaluation together with the correspondent input (or base) (McCarthy 1995: 14). Correspondence constraints militate against divergences of input and output along one dimension,⁷⁶ and thus their effect is to prohibit changes in mapping from inputs to outputs (or from base to reduplicant). Examples of correspondence constraints regarding complex margins are as follows.

⁷⁶ Correspondence constraint: the output equals the input for some property *P* (Kager 1999: 67).

(3) Faithfulness constraints (Kager 1999: 101-102)

MAX-IO: Input segments must have output correspondents. ('No deletion')

DEP-IO: Output segments must have input correspondents. ('No epenthesis')

MAX-IO requires that, for every output segment, there is an input segment corresponding to it: deletion of a segment is not allowed. On the other hand, DEP-IO is violated if the output has a segment that lacks a correspondent in the input: insertion of a segment in the output is not allowed.

Note that faithfulness constraints and markedness constraints are inherently conflicting in the sense that the former preserves lexical properties that the latter may ban on the surface. Stated another way, whenever some lexical contrast is being preserved by keeping faithfulness constraints, there will be some cost associated in terms of markedness.

Whether a language has marked structures or not depends on the relative ranking of markedness constraints and faithfulness constraints: higher ranking of faithfulness constraints over markedness constraints generate marked structures, while higher ranking of markedness constraints over faithfulness constraints generate unmarked structures. Comparison between syllable structures of Korean and English shows how the conflicting relation between markedness and faithfulness constraints generates different syllable structures.

English and Korean present two distinct cases of syllable margins. English allows complex margins in any of the positions, generating V, VC, CV, CVC, CVCC, CCV, CCVC, CVCC, CCCVC, and CVCCC. Note that in English the markedness constraint *COMPLEX is ranked lower than faithfulness constraints. Therefore, with respect to OT

constraints, the greater variety of English syllable structures is made possible at the cost of violating markedness constraint *COMPLEX.

In contrast, Korean ranks markedness constraint *COMPLEX higher than faithfulness constraints, limiting possible syllable types (syllable types in Korean include V, VC, CV, C(G)VC) and using deletion and epenthesis to resolve complex margins. English loanwords in Korean demonstrate a good example of how epenthesis and deletion are employed because *COMPLEX is ranked higher than MAX-IO or DEP-IO in Korean phonology. English loanwords with complex margins are adapted as either vowel insertion (4a) or consonant deletion (4b).

(4) English loanwords with complex margins

a. *grass* [kɪres] vowel insertion

b. *mark* [makɪ] /r/ deletion

To summarize, OT models linguistic well-formedness by a set of conflicting constraints: markedness constraints and faithfulness constraints. Markedness constraints express universal preferences for certain types of structure (e.g., simple onset vs. complex onset), while faithfulness constraints require that surface forms match specific properties of other forms (e.g., input vs. output correspondence).

As was demonstrated by the comparison between syllable structures of Korean and English, under OT, language-specific phonological patterns are attributed to particular rankings of universally determined constraints, not to the introduction of special language-specific constraints. In Korean *COMPLEX dominates MAX-IO and DEP-IO, while in English *COMPLEX is dominated by MAX-IO and DEP-IO. This language-specific ranking generates different syllable structures in the two languages.

7.2.2. Phonetically-Based Constraints in Optimality Theory

This section discusses how OT improves the way that phonetically-based constraints are expressed in a formal system. In so doing, I will first illustrate how the speakers' knowledge of perceptual saliency differences in different contexts (i.e., P-map; Steriade 2001a, 2001b) is incorporated into Optimality Theory. In the OT system, the P-map serves as a mechanism that relates rankings between correspondence constraints to perceived differences of similarity degree. It is important to understand this procedure because I will use the same basic mechanism when translating articulatory and perceptual factors defining consonant cluster markedness in the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM) into OT constraints.

I will directly compare the formalization of nasal place assimilation under a rule-based approach and that of OT to show the advantage of an OT analysis over a rule-based approach. The comparison of nasal place assimilation formalizations in a rule-based approach and OT will illustrate that the inherent structure of OT enables us to provide a deeper explanation for the given phonological pattern. Of course, this is made possible in the OT system, because OT provides a way to incorporate a phonetic observation into the system.

Phonetically-based constraints: word-final obstruent devoicing

Recall from § 4.2.1 that the P-map is a mental representation of the degree of distinctiveness of different contrasts in various positions.

The P-map is 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 in others. One function of the P-map is to identify the margins of articulatory freedom for the speaker: the regions of relative safety within which he may deviate from established pronunciation norms while minimizing the risk of being "found out". Thus, if an innovative speaker contemplates articulatory simplification in a VC₁C₂V cluster, it is the P-map that will identify the optimal target of simplification: the consonant whose modification is least likely to be detected by his interlocutors (Steriade 2001a: 19).

The basic idea of the P-map is that the optimal output would be the one which changes the input minimally when complying with the phonotactic constraints. Following this reasoning, other alternative outputs are avoided because they result in more drastic input-output differences.

The P-map is incorporated into Optimality Theory by projecting and ranking correspondence constraints in the system. That is, based on the perceptual similarity between the input and output form, the P-map projects faithfulness constraints and determines their ranking. For instance, if a contrast x-y/_A is more perceptible (or distinctive) than a contrast x-z/_A, then the P-map's effect on the grammar will rank the correspondence constraint x-y/_A higher than correspondence x-z/_A. I will schematize these ideas with the example of word-final obstruent devoicing as is follows.

(5) P-map effects on the ranking of correspondence constraints (Steriade 2001b: 5)

P-map Comparisons	<i>More distinctive contrast</i> [b]-[m] in V_ vs.	<i>Less distinctive contrast</i> [b]-[p] in V_
Ranking of Correspondence Constraints	<i>Higher ranked constraint</i> IDENT [±nas]/ V_ >>	<i>Lower ranked constraint</i> IDENT [±voice]/ V_

The first line in (5) presents the P-map comparison between the word-final obstruent nasalization and word-final obstruent devoicing. The P-map states that the perceptual contrast between word-final [b] and [m] are more distinctive than the perceptual contrast between word-final [b] and [p] (Greenberg & Jenkins 1964; refer to § 4.2.1 for the summary of this study).

These perceptual similarity differences project context-sensitive faithfulness constraints and determine their ranking relation in (5). IDENT [$\pm$ nas]/ V _ (the case of final obstruent nasalization) militates against a nasal feature change in word-final position. On the other hand, IDENT [$\pm$ voice] (the case of devoicing) militates against a voicing feature change in word-final position. Also, the ranking relation between these two context-sensitive faithfulness constraints (i.e., IDENT [$\pm$ nas]/ V _ ranks higher than IDENT [$\pm$ voice]/ V _)) is determined based on the observation that obstruency difference in word-final position is more perceptible than voicing difference in the same position.

Tableau (6) shows how the markedness constraint *[+voice]/ _ and the two faithfulness constraints IDENT [$\pm$ nas]/ V _ and IDENT [$\pm$ voice]/ V _ interact with each other, resulting in [tæp] as an optimal output.

(6) Word-final obstruent devoicing

/tæb/	*[+voice]/ _	IDENT [$\pm$ nas]/ V _	IDENT [$\pm$ voice]/ V _
[tæb]	*!		
☞ [tæp]			*
[tæm]		*!	

In (6), the most faithful output [tæb] loses its candidacy as an optimal output because it violates the highest ranked markedness constraint, *[+voice]/ _ . Between the remaining [tæp] and [tæm], the output [tæp] wins over [tæm] due to the constraint ranking IDENT

[±nas]/ V_] >> IDENT [±voice]/ V_]. Recall that the ranking relation between the two IDENT IO constraints is grounded in the perceptual similarity differences: the input /tæb/ and its modified output [tæp] are perceptually more similar to each other than the input /tæb/ and its modified output [tæm].

Taking an example from the case of word-final obstruent devoicing, I have shown how the perceptual similarity differences are translated to phonological constraints. In what follows, I will compare the formalization of nasal place assimilation under a rule-based approach and that of OT to show the advantage of an OT analysis over a rule-based approach.

Comparison between a rule-based approach and phonetically-based constraints

It has been observed that many languages assimilate a nasal consonant in place to a following obstruent (/anpa/ → [ampa]), while assimilation to a preceding obstruent (/apna/ → [apma]) is less common (Steriade 2001a). Under a rule-based approach, nasal place assimilation is expressed by the following rule.

(7) Nasal place assimilation under a rule-based approach (LaCharité & Paradis 2000: 215)

$$[\text{nasal}] \rightarrow \begin{bmatrix} \alpha \text{ coronal} \\ \beta \text{ anterior} \\ \gamma \text{ back} \end{bmatrix} / \text{---} \begin{bmatrix} \alpha \text{ coronal} \\ \beta \text{ anterior} \\ \gamma \text{ back} \end{bmatrix} \quad \text{where } \alpha, \beta, \gamma \text{ are variable over + and -}$$

The rule in (7) states that a nasal takes on the feature values of the following consonant.

LaCharité & Paradis (2000) note, regarding the rule based approach to nasal place assimilation, that “among other deficiencies of this approach, the formalism offers no insights into why the language, indeed many languages, have this process, as opposed to other processes that are formally conceivable, though unusual or even unheard of” (p. 215).

On the other hand, nasal place assimilation can be formalized under OT as follows. Tableau (8) illustrates the interaction between the markedness constraints AGREE and the faithfulness constraints IDENT [PLACE]. AGREE militates against heterorganic consonant clusters, and it basically serves as a device to allow place assimilation. In comparison, the faithfulness constraint IDENT [PLACE] requires corresponding segments in the output to share specifications for place feature in the input.

(8) Nasal place assimilation under OT (constructed based on Steriade 2001a)

/anpa/	AGREE	IDENT [PLACE]/C_V	IDENT [PLACE]/V_C
ampa			*
anta		*!	
anpa	*!		

Note in Tableau (8) that IDENT [PLACE] is further divided into two context-sensitive faithfulness constraints: IDENT [PLACE]/C_V and IDENT [PLACE]/V_C. In other words, IDENT [PLACE] in Tableau (8) does not simply refer to the identity of place features in the input and output, but it refers to the place features as realized in specific contexts. IDENT [PLACE]/C_V militates against place feature change in CV transition, while IDENT [PLACE]/V_C constraint militates against place feature change in VC transition.

As is the case with word-final obstruent devoicing, the ranking relation between

the two context-sensitive faithfulness constraints (i.e., IDENT [PLACE]/C_V >> IDENT [PLACE]/V_C) plays a crucial role in determining whether the assimilation is regressive [ampa] or progressive *[anta]. The reverse ranking would have determined *[anta] as an optimal output.

In a broad stroke, the basic mechanism of locating the target of place assimilation is that the weaker the perceptual cues of a segment, the more likely it undergoes place assimilation (Jun 2004, Steriade 2001a). Applying this mechanism to the case of place assimilation, that C₁ is the target of place assimilation in the VC₁C₂V is attributable to the observation that the place cues of a consonant are stronger in prevocalic position than in preconsonantal position. That is, /anpa/ becomes [ampa] rather than *[anta] because changing a place feature of C in VC sequence (whose acoustic cues are weaker than C in CV sequence) is more similar to the unmodified input than changing place feature of C in CV sequence.

To summarize, the observation that consonant's place of articulation is perceived differently in VC transitions and CV transitions generates two context-sensitive faithfulness constraints, IDENT [PLACE]/C_V and IDENT [PLACE]/V_C. In addition, the observation that the weaker the perceptual cues of a segment, the more likely it undergoes place assimilation projects the ranking relation, IDENT [PLACE]/C_V >> IDENT [PLACE]/V_C. In this ranking relation, changing a place feature in CV transition (IDENT [PLACE]/C_V) is more harshly penalized than a changing a place feature in VC transition (IDENT [PLACE]/V_C), which correctly generates regressive place assimilation in /anpa/ → [ampa].

Returning to the representation of nasal place assimilation under a rule-based

approach and OT, we can see that under a rule-based approach, the rule, in itself, is the inherent reason for change, and there is no formal explanation for why a language has a particular rule. By contrast, OT enables us to provide a deeper level of explanation for the process by projecting context-sensitive faithfulness constraints which embody phonetic observation. We have seen that the perceptual similarity differences of a segment in a different context are translated to context-sensitive faithfulness constraints and their relative ranking relation.

Adopting Steriade's (2001a, 2001b) P-map proposal, I assume that phonetic observation (articulatory and perceptual) projects correspondence constraints and determines the ranking relation among the correspondence constraints. With this background, § 7.3.1 and § 7.3.2 analyse the patterns of vowel insertion and consonant deletion in Korean speakers' production of English consonant clusters.

7.3. OT Constraints

7.3.1. Phonological Acquisition Under OT⁷⁷

I will begin this section by briefly discussing how the phonological acquisition is understood under OT to provide a general background about the constraint demotion process (i.e., demotion of *COMPLEX in Korean speakers' production) illustrated in the OT tableaux in this section.

According to the learning theory proposed by Tesar & Smolensky (1993, 1998,

⁷⁷ Throughout this section, *C* stands for a consonant, *V* stands for a vowel, *v/* stands for a voiceless consonant, and *vd* stands for a voiced consonant.

2000), language acquisition involves the acquisition of a language-specific hierarchy of universal constraints. It is assumed in OT that all the universal constraints are innate.⁷⁸ The task given to the language learner is to extract the information that is necessary to rank all constraints in a way that predicts the target outputs.

Given the task of the language learner, there must be a stage when the learner's current grammar declares that the optimal output for a word is an erroneous pronunciation. This is the point where constraints are minimally⁷⁹ demoted (rather than promoted) based on the positive evidence in the ambient language so that the correct pronunciation is declared optimal by the revised grammar. In short, language acquisition in OT is a process of *Constraint Demotion* until a learner's grammar reaches the successful target grammar.

What would language learners do when confronted with a different language? For second language learning, OT assumes that acquisition uses basically the same mechanism as used in first language acquisition except for the facts that a) an underlying form would probably be equal to an adult form of the target language, as perceived by the learner and b) the learner is supposed to re-rank the constraint system of his/her first language (Boersma et al. 2000: 39). The constraint demotion process in second language learning is demonstrated in (9).

⁷⁸ A strong theory of Universal grammar maintains that all the constraints the learner will ever need are available from the beginning (Tesar & Smolensky 1993:1). However, some OT researchers take rather different positions on the issue of whether constraints are innate. For instance, Boersma (1998, 2000) and Hayes (1999, 2003) propose that constraints are emergent from functional properties such as articulatory and perceptual factors which are active during acquisition.

⁷⁹ "Demotion is minimal in the sense that a constraint is demoted to a position immediately below the highest-ranking constraint that induces its violation in the optimal output" (Kager 1999: 298).

(9) Illustration of constraint demotion in L2 learning

a. Initial state

/bret/	*COMPLEX	MAX-C	DEP-V	CONSTRAINT X
a. [bret]	*!			
b. [biret]			*	
c. [bet]		*!		
d. [ret]		*!		

b. Final state

/bret/	MAX-C	DEP-V	CONSTRAINT X	*COMPLEX
a. [bret]				*
b. [biret]		*!		
c. [bet]	*!			
d. [ret]	*!			

Assume a situation in which native Korean speakers are learning English. The input in Tableau (9) is the adult form of the target language (i.e., /bret/).⁸⁰ The initial state of the Korean speakers' grammar is demonstrated in (9a), which is the same as the constraint ranking in the learners' native language, Korean. In other words, complex margins are not allowed in Korean and thus *COMPLEX (no complex margins) dominates MAX-C (no consonant deletion) and DEP-V (no vowel insertion).⁸¹ This ranking relation illustrates that at the initial state when Korean speakers produce English consonant clusters, they are likely to insert an epenthetic vowel between clusters.

In order to correctly produce complex margins in English, the markedness constraint *COMPLEX should be demoted to lower than MAX-C and DEP-V. It may be that

⁸⁰ Note that nonsense words are used as an input in this demonstration.

⁸¹ The higher ranking of MAX-C constraint over DEP-V constraint reflects the general findings in L2 acquisition of consonant clusters that in onset clusters, vowel insertion is more likely to occur than consonant deletion (e.g., Hancin-Bhatt & Bhatt 1997; Kwon 2005).

exposure to a target language input and practice of articulatory gestures of a target language result in the constraint demotion in Korean-English interlanguage grammar. By the process of the markedness constraint demotion, the final state of Korean-English interlanguage grammar regarding complex margins surfaces as in (9b).

7.3.2. Patterns of Vowel Insertion in Onset Clusters

To follow this background on Optimality theoretic phonological acquisition, I will turn to the analysis of the patterns of vowel insertion in Korean speakers' production of English consonant clusters. It is essential to keep in mind throughout this section that the constraint rankings in OT analysis are based on the patterns of vowel insertion found in Korean speakers' production of English consonant clusters rather than the proposed consonant cluster markedness hierarchy in PCCAM. Epenthesis rates in onset clusters are presented in the table below to aid our understanding of the constraint ranking in OT analysis which models the actual production of English consonant clusters by Korean speakers.

Table 7.1. Epenthesis by type of C₁ in onset clusters

Type of C ₁ (onset)	Epenthesis (n)	Total (n)	Epenthesis/Total (%)
voiced stop	76	597	12.7
voiceless stop	19	596	3.2
voiceless strident	11	716	1.5
voiceless strident < voiceless stop < voiced stop			

Note in Table 7.1 that vowel insertion is most likely to occur after a voiced stop and least likely to occur after a voiceless strident.

The patterns of vowel insertion in Korean speakers' production of English consonant clusters are captured by the interaction between the markedness constraint *COMPLEX and the faithfulness constraint DEP-V. The basic DEP -V constraint is further divided into the family of context-sensitive faithfulness constraints depending on the context where vowel insertion occurs (refer to Table 7.1). The context-sensitive faithfulness constraints for the patterns of vowel insertion in onset consonant clusters are given below.

(10) Onset clusters: A family of context-sensitive DEP-V constraints

a) DEP -V / vl strident C₁_ vl stop C₂

A vowel present in the output context of vl strident C₁_ vl stop C₂ has a correspondent vowel in the input context of vl strident C₁_ vl stop C₂.

b) DEP -V / vl strident C₁_ nasal C₂

A vowel present in the output context of vl strident C₁_ nasal C₂ has a correspondent vowel in the input context of vl strident C₁_ nasal C₂.

c) DEP -V / vl strident C₁_ liquid C₂

A vowel present in the output context of vl strident C₁_ liquid C₂ has a correspondent vowel in the input context of vl strident C₁_ liquid C₂.

d) DEP-V / vl stop C₁_ liquid C₂

A vowel present in the output context of vl stop C₁_ liquid C₂ has a correspondent vowel in the input context of vl stop C₁_ liquid C₂.

e) DEP-V / vd stop C₁_ liquid C₂

A vowel present in the output context of vd stop C₁_ liquid C₂ has a correspondent vowel in the input context of vd stop C₁_ liquid C₂.

f) Ranking relation

DEP-V / vl strident C₁_ vl stop C₂, DEP -V / vl strident C₁_ nasal C₂, DEP -V / vl strident C₁_ liquid C₂,

>>

DEP-V / vl stop C₁_ liquid C₂

>>

DEP-V / vd stop C₁_ liquid C₂

The constraint DEP-V / vl strident C₁_ vl stop C₂ militates against vowel insertion in the context of vl strident C₁_ vl stop C₂. On the other hand, the constraint DEP-V / vd stop C₁_ liquid C₂ militates against vowel insertion in the context of vd stop C₁_ liquid C₂.

Note also that no ranking relation is established between two constraints if the C₁ in a cluster shares the same voicing and stridency feature. This is because the major phonetic factors accounting for vowel insertion (i.e., the articulatory and perceptual voicing effects, and the articulatory and perceptual stridency effects) refer to the stridency and voicing feature of a segment. The three constraints DEP-V / **vl strident** C₁_ vl stop C₂, DEP -V / **vl strident** C₁_ nasal C₂, and DEP-V / **vl strident** C₁_ liquid C₂ share [+continuant][-voicing], and thus are grouped together.

The ranking relation in (10f) demonstrates that the DEP-V constraints in (10a-e) differently penalize vowel insertion depending on the context where it occurs. We can see that DEP-V / vl stop C₁_ liquid C₂ dominates DEP-V / vd stop C₁_ liquid C₂, which means that vowel insertion in the context of vl stop C₁_ liquid C₂ is penalized more harshly than vowel insertion in the context of vd stop C₁_ liquid C₂.

Recall that the ranking relation in (10f) is grounded in the articulatory and perceptual factors identified in the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM). In § 5.3.2, I showed that vowel insertion is more likely to occur in voiced stop + liquid than in voiceless stop + liquid due to the articulatory voicing effect. In turn, vowel insertion is less likely to occur in /s/ + C clusters than in stop + liquid clusters because of the articulatory and perceptual stridency effect.

Tableaux (11) demonstrate the patterns of vowel insertion in onset consonant clusters. Note that it is assumed in the present study that different types of IL grammars are attested depending on the degree the markedness constraint *COMPLEX is demoted in Korean-English IL grammar. For instance, the tableaux in (11) present four different types of grammars (Grammar A-D), which are determined by the position of the markedness constraint *COMPLEX in the ranking.

For Grammar A, *COMPLEX dominates all context-sensitive faithfulness constraints, and accordingly all English onset clusters are produced with an epenthetic vowel. In the production experiment presented in Chapter 6, one participant (out of 30) demonstrated this type of IL grammar.

(11) Patterns of vowel insertion in onset clusters

(a) Grammar A

		*COMPLEX	DEP-V / vl strident C ₁ _ * vl stop C ₂	DEP-V / vl strident C ₁ _ nasal C ₂	DEP-V / vl strident C ₁ _ liquid C ₂	DEP-V / vl stop C ₁ _ liquid C ₂	DEP-V / vd stop C ₁ _ liquid C ₂
/sket/	[siket]						
	*[sket]	*!					
/smet/	[simet]			*			
	*[smet]	*!					
/slet/	[silet]				*		
	*[slet]	*!					
/pret/	[piret]					*	
	*[pret]	*!					
/bret/	[biret]						*
	*[bret]	*!					

(b) Grammar B

		DEP-V / vl strident C ₁ _ vl stop C ₂	DEP-V / vl strident C ₁ _ nasal C ₂	DEP-V / vl strident C ₁ _ liquid C ₂	*COMPLEX	DEP-V / vl stop C ₁ _ liquid C ₂	DEP-V / vd stop C ₁ _ liquid C ₂
/sket/	[sket]				*		
	*[sket]	*!					
/smet/	[smet]				*		
	*[smet]		*!				
/slet/	[slet]				*		
	*[slet]			*!			
/pret/	[piret]					*	
	*[pret]				*!		
/bret/	[biret]						*
	*[bret]				*!		

(c) Grammar C

	DEP-V / vl strident C ₁ ₁	DEP-V / vl strident C ₁ ₁ nasal C ₂	DEP-V / vl strident C ₁ ₁ liquid C ₂	DEP-V / vl stop C ₁ ₁ liquid C ₂	*COMPLEX	DEP-V / vd stop C ₁ ₁ liquid C ₂
/sket/ [sket]					*	
*[siket]	*!					
/smet/ [smet]					*	
*[simet]		*!				
/slet/ [slet]					*	
*[silet]			*!			
/pret/ [pret]					*	
*[piret]				*!		
/bret/ [bret]						*
*[bret]					*!	

(d) Grammar D

	DEP-V / vl strident C ₁ ₁	DEP-V / vl strident C ₁ ₁ nasal C ₂	DEP-V / vl strident C ₁ ₁ liquid C ₂	DEP-V / vl stop C ₁ ₁ liquid C ₂	DEP-V / vd stop C ₁ ₁ liquid C ₂	*COMPLEX
/sket/ [sket]						*
*[siket]	*!					
/smet/ [smet]						*
*[simet]		*!				
/slet/ [slet]						*
*[silet]			*!			
/pret/ [pret]						*
*[piret]				*!		
/bret/ [bret]						*
*[biret]					*!	

For Grammar B in (11b), *COMPLEX is demoted to lower than vl strident C₁_liquid C₂ and thus for those speakers with this type of grammar, vowel insertion occurs in a cluster of voiceless stop + liquid and voiced stop + liquid. Five participants demonstrated this type of grammar in the production experiment. The case when *COMPLEX is demoted to lower than vl stop C₁_liquid C₂ is demonstrated in (11c), and thus vowel insertion occurs only in a cluster of voiced stop + liquid. There were 8 participants with this type of grammar. The final state of the grammar for onset consonant clusters is demonstrated in (11d). When complex margins are produced correctly, *COMPLEX is demoted to the invisible position.⁸² There were 12 participants with this type of grammar in the production experiment.

It should be noted that there were four participants (out of 30) whose current grammars did not fit any of the types of grammars in (11). For instance, two participants who inserted an epenthetic vowel in sC onset clusters did not insert any epenthetic vowel in other types of onset clusters. The other two participants inserted an epenthetic vowel in /s/ + liquid clusters and voiceless stop + liquid clusters, respectively, but not in other clusters.

It should come as no surprise that four participants (out of 30) did not fit in any of the types of grammars for the patterns of vowel insertion in onset clusters since “one of the characteristics of the output of second language learners is that it is quite variable” (Archibald 1998: 6). Despite this variability in L2 acquisition, the majority of the participants in the present study fit with the constraint ranking in (11). Therefore, we can say that context-sensitive DEP-V constraints for English onset clusters provide a clear

⁸² “Invisible” in the sense that it is not active in the IL grammar.

illustration of the patterns of vowel insertion in Korean speakers' production of English onset clusters.

7.3.3. Patterns of Vowel Insertion in Coda Clusters

We now turn to the patterns of vowel insertion in coda clusters. As was shown in Chapter 6, vowel insertion occurred after C_1 ($VCC \rightarrow VCvC$) or after C_2 ($VCC \rightarrow VCCv$) in coda clusters. I will only provide a formal analysis for the patterns of vowel insertion after C_1 as the patterns of vowel insertion after C_2 is basically the same as that after C_1 : vowel insertion is most likely to occur after a voiced stop and least likely to occur after a voiceless strident. In addition, the patterns of vowel insertion in sonorant + C clusters are not formalized, since vowel insertion after C_1 occurs only in a cluster of liquid + nasal (but not in other sonorant + C clusters).

To help our understanding of the constraint ranking in OT analysis which models the actual production of English consonant clusters by Korean speakers, epenthesis rates in coda clusters are presented in Table 7.2 and Table 7.3 below.

Table 7.2. Epenthesis by type of C_1 in coda clusters

Type of C_1 in Coda	Epenthesis (n)	Total (n)	Epenthesis/Total (%)
voiced stop	80	455	17.6
voiceless stop	5	583	0.9
voiceless strident	0	354	0.0
voiceless strident < voiceless stop < voiced stop			

Table 7.3. Epenthesis by type of C₂ in coda clusters⁸³

Type of Coda Cluster C ₂	Epenthesis (n)	Total (n)	Epenthesis/Total (%)
voiced stop	158	694	22.8
voiceless stop	141	1179	12.0
voiceless strident	50	584	8.6
voiceless strident < voiceless stop < voiced stop			

A family of context-sensitive DEP-V constraints for the patterns of vowel insertion in English coda clusters is presented in (12).

(12) Coda clusters: A family of context-sensitive DEP-V constraints

a) DEP-V / vl strident C₁_ vl stop C₂

A vowel present in the output context of vl strident C₁_ vl stop C₂ has a correspondent vowel in the input context of vl strident C₁_ vl stop C₂.

b) DEP-V / vl stop C₁_ vl strident C₂

A vowel present in the output context of vl stop C₁_ vl strident C₂ has a correspondent vowel in the input context of vl stop C₁_ vl strident C₂

c) DEP-V / vl stop C₁_ vl stop C₂

A vowel present in the output context of vl stop C₁_ vl stop C₂ has a correspondent vowel in the input context of vl stop C₁_ vl stop C₂

d) DEP-V / vd stop C₁_ vd strident C₂

A vowel present in the output context of vd stop C₁_ vd strident C₂ has a correspondent vowel in the input context of vd stop C₁_ vd strident C₂.

e) DEP-V / vd stop C₁_ vd stop C₂

A vowel present in the output context of vd stop C₁_ vd stop C₂ has a correspondent vowel in the input context of vd stop C₁_ vd stop C₂.

⁸³ Epenthesis by type of C₂ in Table 7.3 is provided to demonstrate that the patterns of vowel insertion (i.e., voiceless strident < voiceless stop < voiced stop) are the same regardless of whether they occur after C₁ or after C₂ in coda clusters. Also, the vowel insertion pattern after a voiced strident is not presented here because constraints for vowel insertion after a *voiced strident* or after a *voiceless stop* cannot be *directly* ranked in terms of the voicing and stridency features. Refer to § 4.3.1 (under the title of *the case where no ranking relationship is established*) for discussion of this issue.

f) Ranking Relation

$$\begin{array}{c}
 \text{DEP-V / vl strident } C_1_ \text{ vl stop } C_2 \\
 >> \\
 \text{DEP-V / vl stop } C_1_ \text{ vl strident } C_2, \text{ DEP-V / vl stop } C_1_ \text{ vl stop } C_2 \\
 >> \\
 \text{DEP-V / vd stop } C_1_ \text{ vd strident } C_2, \text{ DEP-V / vd stop } C_1_ \text{ vd stop } C_2
 \end{array}$$

The constraint DEP-V/vl strident $C_1_ \text{ vl stop } C_2$ militates against vowel insertion in the output context of DEP-V/vl strident $C_1_ \text{ vl stop } C_2$. On the other hand, DEP-V/vd stop $C_1_ \text{ vd stop } C_2$ militates against vowel insertion in the output context of vd stop $C_1_ \text{ vd stop } C_2$. The ranking relation in (12f) demonstrates that DEP-V/vl strident $C_1_ \text{ vl stop } C_2$ dominates all the other constraints, which means that vowel insertion in a voiceless strident + voiceless stop coda cluster is most harshly penalized.

Note also that no ranking relation is established in DEP-V/**vl stop** $C_1_ \text{ vl strident } C_2$ **vs.** DEP-V/**vl stop** $C_1_ \text{ vl stop } C_2$. You can notice that the voicing and stridency features of C_1 are the same in these constraints, and thus are grouped together. The same rule applies to the ranking relation between DEP-V/**vd stop** $C_1_ \text{ vd strident } C_2$ and DEP-V/**vd stop** $C_1_ \text{ vd stop } C_2$.

Tableaux in (13) illustrate the patterns of vowel insertion in coda clusters.

(13) Patterns of vowel insertion in coda consonant clusters

(a) Grammar A

		*Complex	DEP-V /vl strident C ₁ _ vl stop C ₂	DEP-V / vl stop C ₁ _ vl strident C ₂	DEP-V / vl stop C ₁ _ stop C ₂	DEP-V / vd stop C ₁ _ vd strident C ₂	DEP-V / vd stop C ₁ _ vd stop C ₂
/tesk/	[tesik]		*				
	* [tesk]	*!					
/teps/	[tepis]			*			
	* [teps]	*!					
/tept/	[tepit]				*		
	* [tept]	*!					
/tebz/	[tebiz]					*	
	* [tebz]	*!					
/tebd/	[tebid]						*
	* [tebd]	*!					

(b) Grammar B

		DEP-V /vl strident C ₁ _ vl stop C ₂	*Complex	DEP-V / vl stop C ₁ _ vl strident C ₂	DEP-V / vl stop C ₁ _ stop C ₂	DEP-V / vd stop C ₁ _ vd strident C ₂	DEP-V / vd stop C ₁ _ vd stop C ₂
/tesk/	[tesk]		*				
	* [tesik]	*!					
/teps/	[tepis]			*			
	* [teps]		*!				
/tept/	[tepit]				*		
	* [tept]		*!				
/tebz/	[tebiz]					*	
	* [tebz]		*!				
/tebd/	[tebid]						*
	* [tebd]		*!				

(c) Grammar C

		DEP-V /vl strident C ₁ _ vl stop C ₂	DEP-V / vl stop C ₁ _ vl strident C ₂	DEP-V / vl stop C ₁ _ vl stop C ₂	*Complex	DEP-V / vd stop C ₁ _ vd strident C ₂	DEP-V / vd stop C ₁ _ vd stop C ₂
☞ /tesk/	[tesk]				*		
	* [tesik]	*!					
☞ /teps/	[teps]				*		
	* [tepis]		*!				
☞ /tept/	[tept]				*		
	* [tepit]			*!			
☞ /tebz/	[tebiz]					*	
	* [tebz]				*!		
☞ /tebd/	[tebid]						*
	* [tebd]				*!		

(d) Grammar D

		DEP-V /vl strident C ₁ _ vl stop C ₂	DEP-V / vl stop C ₁ _ vl strident C ₂	DEP-V / vl stop C ₁ _ vl stop C ₂	DEP-V / vd stop C ₁ _ vd strident C ₂	DEP-V / vd stop C ₁ _ vd stop C ₂	*Complex
☞ /tesk/	[tesk]						*
	* [tesik]	*!					
☞ /teps/	[teps]						*
	* [tepis]		*!				
☞ /tept/	[tept]						*
	* [tepit]			*!			
☞ /tebz/	[tebz]						*
	* [tebiz]				*!		
☞ /tebd/	[tebd]						*
	* [tebid]					*!	

As was the case with the patterns of vowel insertion in onset clusters discussed in Tableaux (11), the position of *COMPLEX determines the type of IL grammar for Korean learners of English. Thus, for Grammar A, *COMPLEX dominates all context-sensitive faithfulness constraints, and accordingly, all English coda clusters are produced with an epenthetic vowel. This type of grammar was not attested in Korean speakers' production in the present study, yet it does not seem unreasonable to assume that L2 learners at their very initial state of L2 learning may demonstrate Grammar A.

For Grammar B in (13b), *COMPLEX is demoted to lower than DEP-V / vl strident C₁_ vl stop C₂, and thus for the speakers with this type of grammar, vowel insertion occurs in a cluster of voiceless stop + voiceless strident, voiceless stop + voiceless stop, voiced stop + voiced strident, and voiced stop + voiced stop. Two participants turned out to have Grammar B. In (13c), *COMPLEX is demoted to lower than DEP-V / vl stop C₁_ vl stop C₂. There were 11 participants with this type of grammar. The final state of the grammar for onset consonant clusters is demonstrated in (13d). When complex margins are produced without vowel insertion, *COMPLEX is demoted to the invisible position. There were 15 participants with this type of grammar.

Lastly, there were two participants whose grammar did not correspond to the constraint ranking in (12). The participants inserted an epenthetic vowel after a voiceless stop in a voiceless stop + voiceless strident cluster, but did not do so in other clusters.

7.3.4. Patterns of Consonant Deletion in Coda Clusters ⁸⁴

This section discusses patterns of consonant deletion in Korean speakers' production of English consonant clusters. A family of context-sensitive MAX-C constraints that figure in the patterns of consonant deletion is presented in (14). Note that context-sensitive MAX-C constraints are divided into two groups depending on the cluster type: a) obstruent + obstruent clusters and b) sonorant + C clusters

(14) Coda clusters: A family of context-sensitive MAX-C constraints

(14-1) Obstruent + obstruent clusters

a) MAX-strident (No deletion of a strident segment)

A [strident] segment present in the input has a correspondent [strident] segment in the output.

b) MAX-stop (No deletion of a stop segment)

A [stop] segment present in the input has a correspondent in the output.

c) Ranking relation: MAX-strident >> MAX-stop

(14-2) Sonorant + C clusters

a) MAX-liquid (No deletion of a liquid segment)

A [liquid] segment present in the input has a correspondent [liquid] segment in the output.

b) MAX-nasal (No deletion of a nasal segment)

A [nasal] segment present in the input has a correspondent [nasal] segment in the output.

⁸⁴ Formal analyses for the patterns of consonant deletion in *onset clusters* are not provided because there were only four tokens of consonant deletion, and they were all liquid deletion.

c) MAX-stop (No deletion of a stop segment)

A [stop] segment present in the input has a correspondent in the output.

d) MAX-liquid/ $V_C_{[+nasal]\#}$

(No deletion of a liquid segment in a liquid + nasal cluster)

A [liquid] segment present in the input context of $V_C_{[+nasal]\#}$ has a correspondent [liquid] segment in the output context of $V_C_{[+nasal]\#}$.

e) MAX-nasal/ $VC_{[+liquid]_}\#$

(No deletion of a nasal segment in a liquid + nasal cluster)

A [nasal] segment present in the input context of $VC_{[+liquid]_}\#$ has a correspondent [nasal] segment in the output context of $VC_{[+liquid]_}\#$.

f) Ranking relation:

MAX-liquid, MAX-nasal $\gg$ MAX-stop, MAX-nasal/ $VC_{[+liquid]_}\#$, MAX-liquid/ $V_C_{[+nasal]\#}$

MAX-strident blocks deletion of a [strident] segment in the output, while MAX-stop blocks deletion of a stop segment. The ranking relation in (14-1, c) reflects the finding that a stop, not a strident, is the main target of deletion in obstruent + obstruent clusters.⁸⁵

The constraints in (14-2) capture the patterns of consonant deletion in sonorant + C clusters. Notice in the constraints in (14-2) that the two context-sensitive faithfulness constraints, MAX-liquid/ $V_C_{[+nasal]\#}$ and MAX-nasal/ $VC_{[+liquid]_}\#$ are generated to capture the large number of liquid and nasal deletions in a liquid + nasal cluster.

More specifically, that a liquid and a nasal are the target of deletion mostly in a liquid + nasal cluster is reflected in the constraint ranking in MAX-liquid, MAX-nasal $\gg$

⁸⁵ It is plausible that stop deletion is more likely to occur in a certain context than in others. For instance, it was observed in the production experiment that in obstruent + obstruent clusters, C_2 stop is more likely to be deleted than C_1 stop. This asymmetry can be captured through context-sensitive faithfulness constraints such as MAX-stop/ $VC_{[+obstruent]_}\#$ and MAX-stop/ $V_C_{[+obstruent]\#}$. That C_2 stop is more likely to be deleted ranks MAX-stop/ $V_C_{[+obstruent]\#}$ higher than MAX-stop/ $VC_{[+obstruent]_}\#$. However, I only use the context-free MAX-stop constraint in this chapter for the sake of simplicity.

MAX-nasal/ VC_[+liquid]_#, MAX-liquid/ V_C_[+nasal]#. Thus, MAX-liquid and MAX-nasal constraints are in general highly ranked and block deletion of a liquid and a nasal in sonorant + obstruent clusters. Meanwhile, the two context-sensitive faithfulness constraints MAX-nasal/ VC_[+liquid]_# and MAX-liquid/ V_C_[+nasal]# are ranked lower than MAX-liquid and MAX-nasal constraints to capture the frequent deletion of a nasal and a liquid in a liquid + nasal cluster.

Tableau (15) and Tableau (16) below present the patterns of consonant deletion in Korean speakers' production.

(15) Patterns of consonant deletion in obstruent + obstruent coda clusters

		*COMPLEX	MAX-strident	MAX-stop
(a)	/tesp/			
	* [tes]			*
	* [tep]		*!	
	* [tesp]	*!		
(b)	/tezd/			
	* [tez]			*
	* [ted]		*!	
	* [tezd]	*!		
(c)	/tept/			
	* [tep]			*
	* [tet]			*
	* [tept]	*!		
(d)	/tebz/			
	* [tez]			*
	* [teb]		*!	
	* [tebz]	*!		
(e)	/tebd/			
	* [teb]			*
	* [ted]			*
	* [tebd]	*!		

In Tableau (15), *COMPLEX dominates all the other constraints; thus output with a complex margin cannot be an optimal output. The ranking relation between MAX-strident and MAX-stop demonstrates that deletion of a strident is penalized more harshly than deletion of a stop. Thus, the interaction between MAX-strident and MAX-stop makes it possible in (a) for the candidate [tes] to be realized as an optimal output when the input is

/tesp/. Note also that in (c) both [tep] and [tet] are the optimal outputs of /tept/ because a stop is the target of deletion regardless of the position in which it occurs.

Tableau (16) presents patterns of consonant deletion in sonorant + obstruent coda clusters.

(16) Patterns of consonant deletion in sonorant + obstruent coda clusters

		*COMPLEX	MAX-liquid	MAX-nasal	MAX-stop	MAX-nasal/ VC _[+liquid] #	MAX-liquid/ V_C _[+nasal] #
(a)	/teld/				*		
	* [ted]		*!				
	* [teld]	*!					
(b)	/tend/				*		
	* [ted]			*!			
	* [tend]	*!					
(c)	/telm/			*		*	
	* [tem]		*				*
	* [telm]	*!					

The constraint ranking MAX-liquid >> MAX-stop reflects the finding that a liquid is less likely to be deleted than a stop if it does not occur in a liquid + nasal cluster. For instance, in (a) [ted] cannot be an optimal output of the input /teld/ because it violates the higher ranked constraint MAX-liquid. Thus, [tel], which violates the lower ranked constraint MAX-stop, becomes the winner. The two lowest ranked constraints MAX-nasal/ VC_[+liquid] # and MAX-liquid/ V_C_[+nasal] # capture the observation that both a nasal and a liquid are the targets of deletion in a liquid + nasal cluster. So, in (c), both [tel] and [tem] are optimal outputs of /telm/.

What follows is the summary of context-sensitive DEP-V and MAX-C constraints

used in the patterns of vowel insertion and consonant deletion in Korean speakers' production of English consonant clusters.

(17) Summary of OT constraints

a) Onset clusters: A family of context-sensitive DEP-V constraints

a) DEP -V/vl strident C₁_ vl stop C₂

A vowel present in the output context of vl strident C₁_ vl stop C₂ has a correspondent vowel in the input context of vl strident C₁_ vl stop C₂.

b) DEP -V / vl strident C₁_ nasal C₂

A vowel present in the output context of vl strident C₁_ nasal C₂ has a correspondent vowel in the input context of vl strident C₁_ nasal C₂.

c) DEP -V / vl strident C₁_ liquid C₂

A vowel present in the output context of vl strident C₁_ liquid C₂ has a correspondent vowel in the input context of vl strident C₁_ liquid C₂.

d) DEP-V / vl stop C₁_ liquid C₂

A vowel present in the output context of vl stop C₁_ liquid C₂ has a correspondent vowel in the input context of vl stop C₁_ liquid C₂.

e) DEP-V / vd stop C₁_ liquid C₂

A vowel present in the output context of vd stop C₁_ liquid C₂ has a correspondent vowel in the input context of vd stop C₁_ liquid C₂.

f) Ranking relation

DEP-V / vl strident C₁_ vl stop C₂, DEP -V / vl strident C₁_ nasal C₂, DEP -V / vl strident C₁_ liquid C₂

>>

DEP-V / vl stop C₁_ liquid C₂

>>

DEP-V / vd stop C₁_ liquid C₂

(b) Coda clusters: A family of context-sensitive DEP-V constraints

a) DEP-V / vl strident C₁_ vl stop C₂

A vowel present in the output context of vl strident C₁_ vl stop C₂ has a correspondent vowel in the input context of vl strident C₁_ vl stop C₂.

b) DEP-V / vl stop C₁_ vl strident C₂

A vowel present in the output context of vl stop C₁_ vl strident C₂ has a correspondent vowel in the input context of vl stop C₁_ vl strident C₂

c) DEP-V / vl stop C₁_ vl stop C₂

A vowel present in the output context of vl stop C₁_ vl stop C₂ has a correspondent vowel in the input context of vl stop C₁_ vl stop C₂

d) DEP-V / vd stop C₁_ vd strident C₂

A vowel present in the output context of vd stop C₁_ vd strident C₂ has a correspondent vowel in the input context of vd stop C₁_ vd strident C₂.

e) DEP-V / vd stop C₁_ vd stop C₂

A vowel present in the output context of vd stop C₁_ vd stop C₂ has a correspondent vowel in the input context of vd stop C₁_ vd stop C₂.

g) Ranking relation

DEP-V / vl strident C₁_ vl stop C₂

>>

DEP-V / vl stop C₁_ vl strident C₂, DEP-V / vl stop C₁_ vl stop C₂

>>

DEP-V / vd stop C₁_ vd strident C₂, DEP-V / vd stop C₁_ vd stop C₂

(c) Coda clusters: A family of context-sensitive MAX-C constraints

Obstruent + obstruent coda clusters

a) MAX-strident (No deletion of a strident segment)

A [strident] segment present in the input has a correspondent [strident] segment in the output.

b) MAX-stop (No deletion of a stop segment)

A [stop] segment present in the input has a correspondent in the output.

c) Ranking Relation: MAX-strident >> MAX- stop

Sonorant + C coda clusters

a) MAX-liquid (No deletion of a liquid segment)

A [liquid] segment present in the input has a correspondent [liquid] segment in the output.

b) MAX-nasal (No deletion of a nasal segment)

A [nasal] segment present in the input has a correspondent [nasal] segment in the output.

c) MAX-stop (No deletion of a stop segment)

A [stop] segment present in the input has a correspondent in the output.

d) MAX-liquid/ V_C_[+nasal]#

(No deletion of a liquid segment in a liquid + nasal cluster)

A [liquid] segment present in the input context of V_C_[+nasal] has a correspondent [liquid] segment in the output context of V_C_[+nasal].

e) MAX-nasal/ VC_[+liquid]_#

(No deletion of a nasal segment in a liquid + nasal cluster)

A [nasal] segment present in the input context of VC_[+liquid]_# has a correspondent [nasal] segment in the output context of VC_[+liquid]_#.

f) Ranking relation:

MAX-liquid, MAX-nasal >> MAX-stop, MAX-nasal/ VC_[+liquid]_#, MAX-liquid/
V_C_[+nasal]#

To conclude, the OT analyses for the patterns of vowel insertion and consonant deletion proposed in this chapter show that a specific ranking of *COMPLEX with respect to the general ranking of context-sensitive faithfulness constraints mirrors attested variations. In other words, the interaction of the context-sensitive faithfulness constraints and the markedness constraint *COMPLEX captures the variations in the patterns of vowel

insertion and consonant deletion found in Korean speakers' production.

In addition, the flexibility of *COMPLEX in the constraint ranking relation demonstrates the plasticity of interlanguage grammars: the demotion of *COMPLEX is in action in the learners' grammar.

CHAPTER 8

CONCLUSION

This dissertation has examined the roles that articulatory and perceptual factors play in Korean speakers' production of English consonant clusters. I argued that the near-failure of the sonority-based model in L2 acquisition of consonant clusters is attributable to the lack of phonetically natural explanations for the consonant cluster markedness.

As an alternative, I constructed the *Phonetically Based Consonant Cluster Acquisition Model* (PCCAM; Chapter 4) which crucially refers to roles and interactions of articulatory and perceptual factors in determining the relative consonant cluster markedness. I tested this model through perception (Chapter 5) and production experiments (Chapter 6), both of which involved Korean learners of English. PCCAM was also tested through existing L1 and L2 consonant cluster acquisition data (Chapter 4).

In particular, I looked at the patterns of vowel insertion and consonant deletion in Korean speakers' production of English consonant clusters, and proposed an extension of the correspondence constraints of Optimality Theory to account for these patterns (Chapter 7). In the formal analysis of the patterns of vowel insertion and consonant deletion, I have shown that the articulatory and perceptual factors are included in the interlanguage grammar in the form of context-sensitive faithfulness constraints. These context-sensitive faithfulness constraints differently penalize vowel insertion and consonant deletion in consonant clusters, and the relative position of the markedness constraint *COMPLEX in the constraint ranking successfully capture the type of current IL grammar for each Korean speaker.

Review of research questions

As a way of summarizing the findings of the present study, I will review the research questions raised at the outset of this dissertation (§ 1.3) and show how each research question has been answered. The research questions of this dissertation were as follows.

Research questions:

- 1) What are the articulatory factors resulting in vowel insertion and consonant deletion in Korean speakers' production of English consonant clusters?
- 2) What are the perceptual factors resulting in vowel insertion and consonant deletion in Korean speakers' production of English consonant clusters?
- 3) How can we define consonant cluster markedness that will determine a) language learners' variable degree of production difficulty with non-native consonant sequences and b) learners' repair strategies when dealing with a certain type of consonant sequence, based on articulatory and perceptual factors?
- 4) Does the *Phonetically Based Consonant Cluster Acquisition Model* have an explanatory power beyond Korean speakers' production of English consonant clusters? In other words, can the *Phonetically Based Consonant Cluster Acquisition Model* be applied to L2 learners with other first language backgrounds and first language acquisition of consonant clusters?
- 5) How can we formalize the articulatory and perceptual factors that define consonant cluster markedness in the grammar?

As an answer to the first research question: What are the articulatory factors resulting in vowel insertion and consonant deletion in Korean speakers' production of English consonant clusters?, four articulatory factors were identified and incorporated into PCCAM: a) the degree of gestural overlap, b) the articulatory release effect, c) the articulatory stridency effect, and d) the articulatory voicing effect.

This dissertation adopted the view of Articulatory Phonology (AP) on phonological acquisition that language learners are supposed to elaborate individual gestures by differentiating, tuning, and coordinating those gestures. In particular, the hypothesis for the degree of gestural overlap parameter was derived based on the principle of AP that well-formed consonant sequences will be least overlapped and least variable (Browman & Goldstein 1990b).

The second research question, on the other hand, deals with the perceptual factors resulting in vowel insertion and consonant deletion in Korean speakers' production of English consonant clusters. The four perceptual factors identified and incorporated into PCCAM were a) the perceptual stridency effect, b) the perceptual voicing effect, c) the principle of perceptual saliency, and d) the maximal perceptual contrast.

The P-map hypothesis (Steriade 2001a, 2001b) that an alternation that involves a less perceptible change is more likely to take place than an alternation that involves a more perceptible change was adopted as a working mechanism to shape the perceptual stridency effect and the perceptual voicing effect. The validity of the two perceptual principles, the perceptual stridency effect and the perceptual voicing effect, was tested through the auditory similarity judgment experiment in Chapter 5.

Taking these articulatory and perceptual factors into account, I have shown that

PCCAM calculates the relative difficulty of producing consonant clusters through a three-way computation of consonant cluster markedness: a) markedness of C₁ (the first consonant in a cluster), b) markedness of C₂ (the second consonant in a cluster), and c) markedness of gestural coordination of C₁ and C₂. The three-way computation mechanism employed in this dissertation provides an answer to the third research question: How can we define consonant cluster markedness based on articulatory and perceptual factors?

The sources for the three-way markedness computation were, of course, the eight articulatory and perceptual factors mentioned above. The three-way computation of consonant clusters markedness generated the following consonant cluster markedness hierarchy.

(1) Consonant cluster markedness hierarchy in PCCAM

less marked < more marked; // = no markedness relation	
Onset Clusters	sp, st, sk < sm, sn < sl < pl, pr, tr, kl, kr < bl, br, dr, gl, gr
Coda Clusters: Sonorant + C	ls, ns, < lz, nz // lp, lt, lk, np, nt, mp < lb, ld, nd // lm
Coda Clusters: Obstruent + Obstruent	ps, ts, ks < sp, st, sk < pt, kt // bz, gz < zd < bd, gd

The validity of the consonant cluster markedness hierarchy in PCCAM (table above) was tested a) by referring to the published L1 and L2 consonant cluster acquisition data (Chapter 4), and also through b) Korean speakers’ production of English consonant clusters (Chapter 6). It was shown that most of the markedness ranking relationships in (1) hold true in Korean speakers’ production data (10 cases out of 13

ranking relationships) as well as the published L1 and L2 consonant cluster acquisition data. The relative markedness between voiceless strident + voiceless stop and voiceless stop + voiceless stop (sp, st, sk < pt, kt) is the only case (among 13 ranking relationships) that the error rate difference is statistically significant in Korean speakers' production, yet the consonant cluster markedness hierarchy fails to provide the correct ranking relationship.⁸⁶

That the proposed consonant cluster markedness hierarchy in (1) was supported in the published L1 and L2 consonant cluster acquisition data provides an answer to the fourth research question: Does PCCAM have an explanatory power beyond Korean speakers' production of English consonant clusters?

The last research question is about the issue of how we can translate the phonological observations to constraints (or rules) in a formal phonological theory. Related to this issue is how to model L2 learners' grammar within a formal phonological theory. These two issues were stated as a research question: How can we formalize the articulatory and perceptual factors that define consonant cluster markedness in the grammar?

I have shown in this dissertation that the intrinsic structure of Optimality Theory (Prince & Smolensky 1993) makes it eminently suitable for the issues of translating phonetic information into phonological constraints and modeling phonological acquisition. For instance, when it comes to the issue of translating phonetic evidence to phonological constraints, the basic mechanism of OT, which separates markedness constraints that embody a particular form of phonetic difficulty from the issue of how and

⁸⁶ The ranking relationships a) between /s/ + stop and /s/ + nasal and b) between /s/ + liquid and voiceless stop + liquid were not supported either. However, I think this is less of a problem because the error rate differences in these clusters were not statistically significant.

whether the difficulty is avoided, provides an efficient way for phonetic observation to be encoded as a form of ranked, conflicting constraints in the system. More specifically, I proposed an extension of the correspondence constraints of Optimality Theory and demonstrated that the articulatory and perceptual factors are included in the interlanguage grammar in the form of context-sensitive faithfulness constraints.

Contribution of this dissertation to L2 phonological acquisition research

The question arises as to: What are the contributions of this dissertation in L2 phonological acquisition? More specifically, how is the *Phonetically-Based Consonant Cluster Acquisition Model* (PCCAM) different from the previous models of L2 acquisition of consonant clusters in making predictions about L2 learners' difficulty with producing non-native consonant clusters and accounting for the asymmetries found in their production?

There are many advantages of PCCAM over previous consonant cluster acquisition models. First, PCCAM showed that incorporating articulatory and perceptual factors into the acquisition model enables us to provide a wider and deeper understanding of the phenomena at hand. Consider the case of the relative markedness between stop + /s/ and /s/ + stop (ps, ts, ks < sp, st, sk). A fair number of L1 and L2 consonant cluster acquisition data demonstrated the relative unmarkedness of stop + /s/ sequences over /s/ + stop sequences (e.g., Dyson 1998; Hansen 2000; Hansen 2001; Jakielski 1998; Kirk & Demuth 2003; Olmsted 1971; Templin 1957; Wellman et al. 1931; Wiltshire 2005).

How can this pattern be accounted for under the sonority-based consonant cluster acquisition model (e.g., sonority distance parameter, sonority dispersion principle)? I

have shown in Chapter 3 that the sonority based model does not have a ready answer for this pattern due to the sonority reversal in stop + /s/ coda clusters.

Some researchers (e.g., Major 1996) rely on the notion of extrasyllabicity to account for the special status of stop + /s/ coda clusters in L2 learners' production. However, to make the account satisfactory, I think those who rely on the notion of extrasyllabicity have to answer questions such as: What is it in the extrasyllabicity that makes the sequence with an extrasyllabic segment easier for language learners to produce? How do language learners access (in particular L2 learners) the prosodic representation of /s/ as extrasyllabic? Is every sequence with an extrasyllabic segment easier for language learners to produce or acquire?

As an alternative to the extrasyllabic account, I have shown that stop + /s/ sequences are most favored among obstruent + obstruent coda clusters in terms of three aspects: a) degree of gestural overlap, b) principle of perceptual saliency, and c) articulatory release effect. In other words, the relative unmarkedness of stop + /s/ coda clusters is grounded in inherent phonetic factors associated with stop + /s/ rather than the abstract notion of extrasyllabicity of a segment.

The advantage of PCCAM over the sonority-based consonant cluster acquisition model is also demonstrated well with /s/ + stop onset clusters and /s/ + stop coda clusters. Recall that previous studies have shown that L2 learners do not show much difficulty with producing /s/ + stop onset clusters (Chen 2003; Major 1996; Kim 2000; and the present study), while it was reported that /s/ + stop coda clusters are relatively problematic for L2 learners (Major 1996; Hansen 2000, 2001; Wiltshire 2005).

How can this pattern be explained under the sonority-based model? When

viewing the two sequences in terms of the minimal sonority distance parameter (Broselow & Finer 1991), /s/ + stop onset clusters have a sonority value of -1 (or 1, if the notion of the sonority reversal is not taken into account), while /s/ + stop coda clusters have a sonority value of 1. That is, there is not much difference between /s/ + stop onset clusters and /s/ + stop coda clusters when only based on the sonority distance between these two clusters (Actually /s/ + stop onset clusters are worse than /s/ + stop coda clusters if the sonority reversal is taken into account). Again, the researchers who want to adhere to the sonority-based model can entertain the notion of “extrasyllabicity”, yet they have nothing to offer other than postulating the abstract notion of extrasyllabicity to handle the data at hand.

From the perspective of PCCAM, the relative unmarkedness of /s/ + stop onset clusters over /s/ + stop coda clusters is apparent. /s/ + stop onset clusters become the most favored onset clusters as a function of the articulatory stridency effect, the perceptual stridency effect, and maximal perceptual contrast: all of these three phonetic factors favor /s/ + stop onset clusters. In contrast, /s/ + stop coda clusters are disfavored by the principle of perceptual saliency, which predicts that a stop in /s/ + stop coda clusters is the target of consonant deletion and vowel insertion (after a stop).

It was shown in Chapter 6 that the predictions of PCCAM on the markedness of /s/ + stop coda clusters hold true in Korean speakers’ production of coda clusters. Even more interesting is that as was predicted by the principle of perceptual saliency, the major error types associated with /s/ + stop coda clusters (in Korean speakers’ production) were vowel insertion after a stop and stop deletion. Given this observation, we can say that PCCAM enables us to connect the type of consonant clusters with repair strategies

associated with these clusters: vowel insertion occurs frequently after a stop in /s/ + stop coda clusters as a function of the principle of perceptual saliency. In contrast, vowel insertion is least likely to occur after /s/ both in /s/ + stop onset and coda clusters because the articulatory stridency effect is strongest in this sequence.

As was shown in the cases with a) stop + /s/ and /s/ + stop coda clusters and b) /s/ + stop onset clusters and /s/ + stop coda clusters, PCCAM provides efficient accounts for the asymmetric patterns which have resisted a satisfactory account under the sonority-based consonant cluster acquisition model.

Nevertheless, I believe that the principal advantage of PCCAM is that since the universal characteristics of phonetic factors are identified in PCCAM, the model has the potential to explain consonant cluster acquisition data from a variety of L1s and L2s. In other words, PCCAM provides a good testing ground for the existence of the universal aspects of articulation and perception shared by human beings. This dissertation provided a test case for the existence of such commonalities. By the same token, PCCAM enables unified understanding of vowel insertion, consonant deletion, and metathesis processes found in the acquisition of consonant clusters.

There are several areas where future research is needed. First of all, the present study did not investigate the weighting of each articulatory and perceptual factor in determining Korean speakers' performance on English consonant clusters. To understand the complex array of articulatory and perceptual factors (along with other factors such as task types, frequency of consonant clusters, and syllable position effects that have not been discussed in the present study) that may influence L2 learners' performance, it is necessary to use multivariate analysis. I believe that future research certainly benefits

from implementing multivariate analysis in examining L2 learners' performance on non-native consonant clusters.

More importantly, PCCAM in its present form cannot account for why voiceless strident + voiceless stop sequences (sp, st, sk) induced more errors in Korean speakers' production than voiceless stop + voiceless stop sequences (pt, kt), contrary to the model's prediction.

In addition, future research needs to answer questions such as a) what is the exact mechanism of the interaction between articulatory and perceptual factors in the patterns of vowel insertion and consonant deletion? b) what is the exact mechanism of the articulatory voicing effect?, c) what is the correct ranking relationship between sonorant + voiced stop and liquid + nasal (which is left undetermined in the present study)?, and d) what are the phonetic factors defining the ranking relationship between these two clusters (sonorant + voiced stop vs. liquid + nasal)?

Future research should be done to elaborate the articulatory and perceptual factors in the *Phonetically Based Consonant Cluster Acquisition Model*. In addition, more empirical work is certainly needed to test the validity of PCCAM as a predictor of L2 learners' performance on non-native consonant clusters.

APPENDIX A

Answer Sheet for the Auditory Similarity Judgment Experiment

Direction:

- ▶ 청취자는 지금부터 영어를 토대로 만들어진 2 쌍의 단어를 듣게 됩니다 (e.g., A-B vs. C-D).
- ▶ B 와 D 는 A 와 C 에 약간씩 변형을 준 단어입니다
(e.g., A (brin) – B (bren) vs. C (stin)- D (sutin).
- ▶ 답안지에는 단어에 변형을 가하기 전의 단계인 A 와 C 단어가 제시되어 있고 각 쌍은 2 번 반복 됩니다.
- ▶ 각 쌍의 단어들을 듣고 A 와 B 쌍이 서로 더 비슷한지 C 와 D 쌍이 서로 더 비슷한지 결정하셔서 더 비슷하다고 생각되는 쌍에 동그라미 표시를 하세요.

1. You will hear two pairs of nonsense words (e.g., A-B vs. C-D)
2. The word B is modified based on word A, and the word D is modified based on word C. (e.g., A (brin) – B (bren) vs. C (stin)- D (sutin).
3. In the answer sheet, you will find the original words A and C, but modified words B and D are not presented.
4. Listen to the two pairs of words carefully and decide which pair is more similar to each other.
5. Each item is repeated twice.
6. Circle the word that you think is more similar to each other.

PRE-TEST

1)	brin	stin
2)	spun	brun
3)	clen	slen

TEST

01)	dren	pren
02)	nigd	nist
03)	nekt	nesp
04)	bren	cren

05)	negz	neks
06)	dren	sten
07)	nigd	nisp
08)	negd	negd
09)	spun	brun
10)	nusp	nugd
11)	brin	stin
12)	nest	nebd
13)	plin	glin
14)	bren	bren
15)	clun	drun
16)	nebz	neps
17)	slin	grin
18)	nesk	neps
19)	snun	brun
20)	nigz	niks
21)	dren	dren
22)	nest	nept
23)	negd	nekt
24)	brin	brin
25)	nugd	nukt
26)	snun	plun
27)	nikt	nisp
28)	prin	stin
29)	nubd	nupt
30)	bren	sken
31)	negd	nesp
32)	nisk	nibz
33)	smen	clen
34)	nesk	nebz
35)	slin	crin
36)	pren	sken
37)	nusp	nukt
38)	smen	glen
39)	nubz	nust
40)	tren	sten
41)	nisk	nips
42)	nikt	nigd
43)	spin	drin
44)	nikt	nist
45)	spun	prun
46)	nups	nust
47)	nubz	nubz
48)	spin	trin

Note that the five items in the list (bren-bren, negd-negd, dren-dren, brin-brin, nubz-nubz) were included in the test items to see the effect of the identity of an epenthetic vowel on the perceptual similarity judgment. The result of this item is not relevant to the topic of this dissertation and thus was not reported.

APPENDIX B

Questionnaire for the Production Experiment

(1) 이름 (Name): 한글 (Korean): _____

영어(English): _____

(2) 성별 (Sex): 남자 (Male) _____ 여자 (Female) _____

(3) 학년 (Academic Status): 1 학년(Freshman) _____ 2 학년(Sophomore) _____
3 학년(Junior) _____ 4 학년 (Senior) _____
대학원생 (Graduate Student) _____

(4) 전공 (Major) _____

(5) 출생연월 (Date of Birth): _____ 년 (year) _____ 월 (month)

(6) 태어난 곳 (Place of Birth): _____

시간순으로 지금까지 살았던 도시 (In chronological order, list the cities you have lived so far.)

(7) 미국에 체재한 기간 (Period of stay in the U.S.): _____ 년 (year(s)) _____ 개월 (months)

(8) 영어 공부한 햇수 (Period of studying English): _____ 년 (years)

(9) 영어이외에 외국어를 공부한 적이 있습니까? (Have you ever learned foreign languages other than English?) 예 (Yes) _____ 아니오 (No) _____

(9)에 대한 대답이 “예”라면 (9-1)-(9-3)에 대답해주세요. (If “yes” in (9), answer (9-1)-(9-3))

(9-1) 어떤 외국어들을 배웠나요? (Specify foreign languages you have learned.)

1) _____ 2) _____ 3) _____

(9-2) 언제 그 언어를 배우기 시작했나요? (When did you begin to learn the language(s)?)

5 살 이전 _____ (Before 5 years old) 5-10 세 사이 _____ (5 - 10 years old)
10-15 세 사이 _____ (10 - 15 years old) 15-20 세 사이 _____ (15 - 20 years old)
20 세 이후 _____ (After 20 years old)

(9-3) 얼마나 그 언어를 배웠나요? (How long have you studied the language(s)?)

_____ 년 (year(s)) _____ 개월 (months)

_____ 년 (year(s)) _____ 개월 (months)

_____ 년 (year(s)) _____ 개월 (months)

(10) 영어 음운론이나 발음학을 공부해 본 적이 있습니까?

(Have you ever studied English phonology or pronunciation?): 예 (Yes) _____ 아니오 (No) _____

만약 “예” 라면, 언제 어디서 얼마동안? (If yes, when, where and how long?)

(11) 영어회화 능력에 대한 자기 평가 (Self-evaluation on the English-speaking proficiency)

상 (advanced) _____ 상중 (high-middle) _____ 중 (middle) _____

중하 (middle-low) _____ 하 (low) _____

(12) 1 일 평균 영어 사용 시간 (listening and speaking) (Average time of English use per day)

1 시간 미만 _____ 1-2 시간 _____ 3-4 시간 _____ 5-6 시간 _____ 6-7 시간 _____

7 시간 이상 _____

(13) 영어 회화 수업을 들어본 적이 있습니까? (Have you ever taken English classes focusing on speaking skills?) 예 (Yes) _____ 아니오 (No) _____

만약 “예” 라면, 언제 어디서 얼마동안? (If yes, when, where and how long?)

(14) 부모님의 방언 (Parents' dialect):

아버지 (Father) _____ 어머니 (Mother) _____

(15) 건강 (Health):

말하기 장애 (Speech problems): 예 (Yes) _____ 아니오 (No) _____

듣기 장애 (Hearing problems): 예 (Yes) _____ 아니오 (No) _____

빠진 이 (Missing teeth): 예 (Yes) _____ 아니오 (No) _____

감기 (Cold):

예 (Yes) _____ 아니오 (No) _____

후두절개수술 (Laryngotomy):

예 (Yes) _____ 아니오 (No) _____

편도절개수술 (Tonsilectomy):

예 (Yes) _____ 아니오 (No) _____

APPENDIX C

Word List for the Production Experiment

Cluster Type	Attested Word	Target Word
voiced stop + liquid	blame	blen/blet
voiced stop + liquid		blon/blot
voiced stop + liquid	brown	bren/bret
voiced stop + liquid		bron/brot
voiced stop + liquid	drain	dren/dret
voiced stop + liquid		dron/drot
voiced stop + liquid	glad	glen/glet
voiced stop + liquid		glon/glot
voiced stop + liquid	green	gren/gret
voiced stop + liquid		gron/grot
voiceless stop + liquid	claim	clen/clet
voiceless stop + liquid		clon/clot
voiceless stop + liquid	cream	cren/cret
voiceless stop + liquid		cron/crot
voiceless stop + liquid	play	plen/plet
voiceless stop + liquid		plon/plot
voiceless stop + liquid	prawn	pren/pret
voiceless stop + liquid		pron/prot
voiceless stop + liquid	train	tren/tret
voiceless stop + liquid		tron/trot
voiceless strident + voiceless stop	sphere	spen/spet
voiceless strident + voiceless stop		spen/spot
voiceless strident + voiceless stop	star	sten/stet
voiceless strident + voiceless stop		ston/stot
voiceless strident + voiceless stop	sky	sken/sket
voiceless strident + voiceless stop		skon/skot
voiceless strident + liquid	slide	slen/slet
voiceless strident + liquid		slon/slot
voiceless strident + nasal	smile	smen/smet
voiceless strident + nasal		smon/smot
voiceless strident + nasal	snow	snen/snet
voiceless strident + nasal		snon/snot
voiced stop + voiced stop	tabed	nebd/nobd
voiced stop + voiced stop		tebd/tebd
voiced stop + voiced stop	taged	negd/nogd
voiced stop + voiced stop		tegd/togd
voiced stop + voiced strident	tabs	nebz/nobz
voiced stop + voiced strident		tebz/tobz
voiced stop + voiced strident	tags	negz/nogz
voiced stop + voiced strident		tegz/togz

voiceless stop + voiceless stop	adopt	nept/nopt
voiceless stop + voiceless stop		tept/topt
voiceless stop + voiceless stop	act	nekt/nokt
voiceless stop + voiceless stop		tekt/tokt
voiceless stop + voiceless strident	lapse	neps/nops
voiceless stop + voiceless strident		teps/tops
voiceless stop + voiceless strident	spitz	nets/nots
voiceless stop + voiceless strident		tets/tots
voiceless stop + voiceless strident	parks	neks/noks
voiceless stop + voiceless strident		teks/toks
voiceless strident + voiceless stop	lisp	nesp/nosp
voiceless strident + voiceless stop		tesp/tosp
voiceless strident + voiceless stop	list	nest/nost
voiceless strident + voiceless stop		test/tost
voiceless strident + voiceless stop	risk	nesk/nosk
voiceless strident + voiceless stop		tesk/tosk
voiced strident + voiced stop	cruised	nezd/nozd
voiced strident + voiced stop		tezd/tozd
liquid + voiced stop	bulb	nelb/nolb
liquid + voiced stop		telb/told
liquid + voiced stop	held	neld/nold
liquid + voiced stop		teld/told
liquid + voiceless stop	help	nelp/nolp
liquid + voiceless stop		telp/tolp
liquid + voiceless stop	belt	nelt/nolt
liquid + voiceless stop		telt/tolt
liquid + voiceless stop	work	nelk/nolk
liquid + voiceless stop		telk/tolk
liquid + voiceless strident	pulse	nels/nols
liquid + voiceless strident		tels/tols
nasal + voiced stop	tend	nend/nond
nasal + voiced stop		tend/tond
nasal + voiceless stop	ant	nent/nont
nasal + voiceless stop		tent/tont
nasal + voiceless stop	lamp	temp/tomp
nasal + voiceless stop		nemp/nomp
nasal + voiceless strident	rinse	nens/nons
nasal + voiceless strident		tens/tons
nasal + voiced strident	lens	nenz/nonz
nasal + voiced strident		tenz/tonz
liquid + nasal	film	nelm/nolm
liquid + nasal		telm/tolm

Fillers

Target Clusters	Attested Word	Target Word
__sp__	whisper	espo/ospe
__st__	master	esto/oste
__sk__	whiskey	esko/oske
__pr__	april	epro/opre
__tr__	citrus	etro/otre
__dr__	audrey	edro/odre
__kr__	sacred	ekro/okre
__bl__	problem	eblo/oble
__kl__	necklace	eklo/okle
__kw__	awkward	ekwo/okwe
__dw__	edward	edwo/odwe
__pt__	captain	epto/opte
__ps__	tipsy	epso/opse
__kt__	doctor	ekto/okte
__nd__	candy	endo/onde
__mp__	temper	empo/ompe
__ns__	rancid	enso/onse
__ld__	seldom	eldo/olde
__lm__	helmet	elmo/olme
__nd__	condemn	endo/onde
__mp__	impel	empo/ompe
__ns__	insist	enso/onse
__gb__	rugby	egbo/ogbe
__sb__	husband	esbo/osbe
__bt__	obtain	ebto/obte
__bs__	subside	ebso/obse
__pk__	napkin	epko/opke
__pt__	option	epto/opte
__km__	acme	ekmo/okme
__kn__	acne	ekno/okne
__gm__	dogma	egmo/ogne
__gn__	magnet	egno/ogne
__dn__	kidney	edno/odne
__tn__	whitney	etno/otne
__dl__	peddler	edlo/odle
__tl__	butler	etlo/otle
__mb__	timber	embo/ombe
__nv__	denver	envo/onve
__mz__	clumsy	emzo/omze
__nm__	enmity	enmo/onme
__mn__	amnesia	emno/omne
__nl__	only	enlo/onle
__nr__	henry	enro/onre
__lr__	already	elro/olre
__lw__	always	elwo/olwe

REFERENCES

- Ahn, Sang-Cheol. 1985. The interplay of phonology and morphology in Korean. Ph.D dissertation, University of Illinois.
- Anderson, Janet I. 1987. The markedness differential hypothesis and syllable structure difficulty. *Interlanguage phonology: The acquisition of a second language sound system*, ed. by Georgette Ioup and Steven Weinberger, 279-91. Cambridge, MA: Newbury House.
- Archibald, John. 1998. *Second language phonology* (Language acquisition & language disorders, v. 17). Amsterdam; Philadelphia, Pa.: J. Benjamins Pub. Co. xi, 313 p.
- Archibald, John. 2003. Learning to parse second language consonant clusters. *The Canadian Journal of Linguistics/La Revue canadienne de Linguistique* 48.3-4.149-77.
- Barlow, J. 1997. A constraint-based account of syllable onsets: Evidence from developing systems. Ph.D dissertation, Indiana University.
- Bayley, Robert. 1996. Competing constraints on variation in the speech of adult Chinese. *Second language acquisition and linguistic variation*, ed. by Robert Bayley and Dannis R. Preston, 97-120. Amsterdam: John Benjamins.
- Beckman, Mary E. 1994. When is a syllable not a syllable? *OSU working papers in Linguistics* 44.50-69.
- Bernhardt, B. H., and J. P. Stemberger. 1998. *Handbook of phonological development*. New York: Academic Press.
- Bladon, Anthony. 1986. Phonetics for hearers. *Language for hearers* (Language & communication library), ed. by Graham McGregor, 1-24. Oxford [Oxfordshire]: Pergamon.
- Blevins, Juliette. 1995. The syllable in phonological theory. *A handbook of phonological theory*, ed. by J. Goldsmith, 206-44. Oxford: Blackwell.
- Blevins, Juliette. 2004. *Evolutionary phonology: the emergence of sound patterns*. Cambridge; New York: Cambridge University Press. xix, 366 p.
- Blevins, Juliette, and Andrew Garrett. 1998. The origins of consonant-vowel metathesis. *Language* 74.508-56.
- Bloch, Bernard, and George L. Trager. 1942. *Outline of linguistic analysis* (Special publications of the Linguistic Society of America). Baltimore, Md.: Published by

Linguistic Society of America at the Waverly Press. 82 p.

- Boersma, Paul. 1998. *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.
- Boersma, Paul. 2000. Learning a grammar in functional phonology. *Optimality theory: Phonology, syntax, and acquisition* (Oxford linguistics), ed. by Joost Dekkers, Frank Reinoud Hugo van der Leeuw, and Jeroen Maarten van de Weijer, 465-523. Oxford; New York: Oxford University Press.
- Boersma, Paul; Joost Dekkers; and Jeroen Maarten van de Weijer. 2000. Introduction-Optimality theory: Phonology, syntax, and acquisition. *Optimality theory: Phonology, syntax, and acquisition* (Oxford linguistics), ed. by Joost Dekkers, Frank Reinoud Hugo van der Leeuw, and Jeroen Maarten van de Weijer. Oxford; New York: Oxford University Press.
- Bortolini, Umberta, and Laurence B. Leonard. 1991. The speech of phonologically disordered children acquiring Italian. *Clinical Linguistics & Phonetics* 5.1.1-12.
- Broselow, Ellen, and Daniel Finer. 1991. Parameter setting in second language phonology and syntax. *Second Language Research* 7.1.35-59.
- Browman, Catherine P., and Louis Goldstein. 1986. Towards an articulatory phonology. *Phonology Yearbook* 3.219-52.
- Browman, Catherine P., and Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology* 6.2.201-51.
- Browman, Catherine P., and Louis Goldstein. 1990a. Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics* 18.3.299-320.
- Browman, Catherine P. , and Louis Goldstein. 1990b. Tiers in articulatory phonology, with some implications for casual speech. *Papers in laboratory phonology I: Between the grammar and the physics of speech*, ed. by John Kingston and Mary E. Beckman, 341 – 76. Cambridge: Cambridge University Press.
- Browman, Catherine P., and Louis Goldstein. 1992. "Targetless" schwa: An articulatory analysis. *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, ed. by Gerard Docherty and D. Robert Ladd. Cambridge: Cambridge University Press.
- Butt, Matthias. 1992. Sonority and the Explanation of Syllable Structure. *Linguistische Berichte*, 45-67: Westdeutcher Verlag.
- Byrd, Dani. 1996. Influences on articulatory timing in consonant sequences. *Journal of Phonetics* 24.2.209-44.

- Carlisle, Robert S. 1998. The acquisition of onsets in a markedness relationship: A longitudinal study. *Studies in Second Language Acquisition* 20.2.245-60.
- Catford, J. C. 1977. *Fundamental problems in phonetics*. Bloomington: Indiana University Press. 278 p.
- Chen, Matthew Y., and William S. Y. Wang. 1975. Sound change: Actuation and implementation. *Language* 51.2.255-81.
- Chen, Szu-Wei. 2003. Acquisition of English onset clusters by Chinese learners in Taiwan. Postgraduate Conference 2003. University of Edinburgh
- Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York,: Harper & Row. xiv, 470 p.
- Clements, G. Nick. 1990. The role of the sonority cycle in core syllabification. *Papers in laboratory phonology I*, ed. by J. Kingston and M. Beckman, 283-333. Cambridge: Cambridge University Press.
- Clements, G. Nick. 1992. The sonority cycle and syllable organization. *Phonologica*, ed. by Wolfgang Dressler, H. Luschutzky, O. Pfeiffer, and J. Rennison, 63-76. Cambridge: Cambridge University Press.
- Collier, Rene; Leigh Lisker; Hajime Hirose; and Tatsujiro Ushijima. 1978. Voicing in intervocalic stops and fricatives in Dutch. *Journal of Phonetics* 7.357-73.
- Corder, S. P. 1971. Idiosyncratic dialects and error analysis. *International review of applied linguistics in language teaching (IRAL)* IX/2.147-60.
- Côté, Marie-Hélène. 2000. Consonant cluster phonotactics: A perceptual approach. Ph.D dissertation, MIT.
- Côté, Marie-Hélène. 2004. Syntagmatic distinctness in consonant deletion. *Phonology* 21.1-41.
- Davidson, Lisa. 2001. Hidden rankings in the final state of the English grammar. *RuLing Papers II*, ed. by G. Horwood and S.-K. Kim, 21-48. New Brunswick: Rutgers University.
- Davidson, Lisa. 2003. The atoms of phonological representation: Gestures, coordination, and perceptual features in consonant cluster phonotactics. Ph.D dissertation, Johns Hopkins University.
- Davidson, Lisa. 2006. Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics* 34.104-37.

- Davidson, Lisa; Peter Jusczyk; and Paul Smolensky. 2004. The initial and final states: Theoretical implications and experimental explorations of richness of the base. *Constraints in phonological acquisition.*, ed. by Rene Kager, Wim Zonneveld, and Joe Pater, 791-839. Cambridge: Cambridge University Press.
- Davis, Stuart. 1990. An argument for the underspecification of [coronal] in English. *Linguistic Inquiry* 21.2.301-06.
- Dewey, Godfrey. 1923. *Relativ frequency of English speech sounds*. Cambridge: Harvard University Press. xii, 3-148 p. incl. tables.
- Dyson, Alice Tanner. 1988. Phonetic inventories of 2-and 3-year-old children. *Journal of Speech and Hearing Disorders* 53.89-93.
- Dyson, Alice Tanner, and E. P. Paden. 1983. Some phonological acquisition strategies used by two-year-olds. *Journal of Childhood Communication Disorders* 7.6-18.
- Eckman, Fred R. 1991. The structural conformity hypothesis and the acquisition of consonant clusters in the interlanguage of ESL learners. *Studies in Second Language Acquisition*, 13.1.23-41.
- Eckman, Fred R., and Gregory K. Iverson. 1993. Sonority and markedness among onset clusters in the interlanguage of ESL learners. *Second Language Research* 9.3.234-52.
- Farris, Ashley W. 2005. Gapped acquisition of harmonically complete inventories: An analysis of children's cluster acquisition. Paper presented at Mid-continental Workshop on Phonology 11, Nov 4-6 2005. University of Michigan, Ann Arbor MI
- Fikkert, Johanna Paula Monique. 1994. On the acquisition of prosodic structure. Thesis (doctoral), Holland Academic Graphics. 358 p.
- Flege, James Emil; Naoyuki Takagi; and Virginia Mann. 1996. Lexical familiarity and English-language experience affect Japanese adults' perception of /r/ and /l/. *The Journal of the Acoustical Society of America* 99.2.1161-73.
- Fleischhacker, Heidi. 2001. Cluster-dependent epenthesis asymmetries. UCLA working papers in linguistics 7: Papers in phonology 5.
- Fleischhacker, Heidi. 2005. Similarity in phonology: Evidence from reduplication and loan adaptation. Ph.D dissertation, UCLA.
- Freitas, Maria Joao, and Maria Celeste Rodrigues. 2003. On the nature of sC-clusters in European Portuguese. *Journal of Portuguese Linguistics* 2.2.55-85.

- French, N. R; C.W. Carter; and W. Koenig. 1930. The words and sounds of telephone conversation. *Bell Systems Technical Journal* 9.290-324.
- Fujimoto, Masako. 2004. Effects of consonant type and syllable position within a word on vowel devoicing in Japanese. *Speech Prosody-2004*. 625-28. Retrieved from http://www.isca-speech.org/archive/sp2004/sp04_625.pdf
- Gafos, Adamantios I. 2002. A grammar of gestural coordination. *Natural Language & Linguistic Theory* 20.2.269-337.
- Gierut, Judith A. 1999. Syllable onsets: Clusters and adjuncts in acquisition. *Journal of Speech, Language, and Hearing Research* 42.3.708-26.
- Gnanadesikan, A. E. 2004. Markedness and faithfulness constraints in child phonology. *Constraints in phonological acquisition*, ed. by René Kager, Joe Pater, and Wim Zonneveld, 73-108. Cambridge, UK: Cambridge University Press.
- Goad, Heather, and Yvan Rose. 2004. Input elaboration, head faithfulness and evidence for representation in the acquisition of left-edge clusters in West Germanic. *Constraints in phonological acquisition*, ed. by René Kager and Wim Zonneveld, 109-57. Cambridge: Cambridge University Press.
- Goldsmith, John A. 1995. *The handbook of phonological theory* (Blackwell handbooks in linguistics ; 1). Cambridge, Mass., USA: Blackwell. xiv, 986 p.
- Greenberg, Joseph, and J. J. Jenkins. 1964. Studies in the psychological correlates of the sound system of American English. *Word* 20.157-77.
- Guy, G. R. 1980. Variation in the group and the individual: the case of final stop deletion. *Locating language in time and space*, ed. by William Labov, 1-36. New York: Academic Press.
- Hancin-Bhatt, Barbara. 2000. Optimality in second language phonology: Codas in Thai ESL. *Second Language Research* 16.3.201-32.
- Hancin-Bhatt, Barbara, and Rakesh M. Bhatt. 1997. Optimal L2 syllables: Interactions of transfer and developmental effects. *Studies in Second Language Acquisition* 19.3.331-78.
- Hansen, Jette Gjaldbaek. 2000. Social and linguistic constraints on the acquisition of an L2 phonology: A case study. Ph.D dissertation, Ohio State University.
- Hansen, Jette Gjaldbaek. 2001. Linguistic constraints on the acquisition of English syllable codas by native speakers of Mandarin Chinese. *Applied Linguistics* 22.3.338-65.

- Harris, James W. 1983. *Syllable structure and stress in Spanish: A nonlinear analysis* (Linguistic inquiry monographs; 8). Cambridge, Mass.: MIT Press. vii, 158 p.
- Harris, John. 1994. *English sound structure*. Oxford, UK ; Cambridge, Mass.: Blackwell. xiii, 316 p.
- Hayes, Bruce. 1999. Phonetically-driven phonology: The role of Optimality Theory and inductive grounding (Available on Rutgers Optimality Archive, ROA-158). *Functionalism and formalism in linguistics, Vol.1: General papers*, ed. by Mike Darnell, E. Moravcsik, M. Noonan, F. Newmeyer, and K. Wheatly, 243-85. Amsterdam ; Philadelphia: J. Benjamins.
- Hayes, Bruce. 2003. Phonetically-driven phonology: The role of Optimality Theory and inductive grounding. *Optimality theory in phonology: A reader*, ed. by John J. McCarthy, 290-309. Malden, MA, USA: Blackwell Pub.
- Hayes, Bruce, and Donca Steriade. 2004. The phonetic bases of phonological markedness. *Phonetically-based phonology*, ed. by Bruce Hayes and Donca Steriade. New York: Cambridge University Press.
- Henderson, J. B., and B. H. Repp. 1982. Is a stop consonant released when followed by another stop consonant? *Phonetica* 39.71-82.
- Hirose, H., and T. Gay. 1972. The activity of the intrinsic laryngeal muscles in voicing control: An electromyographic study. *Phonetica* 25.3.140-64.
- Hooper, Joan B. 1976. *An introduction to natural generative phonology*. New York: Academic Press. xvi, 254 p.
- Hudson, Grover. 1995. Consonant release and the syllable. *Linguistics* 33.655-72.
- Hume, Elizabeth. 2002. Predicting metathesis: The ambiguity/attestation model. *ROA*.546.
- Imai, Terumi. 2004. Vowel devoicing in Tokyo Japanese: A variationist approach, Michigan State University.
- Ingram, D. 1974. Phonological rules in young children. *Journal of Child Language* 1.233-41.
- Ingram, D. 1989. *Phonological disability in children*. London: Whurr Publishers, Ltd.
- Jakielski, Kathy Jo. 1998. Motor organization in the acquisition of consonant clusters. Ph.D dissertation, University of Texas at Austin.
- Jun, Jongho. 2004. Place assimilation. *Phonetically-based phonology*, ed. by Bruce Hayes, Robert Martin Kirchner, and Donca Steriade, 58-86. New York:

Cambridge University Press.

- Jun, Sun.-Ah., and M. E Beckman. 1993. A gestural-overlap analysis of vowel devoicing in Japanese and Korean. Paper presented at the 1993 Annual Meeting of the Linguistic Society of America. 7-10 January, Los Angeles, CA, USA
- Jun, Sun-Ah. 1993. The phonetics and phonology of Korean prosody. Ph.D. dissertation, The Ohio State University.
- Jun, Sun-Ah; Mary E. Beckman; and Hyuck-Joon Lee. 1998. Fiberscopic Evidence for the influence on vowel devoicing of the glottal configurations for Korean obstruents. *University of California Working Papers in Phonetics* 96.43-68.
- Jun, Sun-Ah; Mary E. Beckman; Seiji Niimi; and Mark Tiede. 1998. Electromyographic evidence for a gestural-overlap analysis of vowel devoicing in Korean. *University of California Working Papers in Phonetics* 96.1-42.
- Kabak, Baris. 2003. The perceptual processing of second language consonant clusters. Ph.D dissertation, University of Delaware.
- Kager, René. 1999. *Optimality theory* (Cambridge textbooks in linguistics). Cambridge, U.K.; New York: Cambridge University Press. xiii, 452 p.
- Kawasaki, Haruko. 1982. The acoustical bases for universal constraints on sound sequences. Ph.D dissertation, University of California, Berkeley.
- Kaye, Jonathan. 1989. *Phonology: A cognitive view* (Tutorial essays in cognitive science). Hillsdale, N.J.: L. Erlbaum Associates. xi, 172 p.
- Kim, Chin W., and Hyoung-Youb Kim. 1991. The Character of Korean glides. *Studies in the Linguistic Sciences* 21.2.113-25.
- Kim, H-Y. 1990. Voicing and tensification in Korean: A multi-face approach. Ph.D dissertation, University of Illinois.
- Kim, Hyouk-Keun. 2000. The interlanguage phonology of Korean learners of English: A computational implementation of optimality theoretic constraints. Ph.D dissertation, Georgetown University.
- Kim, Hyunsoon, and Allard Jongman. 1996. Acoustic and perceptual evidence for complete neutralization of manner of articulation in Korean. *Journal of Phonetics* 24.295-312.
- King, R. D. 1976. *The history of final devoicing in Yiddish*. Bloomington: Indiana University Linguistics Club.

- Kirk, Cecilia, and Katherine Demuth. 2003. Onset/coda asymmetries in the acquisition of clusters. *Proceedings of the Annual Boston University Conference on Language Development* 27.2.437-48.
- Kuehn, D. P., and K. L. Moll. 1976. A cineradiographic study of VC and CV articulatory velocities. *Journal of Phonetics* 4.303-20.
- Kwon, Bo-Young. 2005. The patterns of vowel insertion in IL phonology: The P-map account. *Studies in Phonetics, Phonology and Morphology* 11.2.21-49.
- LaCharité, and Paradis. 2000. Derivational residue: Hidden rules. *Optimality theory: Phonology, syntax, and acquisition* (Oxford linguistics), ed. by Joost Dekkers, Frank Reinoud Hugo van der Leeuw, and Jeroen Maarten van de Weijer, 211-33. Oxford ; New York: Oxford University Press.
- Ladefoged, Peter. 2001. *A course in phonetics*. Fort Worth: Harcourt College Publishers. xiv, 289 p.
- Lee, Iksop, and S. Robert Ramsey. 2000. *The Korean language* (SUNY series in Korean studies). Albany: State University of New York Press. xiii, 374 p.
- Lee, J-S. 1991. Metrical foot and vowel length in verbal morphology in Korean. *Harvard Studies in Korean linguistics* IV.133-42.
- Lee, Yongsung. 1993. Topics in the Vowel Phonology of Korean. Ph.D dissertation, Indiana University.
- Lehmann, Winfred Philipp. 1962. *Historical linguistics: an introduction*. New York: Holt. 297 p.
- Lin, Yuh-Huey. 2001. Syllable simplification strategies: A stylistic perspective. *Language Learning* 51.4.681-718.
- Lin, Yuh-Huey. 2003. Interphonology variability: Sociolinguistic factors affecting L2 simplification strategies. *Applied Linguistics* 24.4.439-64.
- Locke, J. L. 1983. *Phonological acquisition and change*. New York: Academic Press.
- Major, Roy C. 1987. A model for interlanguage phonology. *Interlanguage phonology: The acquisition of a second language sound system*, ed. by Georgette Ioup and Steven Weinberger, 101-24. Cambridge, MA: Newbury House Publishers.
- Major, Roy C. 1996. Markedness in second language acquisition of consonant clusters. *Second language acquisition and linguistic variation*, ed. by Robert Bayley and Dannis R. Preston, 75-96. Amsterdam: John Benjamins.

- Major, Roy C. 1999. Chronological and stylistic aspects of second language acquisition of consonant clusters. *Language Learning* 49.supplement 1.123-50.
- Malécot, A. 1958. The role of releases in the identification of released final stops. *Language* 34.370-80.
- Mattingly, Ignatius G. 1981. Phonetic representation and speech synthesis by rule. *The cognitive representation of speech*, ed. by Terry Myers, John Laver, and John Anderson, 415-20. Amsterdam: North-Holland.
- McCarthy, John J. 1995. Extensions of faithfulness: Rotuman Revisited (ROA-110). Ms. University of Massachusetts, Amherst.
- McCarthy, John J., and Alan S. Prince. 1995. Faithfulness and reduplicative identity. *Papers in Optimality Theory* (University of Massachusetts Occasional Papers 18; ROA 60), 249-384. Amherst Mass: Graduate Linguistic Student Association.
- McLeod, S.; van Doorn J.; and V. A. Reed. 2002. Typological description of the normal acquisition of consonant clusters. *Themes in Clinical Phonetics and Linguistics*, ed. by F. Windsor, L. Kelly, and N. Hewlett, 185 – 200. Hillsdale, NJ: Lawrence Erlbaum.
- Menn, Lise. 1983. Development of articulatory, phonetic, and phonological capabilities. *Language Production: Development, Writing, and Other Language Processes*, ed. by Brian Butterworth, 3-50: Academic Press.
- Mills, Roger E. 1975a. Proto South Sulawesi and Proto Austronesian phonology. Ph.D dissertation, University of Michigan.
- Mills, Roger E. 1975b. The reconstruction of Proto South Sulawesi. *Archipel* 10.205-24.
- Nemser, William. 1971. Approximative systems of foreign language learners. *International review of applied linguistics in language teaching (IRAL)* IX/2.115-23.
- No, Gyeonghee. 1997. Developmental and transfer factors in English pronunciation of Korean Elementary school and college students. *English Teaching* 52.1.21-39.
- Ohala, Diane Kathleen. 1999. The influence of sonority on children's cluster reductions. *Journal of Communication Disorders* 32.6.397-422.
- Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. *The production of speech*, ed. by Peter F. MacNeilage, 189-216. New York: Springer Verlag.
- Ohala, John J. 1990. Alternatives to the sonority hierarchy for explaining segmental

- sequential constraints. *Papers from the Regional Meetings, Chicago Linguistic Society* 2.319-38.
- Ohala, John J. 1994. Speech aerodynamics. *The Encyclopedia of Language and Linguistics*, ed. by R. E. Asher and J. M. Y. Simpson, 4144-48. Oxford: Pergamon.
- Ohala, John J. 1997. Aerodynamics of phonology. *Proceedings of the 4th Seoul International Conference on Linguistics [SICOL]*. Retrieved from <http://trill.berkeley.edu/users/ohala/papers/SEOUL2-aero.pdf>
- Ohala, John J., and Haruko Kawasaki-Fukumori. 1997. Alternatives to the sonority hierarchy for explaining the segmental sequential constraints. *Studies for Einar Haugen*, ed. by S. Eliasson and E. H. Jahr, 343-65. Berlin: Mouton de Gruyter.
- Olmsted, D. L. 1971. *Out of the mouth of babes: Earliest stages in language learning*. The Hague: Mouton. 260 p.
- Osburne, Andrea. 1996. Final cluster reduction in English L2 speech: A case study of a Vietnamese speaker. *Applied Linguistics* 17.2.164-81.
- Pater, Joe, and Jessica A. Barlow. 2003. Constraint conflict in cluster reduction. *Journal of Child Language*, 30.3.487-526.
- Parker, Frank. 1981. A functional-perceptual account of final devoicing. *Journal of Phonetics* 9.129-37.
- Preston, Dennis R. 1989. *Sociolinguistics and second language acquisition*. Oxford, UK; New York, NY, USA: B. Blackwell. xv, 326 p.
- Prince, Alan, and Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar, ms, Rutgers University and University of Colorado at Boulder.
- Rhee, Seok-Chae. 1998. Aspects of release and non-release in phonology. Ph.D dissertation, University of Illinois, UC.
- Roberts, A. Hood. 1965. *A statistical linguistic analysis of American English* (Janua linguarum. Series practica, 8). The Hague,: Mouton. 437 p.
- Ruffner, Peter T. 2000. Epenthesis and fricatives in final consonant clusters. *George Mason University working papers in linguistics*.15-32.
- Sato, Charlene J. 1984. Phonological processes in second language acquisition: Another look at interlanguage syllable structure. *Language Learning* 34.4.43-57.
- Schane, Sanford A. 1972. Natural rules in Phonology. *Linguistic Change and Generative*

- Theory*, ed. by R.P. Stockwell and R.K.S. Macaulay, 199-229: Indiana University Press.
- Scheer, Tobias. 2002. There is no extrasyllabicity. Poster presented at the 10th Manchester Phonology Meeting, 23-25 May 2002. Retrieved from <http://www.unice.fr/dsl/tobweb/papers/posterManch.pdf>
- Greenberg, Joseph, and J. J. Jenkins. 1964. Studies in the psychological correlates of the sound system of American English. *Word* 20.157-77.
- Greenberg, Joseph. 1978. Some generalization concerning initial and final consonant clusters. *Universals of human language: Phonology*, ed. by Joseph Greenberg, 243-79. Stanford, CA: Stanford University Press.
- Selinker, Larry. 1972. Interlanguage. *International review of applied linguistics in language teaching (IRAL)* 10.209-31.
- Selkirk, Elisabeth. 1982. The syllable. *The structure of phonological representation (part III)*, ed. by H. Van der Hulst and N. Smith, 337-84. Dordrecht: Foris Publications.
- Selkirk, Elisabeth. 1984. On the major class features and syllable theory. *Language and sound structure: Studies in phonology presented to Morris Halle by his teacher and students*, ed. by Mark Aronoff and Richard Oerhle. Cambridge, MA: MIT Press.
- Sievers, E. 1881. *Grundzüge der Phonetik*. Leipzig: Breitkopf and Hartel.
- Smit, A. B. 1993. Phonologic error distributions in the Iowa-Nebraska Articulation Norms Project: Word-initial consonant clusters. *Journal of Speech and Hearing Research* 36.931-47.
- Smit, A. B.; L. Hand; J. J. Freilinger; J. E. Bernthal; and A. Bird. 1990. The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders* 55.779-98.
- Smith, N. V. 1973. *The acquisition of phonology: A case study*. Cambridge: Cambridge University Press.
- Sohn, Ho-min. 1999. *The Korean language* (Cambridge language surveys). Cambridge, UK ; New York: Cambridge University Press. xx, 445 p.
- Sohn, Hyang-Sook. 1987. Underspecification in Korean Phonology. Ph.D dissertation, University of Illinois.
- Stampe, David. 1969. The acquisition of phonetic representation. vol. 5. Papers from the Fifth Regional Meeting of the Chicago Linguistic Society. 443-54

- Steriade, Donca. 1982. Greek prosodies and the nature of syllabification. Ph.D dissertation, MIT.
- Steriade, Donca. 1999a. Alternatives to syllable-based accounts of consonantal phonotactics. *Proceedings of the 1998 Linguistics and Phonetics Conference*, ed. by Osamu Fujimura, Brian Joseph, and Bohumil Palek, 205-45. Prague: Karolinum. Retrieved from http://www.linguistics.ucla.edu/people/steriade/papers/Alternatives_to_Syllables.pdf.
- Steriade, Donca. 1999b. Phonetics in phonology: the case of laryngeal neutralization. *UCLA Working Papers in Linguistics 2: Papers in Phonology* 3, 25-146.
- Steriade, Donca. 2001a. Directional asymmetries in place assimilation: A perceptual account. *The role of speech perception in phonology*, ed. by E. Hume and K. Johnson. San Diego: Academic Press.
- Steriade, Donca. 2001b. The phonology of perceptibility effects: the P-map and its consequences for constraint organization. Unpublished manuscript. UCLA.
- Stevens, Kenneth N. 1998. *Acoustic phonetics*. Cambridge, Mass.: MIT Press. viii, 607 p.
- Stoel-Gammon, Carol, and Carla Dunn. 1985. *Normal and disordered phonology in children* (Child language acquisition series). Baltimore: University Park Press. xiv, 238 p.
- Studdert-Kennedy, Michael. 1987. The phoneme as a perceptuomotor structure. *Language perception and production: relationships between listening, speaking, reading, and writing*, ed. by A. Allport, D. MacKay, W. Prinz, and E. Scheerer, 67-84. London; Orlando: Academic Press.
- Tak, Jin-young. 1997. Issues in Korean Consonantal Phonology. Ph.D dissertation, Indiana University.
- Tarone, Elaine E. 1987. Some influences on the syllable structure of interlanguage phonology. *Interlanguage phonology: The acquisition of a second language sound system*, ed. by Georgette Ioup and Steven Weinberger, 232-47. Cambridge, MA: Newbury House.
- Templin, Mildred C. 1957. *Certain language skills in children, their development and interrelationships*. Minneapolis: University of Minnesota Press. xviii, 183 p.
- Tench, Paul. 2001. An applied interlanguage experiment into phonological misperceptions of adult learners. *International Journal of English Studies*, 1.1.257-76.

- Tesar, Bruce, and Paul Smolensky. 1993. The learnability of Optimality Theory: An algorithm and some basic complexity results. Unpublished manuscript. Department of Computer Science & Institute of Cognitive Science, University of Colorado at Boulder. Available as Rutgers Optimality Archive 2, <http://ruccs.rutgers.edu/roa.html>.
- Tesar, Bruce, and Paul Smolensky. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29.2.229-68.
- Tesar, Bruce, and Paul Smolensky. 2000. *Learnability in Optimality Theory*. Cambridge: Massachusetts Instit Technology Press. vi, p. 140.
- Törkenczy, Miklós, and Péter Siptár. 1999. Hungarian syllable structur: arguments for/against complex constituents. *The syllable: Views and facts* (Studies in generative grammar 45), ed. by Harry van der Hulst and Nancy A. Ritter, 249-84. Berlin/New York: Mouton de Gruyter.
- Törkenczy, Miklós, and Péter Siptár. 2000. *The phonology of Hungarian, Phonology of the world's languages*. Oxford / New York: Oxford University Press / Clarendon Press.
- Treiman, Rebecca. 1983. The structure of spoken syllables: Evidence from novel word games. *Cognition* 15.1-3.49-74.
- Treiman, Rebecca; Jennifer Gross; and Annemarie Cwikel-Glavin. 1992. The syllabification of /s/ clusters in English. *Journal of Phonetics* 20.3.383-402.
- Trope, Herbert S. 1987. Sonority as a variability factor in second language phonology. *Sound patterns in second language acquisition*, ed. by Allan James and Jonathan Leather, 173-91. Dordrecht: Foris.
- Umeda, Noriko. 1977. Consonant duration in American English. *Journal of the Acoustical Society of America* 61.3.846-58.
- Vihman, Marilyn May; Marlys A. Macken; Ruth Miller; Hazel Simmons; and Jim Miller. 1985. From babbling to speech: A re-assessment of the continuity issue. *Language* 61.2.397-445.
- Walden, Brian, and Montgomery. 1975. Dimensions of consonant perception in normal and hearing impaired listeners. *Journal of Speech and Hearing Research* 18.445-55.
- Weinberger, Steven. 1987. The influence of linguistic context on syllable simplification. *Interlanguage phonology: The acquisition of a second language sound system*, ed. by Georgette Ioup and Steven Weinberger, 401-17. Cambridge, MA: Newbury

House.

Weinberger, Steven. 1994. Functional and phonetic constraints on second language. *First and second language phonology*, ed. by Mehmet S. Yavas, 283-302. San Diego, Calif.: Singular Pub. Group Inc.

Wellman, Beth Lucy; Ida Mae Case; Ida Gaarder Mengert; and Dorothy Edith Bradbury. 1931. *Speech sounds of young children*. Iowa City: The University. 82 p.

Wilson, Colin. 2001. Consonant cluster neutralisation and targeted constraints. *Phonology* 18.147-97.

Wiltshire, Caroline R. 2005. Word-final consonant and cluster acquisition in Indian English(es). Boston University Conference on Language Development 30. Boston University: Retrived from <http://psy-static01.bu.edu/posters/30/WiltshireBUCLD2005.pdf> 06/13/06

Wright, Richard. 1996. Consonant clusters and cue preservation in Tsou. Ph.D dissertation, University of California, Los Angeles.

Wright, Richard. 2004. A review of perceptual cues and cue robustness. *Phonetically-based phonology*, ed. by Bruce Hayes and Donca Steriade, 34-57. New York: Cambridge University Press.

Yoshioka, Hirohide; Anders Löfqvist; and Hajime Hirose. 1982. Laryngeal adjustments in Japanese voiceless sound production. *Journal of Phonetics* 10.1, Jan.1-10.

Young-Scholten, Martha, and John Archibald. 2000. Second language syllable structure. *Second language acquisition and linguistic theory*, ed. by John Archibald, 64-101. Malden: Blackwell.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 02845 3946