

Systematic Asymmetries in Perception and Production of L2 Inflections in Mandarin L2  
Learners of English: The Effects of Phonotactics, Salience, and Processing Pressure on  
Inflectional Variability

by

Timothy E. Bonner

A dissertation submitted to the Graduate Faculty in Linguistics in partial fulfillment of the  
requirements for the degree of Doctor of Philosophy, The City University of New York

2013

© 2013

Timothy E. Bonner

All Rights Reserved

This manuscript has been read and accepted for the Graduate Faculty in Linguistics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Dr. Gita Martohardjono

Date

Chair of Examining Committee

Dr. Gita Martohardjono

Date

Executive Officer

Dr. Elaine Klein

Dr. Michael Newman

Supervisory Committee

## Abstract

# Systematic Asymmetries in Perception and Production of L2 Inflections in Mandarin L2 Learners of English: The Effects of Phonotactics, Salience, and Processing Pressure on Inflectional Variability

by

Timothy E. Bonner

Advisor: Dr. Gita Martohardjono

The study of language production by adults who are learning a second language (L2) has received a good deal of attention especially when it comes to omission of inflectional morphemes within L2 utterances. Several explanations have been proposed for these inflectional errors. One explanation is that the L2 learner simply does not have the L2 syntactic or prosodic representation in his grammar leading to omission of surface inflections (Hawkins & Liszka, 2003; Goad, White, & Steele, 2003, respectively). Others attribute L2 errors to mapping problems between the lexicon and syntax (Prévost & White, 2000; Lardiere, 1998, 2003). Another potential explanation for the variable production of inflectional endings is that it may be due to performance factors as in Hopp (2009) and Martohardjono, Valian, and Klein (2012) or to “Extra-syntactic” factors as proposed in Klein (2004) or to syllable repairs due to L1 phonotactic interference as proposed in Davidson (2005, 2006a, 2006b). This dissertation claims that when L2 morphosyntactic representations are shown to be available in the L2 learner’s grammar, L2 inflectional variability can be attributed to L1 phonotactic interference, salience of the L2 inflection, and performance factors leading to systematic, but asymmetrical patterns of perception and production of the allomorphs that represent the surface L2 inflections. The results

revealed that the target inflections were not omitted across the board as would be expected under deficit accounts. On the contrary, repairs of the final target coda clusters (i.e., schwa epenthesis before and after the final inflectional consonant and devoicing of the word-final consonants) revealed patterns that are consistent with the degree of syllabicity (e.g., [Vd] vs. [t] and [d]) and sonority (e.g., [s] vs. [t]) of the allomorph or coda and are not indicative of morphosyntactic deficits. Importantly, schwa epenthesis was applied asymmetrically (i.e., particularly to stops [t] and [d]) in clusters that contained target codas and inflectional allomorphs in real, nonce, and monomorphemic items, and thus, this repair pattern is contra the Prosodic Transfer Hypothesis of Goad et al. (2003). Overall, this dissertation presents an alternative explanation for L2 inflectional errors outside of the morphosyntactic and prosodic deficit arguments.

## Acknowledgements

This dissertation is dedicated first and foremost to my loving and deeply supportive wife and partner, Hala, who never failed to lend me her confidence, faith, and trust when my own well of confidence and faith was running on empty.

I also cannot express my true gratitude to my family and friends who have waited patiently for eight long years for me to return to my pre-PhD life. Thank you, especially, mom and dad for your patience and, most importantly, your intense curiosity and desire to learn, which instilled in me that same desire to learn and to always ask, “Why?”

I am indebted to my advisor, friend, and lifelong colleague, Gita Martohardjono, who from the day that I met her always gave me the optimism to continue, even when my notions of research were out in left field. Thank you Gita; you have taught me how to be a true researcher.

I would also like to thank Elaine Klein for her very careful and thoughtful feedback and teaching throughout the dissertation process. You are a consummate educator and a wonderful colleague. Finally, I would like to thank my committee members Kathleen Currie Hall, who taught me a great deal about the very useful and challenging skill of task design and data analysis, and Michael Newman, who contributed a new perspective to my analysis and dissertation even in the eleventh hour. I have learned a great deal from all of you!

## Table of Contents

List of Figures.....	xii
List of Tables.....	xiii
Chapter 1: Introduction.....	1
1.1 Background of the Problem – Summary of Major Proposals	
1.2 Statement of the Problem	
1.3 Goals, Purpose, and Significance of the Current Study	
1.4 Overview of the Dissertation	
Chapter 2: Review of the Literature.....	13
2.1 Mapping Deficits	
2.2 Representational Deficits	
2.3 L1 Phonological Interference on L2 Segmental Structure: The Importance of Perception	
2.4 L1 Phonotactic Constraints and Effects on Perception and Production	
2.5 Variable Perception and the Effects on Performance	
2.6 General Performance-Based Accounts and the Effects on Performance	
2.7 Summary of the Literature	
Chapter 3: Theoretical Background and Research Questions.....	52
3.1 A Model of Competence and Performance	
3.1.1 Parsing L2 Inflection: A Brief Summary	
3.1.2 Determining Grammatical Competence or Knowledge	
3.2 Results of a Pilot Study of Competence and Performance	
3.3 Possible Performance “Deficits” Assuming Underlying Knowledge	

3.3.1 L1 Phonotactic Interference	
3.3.2 Perceptual Saliency of Inflections	
3.3.3 Processing Load Effects on Performance	
3.3.4 L2 Proficiency Level and Overall Performance	
3.4 A Comparison of Mandarin and English: Possible Interference Effects	
3.4.1 Morphological and Morphosyntactic Differences	
3.4.2 Phonological and Phonotactic Differences	
3.4.3 Bound Morphemes in Mandarin and English: A Comparison of Saliency	
3.4.4 L1 Mandarin Speakers' Perception and Production of L2 English	
3.5 Research Questions and Hypotheses	
3.5.1 Research Questions	
3.5.2 Hypotheses	
Chapter 4: Methodology.....	76
4.1 Subjects	
4.2 Materials and Procedures	
4.2.1 Written Grammaticality Screening Task	
4.2.1.1 WGS, Part 1	
4.2.1.2 WGS, Part 2	
4.2.2 Experiment 1: Aural Perception Tasks	
4.2.2.1 Aural Word Perception Task	
4.2.2.2 Aural Sentence Perception Task	
4.2.3 Experiment 2: Oral Production Tasks	
4.2.3.1 Oral Word Production Task	

4.2.3.2 Oral Sentence Production Task	
4.2.4 Cross-Modal Design of Target Items	
4.2.5 General Procedure	
Chapter 5: Results.....	97
5.1 Experiment 1: Aural Word and Sentence Perception Tasks	
5.1.1 Aural Word Task Results – Real and Nonce Items	
5.1.2 Aural Sentence Task Results – Real and Nonce Items	
5.1.3 Sentence Length Effects in Aural Sentence Perception	
5.2 Experiment 2: Oral Word and Sentence Production Tasks	
5.2.1 Analysis of Suppliance of Real Items in the Oral Word Task	
5.2.2 Analysis of Suppliance of Nonce Items in the Oral Word Task	
5.2.3 Analysis of Suppliance of Monomorphemic Items in the Oral Word Task	
5.2.4 Analysis of Suppliance of Real Items in the Oral Sentence Task	
5.2.5 Sentence Length Effects on Suppliance of L2 Inflection in Real Verbs	
5.2.5 Analysis of Suppliance of Nonce Items in the Oral Sentence Task	
5.2.6 Sentence Length Effects on Suppliance of L2 Inflection in Nonce Verbs	
5.3 Experiment 2: Analysis of Repair Strategies in the Oral Tasks	
5.3.1 Devoicing of Real Items in the Oral Word Task	
5.3.2 Devoicing of Nonce Items in the Oral Word Task	
5.3.3 Schwa Epenthesis of Real Items in the Oral Word Task	
5.3.4 Schwa Epenthesis of Nonce Items in the Oral Word Task	
5.3.5 Devoicing Plus Schwa Epenthesis of Real Items in the Oral Word Task	
5.3.6 Devoicing Plus Schwa Epenthesis of Nonce Items in the Oral Word Task	

- 5.3.7 Schwa Epenthesis Totals for Real and Nonce Items in the Oral Word Task
- 5.4 Experiment 2: Analysis of Repair Strategies in the Oral Sentence Task
  - 5.4.1 Devoicing of Real Items in the Oral Sentence Task
  - 5.4.2 Devoicing of Nonce Items in the Oral Sentence Task
  - 5.4.3 Schwa Epenthesis of Real Items in the Oral Sentence Task
  - 5.4.4 Schwa Epenthesis of Nonce Items in the Oral Sentence Task
  - 5.4.5 Schwa Epenthesis Plus Devoicing of Real Items in the Oral Sentence Task
  - 5.4.5 Schwa Epenthesis Plus Devoicing of Nonce Items in the Oral Sentence Task
- 5.5 Analysis of Repairs in Monomorphemic Items
  - 5.5.1 Devoicing in Monomorphemic Items
  - 5.5.2 Schwa Epenthesis in Monomorphemic Items
  - 5.5.3 Schwa Epenthesis Plus Devoicing in Monomorphemic Items
  - 5.5.4 Schwa Epenthesis Totals in Monomorphemic Items
- 5.6 Correlations between Aural Perception and Oral Production
  - 5.6.1 Different Items in Aural Word Task to Suppliance in Oral Word Task
  - 5.6.2 Different, Voiced Items in Aural Word Task to Devoicing in Oral Word Task
  - 5.6.3 Different Items in Aural Word Task to Schwa Epenthesis in Oral Word Task
  - 5.6.4 Different Items in Aural Sentence Task to Suppliance in Oral Sentence Task
  - 5.6.5 Different, Voiced Items in Aural Sentence Task to Devoicing in Oral Sentence Task
  - 5.6.6 Different Items in Aural Sentence Task to Schwa Epenthesis in Oral Sentence Task

5.7 Oral Word Repairs in Real and Nonce Items to Monomorphemic Items	
Chapter 6: General Discussion .....	144
6.1 Evidence of Morphosyntactic Representation	
6.2 Effects of L1 Phonotactics, Saliency, and Processing Load on Perception and Production	
6.2.1 L1 Phonotactic Effects	
6.2.2 Saliency Effects	
6.2.3 Processing Load Effects	
6.3 Repair Strategies in Oral Tasks	
6.3.1 Devoicing Word-Finally: Evidence of Markedness Effects	
6.3.2 Schwa Epenthesis: Evidence Against a Strict Prosodic Transfer Hypothesis	
6.4 Proficiency Level Effects	
6.5 Correlations between Perception and Production	
Chapter 7: Conclusions and Limitations .....	170
7.1 Conclusions	
7.2 Limitations	
Appendix A: Data Tables from Pilot Study.....	175
Appendix B: Pilot Test Items .....	177
References .....	181

## List of Tables

Table 1: Aural Word Perception – Same Items (Real).....	97
Table 2: Aural Word Perception – Different Items (Real).....	98
Table 3: Aural Word Perception – Same Items (Nonce).....	99
Table 4: Aural Word Perception – Different Items (Nonce).....	100
Table 5: Aural Sentence Perception - Same Items (Real).....	101
Table 6: Aural Sentence Perception - Different Items (Real).....	102
Table 7: Aural Sentence Perception – Same Items (Nonce).....	103
Table 8: Aural Sentence Perception –Different Items (Nonce).....	104
Table 9: Aural Sentence Perception - Sentence Length Effects (Lexical).....	106
Table 10: Aural Sentence Perception - Sentence Length Effects (Nonce).....	107
Table 11: Oral Word Production - Target Coda Suppliance (Real).....	109
Table 12: Oral Word Production - Target Coda Suppliance (Nonce).....	110
Table 13: Oral Word Production – Monomorphemic Suppliance.....	111
Table 14: Oral Sentence Production - INFL Suppliance (Real).....	112
Table 15: Oral Sentence Production - Sentence Length Effects (Real).....	113
Table 16: Oral Sentence Production - INFL Suppliance (Nonce).....	114
Table 17: Oral Sentence Production - Sentence Length Effects (Nonce).....	115
Table 18: Oral Word Production - Target Coda Devoicing (Real).....	118
Table 19: Oral Word Production - Target Coda Devoicing (Nonce).....	119
Table 20: Oral Word Production - Target Coda Schwa Epenthesis (Lexical).....	120
Table 21: Oral Word Production - Target Coda Schwa Epenthesis (Nonce).....	121
Table 22: Oral Word Production - Target Coda Devoicing + Schwa Epenthesis (Real).....	122

Table 23: Oral Word Production - Target Coda Devoicing + Schwa Epenthesis (Nonce)...	123
Table 24: Oral Word Production - Target Coda Schwa Epenthesis Totals (Real).....	124
Table 25: Oral Word Production - Target Coda Schwa Epenthesis Totals (Nonce).....	124
Table 26: Oral Sentence Production - INFL Devoicing (Real).....	125
Table 27: Oral Sentence Production - INFL Devoicing (Nonce).....	127
Table 28: Oral Sentence Production - INFL Schwa Epenthesis (Real).....	128
Table 29: Oral Sentence Production - INFL Schwa Epenthesis (Nonce).....	129
Table 30: Oral Sentence Production - INFL Devoicing + Schwa Epenthesis (Real).....	130
Table 31: Oral Sentence Production - INFL Devoicing + Schwa Epenthesis (Nonce).....	131
Table 32: Oral Word Production – Monomorphemic Devoicing.....	133
Table 33: Oral Word Production – Monomorphemic Schwa Epenthesis .....	134
Table 34: Oral Word Production – Monomorphemic Epenthesis + Devoicing .....	135
Table 35: Oral Word Production – Total Monomorphemic Schwa Epenthesis .....	136
Table A1: Overall performance on WGJ Task - mean percentage of correct inflections by group .....	175
Table A2: WGJ Task – number and percent correct by verb type and participant .....	175
Table A3: Aural Task – number and percent correct by verb type and participant .....	175
Table A4: SPEAK Task – number and percent correct by type and participant.....	176

## List of Figures

Figure 1: Competence/Performance Model.....	53
Figure 2: Overview of Task Types and Intended Purpose .....	78

## Chapter 1 – Introduction

The study of second language (L2) production by adults who are learning a second or other language has received a good deal of attention especially when it comes to omission of inflectional morphemes within target language utterances as exemplified in (1) and (2):

(1) \*He sing well.

(2) \*Yesterday, the boy walk home very late.

Several explanations have been proposed for these inflectional errors within the first language (L1) beginning with the claim that there may be input frequency effects leading to a specific *order of acquisition* (Brown, 1973; de Villiers & de Villiers, 1973) to possible brain lateralization effects leading to deficits in a L2 learner's grammar due to a biological *critical period* (Lenneberg, 1967). In fact, since the *Morpheme Order Studies* of the 1970s (cf. Brown, 1973; Dulay & Burt, 1974), the debate over the causes for inflectional acquisition, and the subsequent variability of inflections in second language (L2) learner's grammars, has been extensive and has included various explanations including an inability of the L2 learner to acquire the L2 grammatical forms due to a language critical period to more recent claims that there may be missing morphosyntactic or prosodic features within the L2 grammar that are not acquirable due to restrictions within the L1 grammar.

It is suggested by those in the "representational deficit" camp that the adult L2 learner may be unable to acquire target or native-like competence in the L2 grammatical representation (i.e., either morphosyntactic or prosodic) if the target L2 grammatical structures are not already available in the L1 grammar. In essence, they claim that certain L2 grammatical structures are subject to a critical period. On the other hand, from the perspective of performance, it could be

claimed that if availability of a L2 grammatical representation can be shown to exist in the L2 learner's grammar (e.g., the L2 learner reveals target-like grammatical knowledge of that inflection), inflectional errors are more likely a result of other factors that lie outside of morphosyntactic and prosodic deficits.

This dissertation systematically investigates whether variability of L2 inflections can be attributed to phonotactics, salience, and performance factors leading to reduced accuracy of aural perception and oral production of L2 inflections and, consequently, an increased variability of the inflections in perception and production even when L2 morphosyntactic representations are shown to exist in the L2 learner's grammar. The results of this study provide further evidence against L2 representational deficit accounts for adult L2 learners, while shedding light on the question of what causes L2 inflectional errors in production by providing a systematic analysis of the causes of variable perception and production of inflections for L2 learners of English.

### 1.1 Background of the Problem – Summary of Major Proposals

The L2 representational deficit accounts differ primarily in terms of the module of the grammar that is thought to be the source of variability and can be summarized as follows: Hawkins & Liszka (2003) propose a *syntactic* deficit account (i.e., the Representation Deficit Hypothesis or RDH), which claims that specific syntactic features in the L2 grammatical representation cannot be acquired if they are not already represented in the L1 grammar (e.g., [+/-past]). Goad, White, & Steele (2003) and Goad & White (2006) offer the *prosodic* deficit account (i.e., the Prosodic Deficit Hypothesis or PTH), which claims a mismatch between the L1 and L2 prosodic representation is the reason for L2 inflectional errors. As will be elaborated in

Chapter 2, these proposals make the claim that the L2 grammar is missing key features that negatively influence the native-like production of L2 surface inflections.

*Mapping* accounts propose that variability is the result of mapping problems between the lexicon and syntax; these mapping accounts claim that while syntactic features are acquired within the L2 learner's grammar, mapping the correct features to the syntax is prone to error, resulting in surface forms that diverge from those produced by native speakers. For example, the Missing Surface Inflection Hypothesis (MSIH) explains non-target production via optional default rules that are selected in L2 grammars, but not in native grammars (e.g., Prévost & White, 2000), while Lardiere's Feature re-Assembly Hypothesis (FAH) (e.g., Lardiere, 1998a, 1998b, 2003) explains variability as the re-arrangement of feature clusters relevant in the L1 to suit the needs of the L2. In these mapping deficit accounts, L2 functional features are available to the L2 learner, but the L2 learners fail to map these features to the syntactic nodes, and thus, these features do not materialize in surface forms.

Several researchers have also proposed influences of L1 and L2 phonotactic constraints leading to variable perception and production. For example, Dupoux, Kakehi, Hirose, Pallier, and Mehler (1999) reveal that when a speaker's first language phonotactic constraints disallow a CC cluster (e.g., in Japanese), the speaker will actually perceive an "illusory" vowel inside of the cluster in order to "break up" the illicit cluster. The study determined that these Japanese speakers could not easily discriminate between VCCV and VCuCV stimuli, and the Japanese speakers mapped both of the stimuli to their L1 phonotactic constraints suggesting a fault in their perception of the L2 that was caused by interference of L1 phonotactic constraints. In an attempt to determine how phonotactic constraints are encoded in English—i.e., at the *segmental* (or syllable) level or at the *representational* (or feature) level—Goldrick (2004) used non-word

sequences of CVC syllables to determine if native speakers of English would first exhibit the phonotactic constraints of English in these non-words and, second, would encode phonotactic constraints at the syllable or featural level. Goldrick found that phonotactic constraints are encoded at both the syllable and featural levels and claims that phonotactic constraints are affected both by processing and representational levels. Goldrick's conclusion is an important distinction when examining the effects of L1 phonotactic constraints on perception and production of a L2 such that it underscores the possibility of transfer between L1 and L2, and hence, it could be an area of interference leading to variability of syllable perception and production. In the case of this dissertation, as will be detailed below, L1 (i.e., Mandarin) phonotactic constraints are proposed to have a significant effect on L2 (i.e., English) syllable structure, specifically negatively affecting the word-final consonants that include the target inflections. Finally, in an Optimality Theoretic (OT) analysis of phonotactic constraints, Davidson (2005, 2006a, 2006b) reveals that non-native output can be variable due to differences in phonotactic constraints. As she claims, the repair strategies produced by L2 learners may be reflective of this variability, but by implication in the OT approach, it is not evidence of a lack of representation within the L2 learner's grammar. As she reiterates regarding the tenets of OT (cf. Smolensky, 1999), "The set of possible outputs, and the set of constraints, is the same in all languages; grammars of languages differ only in the way constraints are ranked" (p. 11). Additionally, she claims that variability is proportional to the strength of ranking within the L2 learner's new grammar. Thus, she suggests that the interlanguage does vary and can move towards native-like fluency, and thus, variability and repair strategies are not evidence of a grammatical deficit. These phonotactic accounts reveal that although perception and production errors can occur due to mismatches in phonotactic constraints, these mismatches are not likely

due to representational deficits, but they are more indicative of both performance effects and surface-level phonetic variability.

Variable suppliance of L2 inflections at varying levels of language proficiency has also been linked to perceptual limitations by Solt et al. (2004) and Klein et al. (2004). This research claims that the varying salience (i.e., acoustic perceptibility) of certain English allomorphs for L2 learners' of English could reduce grammatical acquisition and reduce overall perception and production. This investigation of acoustic perceptibility of inflections and varying performance for L2 learners of English provides an additional approach to examine the variability of L2 inflections outside of the representational deficit accounts.

Inflectional variability has also been explained in terms of general performance phenomena. Epstein, Flynn, and Martohardjono (1996) and Klein and Martohardjono (1999) have suggested that general processing mechanisms could be responsible for non-target production of inflections. More recently, Martohardjono, Valian, & Klein (2012)<sup>1</sup> investigate the nature of performance and its effect on L2 inflectional variability and have suggested that target input factors such as frequency, salience, and processing overload may be the primary reasons for errors in production. Martohardjono et al. claim that when syntactic representation can be shown to be available to the L2 learner, inflectional errors are the result of task demands and performance factors rather than deficits in representation. Arguing in particular against syntactic deficit accounts, Martohardjono et al. propose that spontaneous production is not optimal for adducing evidence regarding competence or representation because it involves discourse planning, increased lexical retrieval abilities, and increased processing, all of which present

---

<sup>1</sup> This study in preparation for publication. Results from this study were also presented as a poster at *50 Years of Linguistics at MIT: a scientific reunion* in December, 2011.

considerable difficulty to the L2 learner due to the increased attentional resources and the possibility of exceeding these resources within spontaneous production tasks.

In another account of performance-based effects on L2 inflectional variability, Hopp (2009) provides evidence for native-like attainment of L2 inflections and reveals that variability (i.e., non-target like performance) in L2 inflections can be related to both L1 transfer and limitations in L2 processing efficiency, concluding that surface errors can be attributed to be both L1 interference and performance effects. In an analysis of several studies across L2 learners in various stages of acquisition, Hopp reveals that suppliance (and comprehension) of inflections is not random, but in fact, it is constrained by the L2 grammar and appears more affected by the availability of processing resources than by a lack of grammatical representation. Under a similar generalized performance-based approach to examining L2 inflectional errors, this dissertation claims that even when L2 learners exhibit morphosyntactic knowledge of inflections, performance effects like salience of the input and increased processing load<sup>2</sup> mask this knowledge and explain the variability of L2 surface inflections. Hopp also provides evidence that even when learners are at advanced stages of acquisition, they will continue to exhibit L1 effects under greater processing load and that even native speakers will reveal these same effects on L2 inflections when subjected to high degrees of similar processing pressures that reduce processing resources.

Since the introduction of Baddeley's (1986) model of working memory, several proposals have been introduced to explain mental processes with regards to language processing. Although not used to examine L2 inflectional variability, these proposals are important to set the stage for the larger context of errors in language performance, and can be useful to consider language

---

<sup>2</sup> As will be explained below, the processing load effects under study are generalized to conditions of higher vs. lower task demands and are not of the type examined on-line as in ERP studies (e.g., Felser & Clahsen, 2009).

error data from the perspective of memory demands and other possible non-linguistic causes of surface inflectional variability. In fact, many of the studies on processing load focus on garden-path effects or research the effects of ambiguous sentences to show the effects of syntactic processing. As an example, Gibson's (1998) *locality* theory proposed specific constraints on syntactic processing, an important proposal in the area of sentence processing. However, more pertinent to this dissertation, some recent research has examined the effects of memory load and reduced attentional resources leading specifically to reduced language comprehension and production. For example, Gordon, Hendrick, and Levine (2002) found that syntactically more complex sentences impair sentence comprehension, leading to their conclusion that memory interference can play a large role in sentence comprehension. In another study that examined the effects of the phonological buffer, which is claimed to retain strings of syllables in working memory and is tasked with phonological rechecking of the strings before access of semantic and syntactic levels, Jacquemot, Dupoux, Decouche, and Bachoud-Lévi (2006) found that sentence perception can be affected at the level of phonological analysis prior to accessing the semantic and syntactic levels even in their English native speaking subjects. This result reveals that speech comprehension can be affected by processing factors that occur specifically at the phonological decoding and lexical segmentation levels, thus, revealing that performance effects can be independent of morphosyntactic processing and can even directly affect access to the grammar.

Finally, in an analysis of variability in L2 article production, Trenkic (2007) provides an alternative processing account for L2 variability that goes beyond representational deficit and mapping accounts to claim that the patterning of omissions in L2 production (i.e., in Trenkic's case with English articles) is better explained by a reduction in attentional resources at the conceptual level prior to encoding at the syntactic or prosodic levels. In other words, for L2

learners that would appear to have representational deficits as a result of missing features or missing representations in the L1, asymmetries in article production cannot be explained by these accounts and must be caused by processing limitations related to general cognition and working memory conditions. These recent processing accounts are rooted in what can be classified as deficiencies in performance rather than deficits in the grammar, which is one important distinction that will be examined closely in this dissertation. Chapter 2 of this dissertation details the literature that underpins the above debate and provides alternative proposals that motivate a performance account of L2 inflectional variability.

## 1.2 Statement of the Problem

As is evident in the above summary of the major arguments for L2 inflectional variability, there are many plausible proposals that could account for the variability of inflections. However, there are three possible gaps in the research that need examination and that underpin the purpose of this dissertation. The first issue at hand is with the data collection methods of the major representational deficit accounts, and one that is noted by Klein and Martohardjono (1999) and Martohardjono, Valian, and Klein (2012), which is the unsystematic and uncontrolled nature of the production data that is collected and utilized by some researchers to determine proof of representational and competence deficits. Production data that is collected via spontaneous speech is unreliable when examining discrete inflectional forms because, by its very nature, spontaneous production data is subject to performance effects of the nature that is examined in several studies on performance. The tasks in this dissertation have been designed to systematically control for and test for the effects of performance pressure on inflectional variability. The second issue is the lack of combined analysis of both perception and production

in most studies and the effects of both perception and production on L2 inflectional variability within the same subject group. Only a few studies have examined both perception and production together. This dissertation will more deeply and systematically examine the connection between perception and production and the effects on variation of L2 surface inflections. Finally, the overall purpose of this dissertation is to attempt to reconcile several of the various performance accounts and provide a more comprehensive analysis of L2 inflectional variability from a performance perspective.

### 1.3 Goals, Purpose, and Significance of the Current Study

This dissertation systematically investigates the effects of performance factors while looking at the perception and production of tense and agreement inflections in English by L1 Mandarin Chinese speakers. As will be elaborated further in Chapter 3, Mandarin does not contain inflections for either tense or agreement, while English does. This difference allows for discrete testing of L1 interference, such that if a Mandarin speaker's L2 English morphosyntax can be shown to be intact, other non-morphosyntactic factors can be pinpointed as the potential cause of L2 inflectional errors. Furthermore, the inflections under study occur verb-finally in English, which allows for the investigation of an additional contrast between English and Mandarin: Mandarin has strict phonotactic constraints against word-final consonants except nasals /n/ and /ŋ/ and variably /ʒ/<sup>3</sup> in the syllable coda. As will be explained in Chapter 3, the subjects in this dissertation are L1 speakers of Mandarin, which severely restricts the types of L2 complex syllable codas under investigation. Overall, these restrictions in both the morphosyntax

---

<sup>3</sup> This voiced retroflex fricative is variably produced across the many varieties of Mandarin spoken in China. Nevertheless, The variability (i.e., presence or absence) of the retroflex fricative /ʒ/ across the different varieties of Chinese spoken in Mainland China is not relevant to the questions under investigation in this dissertation, so it was not necessary to control for this variation among the different varieties spoken by the subjects in this study.

and phonology of the L1 grammar compared to the L2 grammar provide a clear distinction to examine L1 effects vs. non-L1 effects in the subjects' perception and production of the L2.

The main question that is under investigation in this dissertation is the following: In speakers where knowledge of English inflectional patterns is established—that is, when morphosyntactic representation can reasonably be assumed to be present in the L2 grammar—what factors influence the aural perception and oral production of these inflections? This dissertation examines this broad question and determines that when morphosyntactic representation is available for L2 learners, the question then becomes how is this morphosyntactic knowledge reflected in perception and production of inflections, such as past tense *-ed* and 3<sup>rd</sup> person singular *-s*, under varying performance conditions and regarding other factors not related to deficits in the morphosyntax or prosody of the target population. As such, this dissertation investigates these effects on perception and production of particular word-final English allomorphic surface variations such as voiced [z] in *reads* vs. unvoiced [s] in *takes* or voiced [d] in *caused* vs. unvoiced [t] in *walked*. The major claim is that variable rates of perception and production of these surface inflections may not necessarily be equated with morphosyntactic, prosodic, or mapping deficits, and thus, the explanations for their surface variability lie outside of these deficit accounts and, in fact, provide counterevidence to these grammatical deficit accounts.

The following general research questions are posed to investigate the possible alternative reasons for L2 inflectional variability:

- 1) First, what role do L1 phonotactic constraints play in the variability of perception and production of English inflections, and is the mismatch between

L1 and L2 phonotactics a possible cause for the well-attested omission of inflections in spontaneous L2 speech?

- 2) Second, what is the effect of salience of English inflectional allomorphs on the perception and production of the corresponding L2 inflections? For example, are syllabic inflections more salient than non-syllabic inflections (e.g., [Vd] vs. [d] in pairs like *wanted* vs. *caused*; [Vz] vs. [z] in pairs like *causes* vs. *reads*)? This dissertation classifies this relative syllabicity in terms the *acoustic perceptibility* of target inflections, which will be elaborated in more detail in Chapters 2 and 3.
- 3) Third, this dissertation considers what non-representational processing factors might account for inflectional variability. For example, would omission of inflections still be a common error when task demands are considerably reduced, as they would be in a controlled production task where lexical selection is predetermined? And would length and complexity of utterances have an effect on the perception and production of inflections in these controlled tasks?
- 4) And finally, if L2 inflection is not subject to a critical period and is attainable as Hopp (2009) claims, and performance effects are the cause of inflectional variability, what would be the effects of proficiency level on perception and production of L2 inflections under varying processing loads?

Taken as a whole, this dissertation investigates several accounts of inflectional variability including L1 phonotactic interference, salience (i.e., acoustic perceptibility), and processing load (i.e., a reduction in processing resources) and attempts to reconcile these accounts under one

study. In sum, it systematically examines the effects on L2 surface inflections that lie in the area of phonotactics, salience, and performance and how these processes affect the perception and production of these inflections leading to surface variability.

#### 1.4 Overview of the Dissertation

Chapter 2 presents the literature and the current state of the debate regarding the causes of L2 surface inflections. Chapter 3 has three purposes as follows: a) it provides the theoretical rationale and construct that underpins the current debate regarding L2 inflectional variability; b) it provides the results of a pilot study conducted by this researcher to provide more evidence against the claim that a morphosyntactic deficit with regards to Mandarin speaking learners of English is the cause of L2 inflectional errors in production; and c) it provides reasoning for the choice of the current subjects under study with regards to critical differences between the L1 (Mandarin) and the L2 (English). Chapter 4 details the research design and methodology. Chapter 5 reveals the results of the study. Chapter 6 provides the general discussion, and Chapter 7 ends with a brief conclusion and implications for future research.

## Chapter 2 – Review of the Literature

The cause of L2 inflectional variability is a hotly debated topic in the field of Second Language Acquisition (SLA). Specifically, the debate rests in four main camps within the field of SLA: 1) *mapping accounts* which claim that there are mapping deficits between the lexicon and other parts of the grammar; 2) *representational deficit* accounts which claim that availability of L2 grammatical representations is lacking; 3) *L1 phonotactic interference* accounts in which mismatches between L1 and L2 phonotactic constraints interfere with perception and production of L2 inflections; and 4) *general processing* and *performance accounts* which claim that L2 inflectional variability can be linked to either memory-load or processing load that reduces attentional resources and/or reduced salience or frequency of the input that reduces performance. The following detailed review of the current literature will examine the major proposals that have been at the center of the debate and will provide an analysis of these proposals, their conclusions, and their methodological shortcomings to provide the background for this study.

### 2.1 Mapping Deficits

In their discussion of UG and adult L2 acquisition, Flynn and Martohardjono (1994) provide an introduction to the mapping deficit account of L2 inflectional variability. They analyze the nature of L2 learner errors via WH-movement principles and claim that even when the L1 (e.g., Indonesian and Chinese) does not instantiate the UG parameter for WH-movement, the L2 learners access UG and, thereby, successfully acquire the appropriate parameter for WH-movement in the L2 (e.g., English). Thus, citing several previous studies, the authors claim that the Indonesian/Chinese L2 learners of English clearly have the ability to recognize movement

violations regardless of the lack of syntactic movement in the L1. In fact, the Indonesian and Chinese L2 learners of English made grammaticality judgments that, as the researchers claim, pattern like those of native speakers. Flynn and Martohardjono (1994) further conclude that L2 acquisition studies reveal that L2 acquisition patterns similarly to L1 acquisition in that they both exhibit developmental stages, which are “necessary for language development in terms of the mapping of the grammatical principles onto the lexicon” (p. 331). Thus, these researchers propose that even when parameters of UG are immediately available to the L2 learner, the L2 parameters may not be immediately visible to the L2 learner, which is likely the cause of errors. They conclude that mapping between the lexicon and the grammar is the likely cause of the errors.

In her longitudinal study of a Mandarin Chinese speaker (“Patty”) who is a L2 learner of English, Lardiere (1998a, 1998b) speculates that the cause of Patty’s inflectional omissions is due to a mapping problem between the syntax and the morphology in which neither the syntax nor the morphology was in deficit. Lardiere questions prior claims that a presence or absence of inflectional morphology (i.e., Eubank, 1993; Eubank & Grace 1996; Vainikka & Young-Scholten, 1994) can be used as direct evidence for claiming that functional features within the syntactic representation is lacking. As proof of her claim, she posits that because Patty shows clear evidence of 100% accurate case marking and the CP projection in her production, she must have the TP (i.e., *Tense Phrase*) projection despite impoverished surface morphology. Mandarin does not have tense specified within the grammar (cf. Lin, 2005), and so tense would have to be acquired via the L2 English, which *is* specified for tense in the grammar. In fact, Lardiere claims that the T(ense) node in the syntax (cf. Chomsky, 1995) was clearly available to her participant. Thus, she claims that it must be a problem with the unidirectional mapping from the syntax to the

morphology such that the inflectional omissions are indicative not of deficits in the T/AGR areas of the grammar, but of a deficit in an autonomous morphological component which interprets lexical and syntactic derivations. And while remaining somewhat vague in her proposal of the mapping system, she claims that the morphological mis-mapping occurs between the syntax and PF (cf. Chomsky's *Phonological Form*), which suggests that it is still a competence phenomenon. Although Lardiere's claims have merit, it is important to note that Lardiere's mapping claim suffers because of the paucity of subjects in her study (n=1), which weakens her overall claim significantly due to methodological concerns.

In reexamining the same data from Patty, Lardiere (2003) proposes a possible phonological explanation when she alludes to the fact that final *t/d* deletion may be due to a phonological "constraint" against final consonant clusters. However, this claim is not substantiated or defined specifically in relation to any specific L1 phonotactic constraints, and it is not examined in any systematic way in her data. She simply concludes among other possible reasons for inflection deletion that "there is a robust phonological constraint against final consonant clusters" (p. 188), but it is unclear what that constraint might be and how it exactly is affecting deletion of the segments. That said, the possibility that constraints within the phonology (e.g., L1 phonotactic constraints) may cause L2 inflectional variability for some L2 learners has been proposed by others and will be examined more closely below.

Prévost and White (2000) also propose that variability in L2 learner production is not due to a lack of functional categories or features related to tense and agreement within the grammar. Rather, they propose that optional omission of inflections in L2 learner production is due to absence at the surface morphological level rather than at the more abstract feature level. They propose the Missing Surface Inflection Hypothesis (MSIH) which asserts that L2 learners

substitute bare non-finite verbs in place of inflected finite forms and, in fact, that these non-finite forms exhibit properties of finite forms. They claim mainly that if a L2 learner lacks the more abstract functional features within his L2 grammar, non-finite forms should appear in surface finite position, and finite forms should appear in surface non-finite positions. However, if the abstract functional features are present in the grammar, as they claim, there should not be a wholesale interchangeability between finite and non-finite forms. Thus, finite forms should appear in their finite positions (i.e., not in random positions as would be the case if the L2 grammar is lacking the functional categories and features), and non-finite forms that appear in finite positions are merely lacking the inflections on the surface. Prévost and White examined spontaneous production data in four L2 French and L2 German subjects, and their results indicated that in cases in which L2 grammatical impairment would have resulted in clear differences in surface location (i.e., verb raising with negatives), the surface location of the verbs showed that raising occurred and, thus, feature checking occurred. In addition, in cases in which non-finite verbs occurred variably in finite contexts, these non-finite verbs were claimed to have raised and were checked, but they simply lacked the surface inflectional morphology of the finite form. The authors also claim that without feature checking (i.e., in an impaired grammar), finite forms would appear randomly in non-finite contexts, a result that was not borne out in their data. Importantly, they conclude that agreement is in place for their L2 learners as evidenced in the largely accurate agreement when inflected forms are produced. Finally, they conclude, similar to Lardiere's conclusion, that when a L2 learner resorts to a default non-finite form, this lack of surface morphology indicates a mapping problem between the abstract features and surface forms.

## 2.2 Representational Deficits

Rather than attributing errors to mapping, Hawkins and Liszka (2003) instead claim that the cause of the variability in L2 production is attributable to defective or missing syntactic features within the learner's syntactic representation, a position that is termed the Representational Deficit Hypothesis (RDH). Under the RDH, they argue that Mandarin Chinese speakers have problems assigning the syntactic feature [+/-past] to the appropriate verb forms in English because this feature is not selected in Mandarin and, as they claim, "is subject to a critical period" (p. 5). If this claim is true, it is expected that native Mandarin speakers who are advanced learners of English, having begun learning the L2 post critical period, should consistently drop the *-ed* inflection in all modes of production. However, Hawkins and Liszka did not find this result. First, using a test of morphological knowledge (i.e., a grammar test used to determine knowledge of target morphology) to determine that their subjects had advanced knowledge of grammar, they found that their Mandarin speaking subjects were at an advanced level similar to their Japanese and German subjects. They then tested their subjects using an written grammaticality task that required their subjects to produce inflected real and nonce verbs in frame sentences to isolate their knowledge of past regular morphology. In this task, they also found that the subjects had advanced knowledge of English morphology with no significant difference in L2 inflectional production accuracy from the other groups, including native speakers of English. For example, the percentage of correctly inflected English past tense verbs and nonce words that required past tense inflections for each language group were as follows: Chinese, 91.3%; Japanese, 94.0%; German, 96.9%; and English, 99.3%. Finally, in two spontaneous oral production (i.e., narrative retelling and recounting) tasks, they found that their Mandarin speakers significantly omitted the past inflection *-ed*. Hawkins and Liszka observed

that it was only in the spontaneous oral performance tasks where their Mandarin speakers omitted the past inflection in any significant way. On the spontaneous oral production tasks, they observed the following percentage of correct regular past tense inflected verbs for each language group: Chinese, 62.5%; Japanese, 91.9%; and German, 96.3%.

Hawkins and Liszka also considered the possibility that word-final consonant clusters might be causing phonological interference resulting in omission of the past tense *-ed* for the Chinese speakers in the spontaneous oral tasks. To ascertain whether this might be the case, the authors examined the subjects' performance in spontaneous production of monomorphemic words such as *most* and *kind* that end in obstruents [t] and [d]. The authors conclude that because the Chinese speakers had an 82% accuracy rate of final consonant retention in monomorphemic uninflected words compared to the Japanese speakers' (96%) and German speakers' (100%) rates, this reduced rate is only "suggestive" of a phonological effect. In fact, the authors claim that because the Chinese speakers' rate of retention of [t] and [d] in monomorphemic uninflected words (82%) was greater than their accuracy of insertion of past inflections (63%), the reason for the dropping of [t] and [d] in inflected past regular verbs cannot be attributed to phonological interference. The authors also claim that increased performance pressures in spontaneous oral production were not the cause of the Mandarin subjects' errors because these pressures should have been evident across the other subject groups as well—i.e., Hawkins and Liszka did not observe this processing pressure effect for the Japanese and German subjects. Thus, the authors conclude that because the feature for tense [+/-past] is not instantiated in the L1 syntax for Mandarin Chinese speakers, the feature is, therefore, not available to the adult Mandarin speakers causing them to variably insert bare Vs into terminal nodes. However, the question arises which is that if this [+/- past] feature is not available to these Mandarin speaking subjects,

it should show a similarly negative effect across the board for the authors' other controlled written production task, which was designed to determine if their subjects could select the appropriate past morphology on real and nonce words. In fact, they did not see a negative effect—that is, their Mandarin-speaking subjects showed an advanced knowledge of the L2 inflectional morphology. This confound is not addressed by Hawkins and Liszka.

There are also some methodological concerns with Hawkins and Liszka's study as well. For example, the number of subjects for each group was quite low (i.e., Chinese n=2, Japanese n=2, German n=5). Additionally, they claim that because one subject group (i.e., Chinese) showed reduced accuracy on the spontaneous oral tasks, but the other groups (i.e., Japanese and German) did not, performance pressures leading to reduced accuracy were not observable in these data. However, this cross-subject group comparison is problematic, as it would be better to compare processing pressure within each separate group under higher and lower processing conditions (e.g., sentence length or complexity) within each task to determine different memory-load effects due to item type. This dissertation will address these methodological concerns.

Further citing deficits in the grammar as the reason for the inflectional errors, Goad, White, and Steele (2003) attribute the omission of surface inflections not to syntactic deficits, but to prosodic differences between the learner's L1 and the L2 in what they term the Prosodic Transfer Hypothesis (PTH). Specifically, they claim that differences in prosodic representations in L1 vs. L2 competence cause oral production errors in the L2. Goad, White, and Steele claim that Mandarin does not allow adjunction to the prosodic word, and therefore, English inflections that would otherwise be available to a Mandarin L1 speaker in the interlanguage are unreliably produced due to differences in prosodic structures across the languages. Specifically, Goad, White, and Steele propose that English attaches inflections to verbs in a prosodically different

way than Mandarin such that an English verb adjoins the past *-ed* inflection as an adjunct Prosodic Word attached to the main Prosodic Word (e.g., [want]<sub>Pwd</sub> + [ed]<sub>Pwd</sub>). On the other hand, the authors claim that because Mandarin words are mainly monosyllabic in their structure and represent past through an Aspect clitic that attaches to the Foot under the main Prosodic Word (e.g., [want]<sub>Pwd</sub> + [Asp]<sub>Ft</sub>), Mandarin prosodic structure disallows the adjunct Prosodic Word structure required for English regular past tense endings. Thus, Goad, White, and Steele claim that Mandarin speakers omit the past tense inflection because their L1 prosodic structure disallows the English prosodic structure needed for English past inflections. Similar to those of Hawkins and Liszka, the subjects in Goad, White, and Steele scored at ceiling on a grammaticality judgment task, which was designed to test the subjects' knowledge<sup>4</sup> of the inflections. The authors also note that their subjects showed native-like evidence of the availability of the T(ense) node via the subjects' suppliance of INFL in other functional features such as nominative case and copula *be*. However, using spontaneous oral production tasks similar to the narrative retelling tasks of Hawkins and Liszka, Goad, White, and Steele observe that their Mandarin speaking subjects are divided into two groups in terms of their accuracy: one group regularly omits the final regular past tense *-ed* inflection across the board or ATB (i.e., at 10% accuracy) because as the authors' claim, "their grammars do not permit adjunction in outputs" (p. 254), and the other group that variably supplies L2 inflections or variable deletion group through various syllabification options (i.e., at 49% accuracy) "where [inflectional morphology] can be incorporated into the PWD" (p. 256). The latter group of subjects are said to be incorporating the inflectional morphology into various syllabification options. However, the question arises which is that if the PTH restricts adjunction to the prosodic word for Mandarin-

---

<sup>4</sup> We use the term *knowledge* here to refer to a language learner's grammatical competence in general and do not differentiate between any specific types of competence. This point will be addressed in more detail in Chapter 3.

speaking subjects, why would the subjects produce native-like (i.e., at ceiling) knowledge of the target inflectional morphology in the authors' controlled task and then subsequently omit or variably produce the target inflections in a spontaneous oral production task? In an attempt to rule out "performance" effects as the cause of the omission and variability in the spontaneous oral production tasks, Goad, White, and Steele examined accuracy of production of word-final consonant clusters in monomorphemic words by both groups and found that both groups had a similar accuracy rate on the monomorphemic items—57% accuracy of production for the ATB group and 68% for the variable deletion group. From this result, the authors conclude that the rates of suppliance with the inflected items "cannot be attributed to more general deletion of clusters" because the ATB group had nearly equal accuracy as the variable deletion group (p. 259). However, this conclusion does not reflect why both groups did not produce these monomorphemic words at higher rates of accuracy, considering that the PTH only targets inflectional morphemes, not word-final consonant clusters that are lexically incorporated into the word (i.e., not inflected).

Additionally, Goad, White, and Steele found that the suppliance of regular past tense and 3<sup>rd</sup> singular agreement inflectional errors were not equal for many of their subjects even though, as they claim under the PTH, both tense and agreement features must adjoin to the adjunct prosodic word in English and that this adjunction is not possible in Mandarin. Accordingly, if adjunction to the prosodic word is illicit for Mandarin speakers causing English inflectional errors in speech as the authors' claim, this effect should be consistent across the board and should reveal below chance correct suppliance of the past *-ed* inflection for ALL of the subjects in all of the tasks, but this effect is simply not attested across the board in their study.

Both of these representational deficit approaches (i.e., RDH and PTH) claim that there are deficits in the language learner's grammar (i.e., syntactic or prosodic) that inhibit the production of L2 grammatical inflections. In other words, both hypotheses assume that features in some area of the L2 grammar, syntax or prosody, are not available for the L2 learner in the same way that they are to native speakers. However, it should be pointed out that in both cases, their subjects revealed target-like knowledge of the L2 inflections on grammaticality judgment tasks. Both studies revealed that the subjects performed more poorly on the spontaneous oral production tasks. This dissertation proposes that these results constitute evidence for performance effects and that this variability in performance can be systematically explained by controlling performance factors. In fact, as will be explained further, this researcher claims that Mandarin speakers' reduced aural perceptual accuracy and variable oral production of inflections is related to their ability to access the grammar due to increased performance pressures and L1 phonotactic interference effects and is not evidence of representational deficits. Accordingly, because the subjects in Hawkins and Liszka and Goad, White, and Steele showed clear knowledge of L2 grammatical inflections (i.e., as indicated by the subjects' high accuracy rates on the controlled sentence production tasks) but still could not consistently produce the L2 inflections in the spontaneous tasks, the data in those studies are more indicative of performance effects. In fact, this researcher claims that perception of the inflections is critically affected at a level in the grammar that is more of a surface-level phenomenon at the level of PF (cf. Chomsky, 1997)<sup>5</sup>. Thus, as is proposed in this dissertation, variable production of the L2 inflections can be traced to effects such as salience, phonotactics, and processing load effects, which are independent of syntactic or prosodic deficits.

---

<sup>5</sup> Chomsky refers to the PF system that is related to this level of the grammar as the Articulatory-Perceptual (A-P) system. Note that the focus of this study is not to empirically examine this theoretical system, but it is important to understand the generative theoretical framework that underpins this dissertation research.

Contrary to Hawkins and Liszka (2003) and Goad, White, and Steele (2003), this dissertation is in agreement with claims that access to Universal Grammar (UG) is still possible for adults who begin to acquire a L2 after puberty. As such, this researcher agrees with researchers such as Flynn and Martohardjono (1994), who claim that access to UG is still possible for adult L2 learners and that surface variability of L2 inflections is not a question of lack of L2 competence, but critically, it is a question of why variability of L2 surface inflections persists even when L2 morphosyntactic representation is available to L2 learners.

### 2.3 L1 Phonological Interference on L2 Segmental Structure: The Importance of Perception

With regard to the question of UG accessibility, Brown (1998) examines the role of L1 phonological features in the acquisition of L2 segmental structure and asks the following three questions: 1) What accounts for the failure of L2 learners to acquire L2 phonological properties?; 2) What aspects of the L1 grammar impinge on L2 phonological acquisition?; And 3) can L2 learners acquire novel phonological representations when the necessary conditions are met? The author describes the differences between L1 and L2 phonological acquisition and explains that the “L1 phonological system may block accurate perception of the input, thereby preventing the acquisition of novel segmental representations” (p. 140). Brown claims that two conditions must be met for phoneme acquisition to occur: 1) UG must be operating and 2) the learner must detect a phonemic contrast in the input. As such, she claims that a L2 learner’s speech perception capacities must be working in order to perceive novel segmental input or intake of that segment will not occur. Thus, a L2 learner who has a mature L1 phonological system may perceive the L2 input through this L1 system leading to misperception of new input. In fact, Brown claims that if the L2 input continues to be mapped to existing L1 categories within the learner’s grammar, L2

acquisition of novel segments will not occur, and the learner will continue to be unable to distinguish between L2 phonological contrasts in the input.

However, unlike Hawkins and Liszka (2003) and Goad, White, and Steele (2003), Brown claims that novel L2 segments can be acquired if certain conditions are met in the input. Brown's conclusion rests upon the critical first step to acquisition: perception of input leading to intake. To make claims about L2 segmental acquisition, Brown first describes the development of L1 segmental acquisition within a feature geometry model, in which the L1 grammar first builds the feature geometry of segments with less structure and then moves to acquire segments that contain more articulatory complexity. She makes the claim that L1 phonological acquisition is a consequence of first perceiving (i.e., *detecting*) segmental contrasts in the input, which becomes intake once it is analyzed or processed, and then that new input can be successfully acquired. Her claim is that the child must encounter contrastive input to build novel phonemes or the structure that differentiates different phonemes will never be postulated such as in cases in which there are two allophones of the same phoneme, for example. However, citing prior studies, she notes that perceptual sensitivity to non-native contrasts slowly declines with age (cf. Werker & Tees, 1984a, 1984b) and that sensitivity to a non-native segment is retained if that segment does not occur in the L1 (cf. Best, McRoberts, & Sithole, 1988). Brown claims that the phonological structure of the L1 becomes a "filter which funnels acoustically distinct stimuli into a single phonemic category" (p. 147). Thus, from this perspective, this 'filter' can be problematic for L2 learners who encounter a new phonemic contrast within the L2 if one of the contrasting segments is already existent in the L1. Nevertheless, Brown concludes that L2 learners still have access to UG and that novel segments can be acquired if the appropriate input conditions are met, which presupposes that the L2 learner can perceive the novel segments in the input.

In her study, Brown employed an AX minimal pair discrimination task and picture identification task (i.e., both perception tasks) to determine if the subjects (i.e., Japanese and Mandarin speakers) were able to distinguish between an /l/ and /r/ contrast, noting that neither of these languages contrast these segments in their phonology. She claims that we must first show that the subjects can perceive the segmental contrasts because NOT producing segmental contrasts is simply not a reliable test for grammatical knowledge of those contrasts. Brown found that because the L1 feature geometry of Chinese and Japanese are different such that Chinese contains the coronal node (i.e., a required feature to make the /l-r/ contrast in segmental feature geometry) and Japanese does not, Japanese L1 speakers perceived the /l-r/ contrast less accurately than Chinese L1 speakers. Brown claims that Chinese speakers can more successfully map the /l/ and /r/ acoustic signals onto separate segments within their L1 grammars, and therefore, Chinese L1 speakers should more easily acquire the L2 /l-r/ contrast. Although she does not elaborate or test the point, Brown concludes that acoustic differences in the task items, not phonological differences, may have provided clues to differences in the speech signal. Brown also found that the Chinese L1 speakers could distinguish onset consonant clusters, which are disallowed in Chinese. She argues that this result is important because it “suggests a fundamental difference between the acquisition of segmental structure and syllable structure” (p. 172).

Brown’s results support claims by some other researchers (namely, Epstein et al., 1996; Flynn & Martohardjono, 1994; Martohardjono et al., 2012) that access to UG in acquisition of the L2 is still available to the L2 learner for acquisition of novel segments when the right conditions are met. She also posits that data from perceptual comprehension tasks are more reliable to judge underlying competence because there are fewer performance issues that could intervene and obscure the data. The issue of performance factors raised by Brown is one that will

be examined in more detail within this dissertation. Finally, Brown's study introduces the notion that the perceptual variability of L2 segments in word-final position is independent of L1 segmental inventory constraints such that L2 learners may be able to overcome L1 phonotactic constraints even when L1 segmental constraints continue to hinder acquisition of L2 segmental contrasts. Brown's overall conclusions confirm some important theoretical and methodological claims and serve as a starting point for a discussion about causes of inflectional variability that lie outside of representational deficit claims.

#### 2.4 L1 Phonotactic Constraints and Effects on Perception and Production

Following Brown's discussion of phonotactics, another potential explanation for the variable production of inflectional endings is that there may be an effect of mismatches between L1 and L2 phonotactic constraints. What if the L2 learner is unable to correctly perceive or produce L2 inflections due to L1 constraints on syllable structure, for example? For L2 learners who receive non-native English speaker input or English input that may be acoustically *less robust* in the input, such as non-syllabic, unstressed regular past tense [t] and 3sg agreement inflection [s] in English, the result could be considerable variability in perception and production. Additionally, for L2 learners who have mis-matches between their L1 and L2 phonotactic constraints, variability in the output of consonant clusters could result in variation in the production of target inflections. The following literature examines these potential interference factors and their effects on performance.

In an example of the effect of faulty perception leading to performance variability, Dupoux, Kakehi, Hirose, Pallier, and Mehler (1999) reveal that when a speaker's first language phonotactic constraints disallow a CC cluster (e.g., Japanese), the speaker will actually perceive

an “illusory” vowel inside of the cluster. Their Japanese subjects regularly identified a vowel in contexts that did not contain the additional vowel while their French subjects did not. They claim that because phonotactic constraints in Japanese disallow the target consonant clusters, Japanese subjects perceived the input through their L1. This same study revealed that these Japanese speakers could not easily discriminate between VCCV and VCuCV stimuli. The Japanese speakers mapped both of the stimuli to their L1 phonotactic constraints revealing a fault in their perception of the L2 at the syllable level. Dupoux et al. also analyzed L2 proficiency to determine if there was an effect on the Japanese speakers’ perception depending on their L2 proficiency level, but they found that even the Japanese speakers who had advanced knowledge of English and French still revealed the vowel epenthesis effect. Dupoux et al. provide evidence to the fact that what is perceived is directly connected with how the input is analyzed and specifically that L1 phonotactic constraints can significantly affect perception of the L2. Unlike Goad et al. who claim that the L2 prosodic structure is inaccessible to the L2 learner leading to deficits in L2 representation (i.e., specifically in terms of how the grammar prosodifies L2 inflectional morphology) causing errors in production, Dupoux et al. reveal that language processing may depend to a large extent on the L1 phonotactic constraints, and these constraints can have a direct effect on the aural perception of input.

Some related research on cross-linguistic loanword adaptations reveal that perceptual differences between the borrowing language and the language of the loanword can explain changes that are made to the loanwords to “assimilate” them into the language (e.g., Kang, 2003; Peperkamp & Dupoux, 2003; Steriade, 2001). These authors claim that perceptual similarity between the two languages has a significant effect on the loanword phonology. For example, these studies claim that loanword vowel insertions such as the ones observed in Dupoux et al.

(1999) are a direct result of a mis-perception as determined by the L1 phonotactics that leads to the epenthesis. Similarly, Peperkamp (2004) revealed that loanword adaptations are not always in accordance with the borrowing language phonology at the phoneme level. The author proposes that alterations in the loanwords are made at the phonetic and/or perceptual level and that “all loanword adaptations are phonetically minimal transformations that apply during speech perception” (p. 1). For example, she cites examples of English loanwords in Korean in which underlying /s/ should have become [Ct] in coda position according to Korean phonology. However, in English loanwords such as *boss* and *mouse*, the final /s/ that should have undergone the change to [Ct] remained, and [ɪ] was epenthesized into the coda. The resulting forms were [posɪ] for *boss* and [mausɪ] for *mouse* as opposed to \*[pot] or \*[maut], respectively. The author concludes that rather than the borrowing phonology having special rules that only apply to loanwords in representation, which would be unlikely, it is more likely that loanword adaptations are phonetic transformations that occur in perception at the syllable level, rather than at the segmental level. The author makes this observation based on fact that many loanword adaptations are similar to the types of misperceptions that occur due to L1 syllable constraints rather than based on the underlying phonological transformational rules of the L1. Importantly, similar to Dupoux et al., Peperkamp concludes that loanword adaptations occur at the syllable level and that these adaptations are realized during speech perception, overcoming L1 phonological transformational rules.

Kabak and Idsardi (2007) revisited Dupoux et al. (1999) to replicate and extend their study to determine the effects of more specific L1 syllable constraints on Korean speakers’ perception of novel syllable structures that are not instantiated in their L1. Similar to Dupoux et al., Kabak and Idsardi were concerned with L1 phonotactic effects on perception. However, they

extended their study to more specific L1 syllable structure conditions, examining the effects of the following two conditions on Korean speakers' perception of non-Korean nonsense syllables: 1) syllable structure restrictions that prohibit the occurrence of certain consonants in coda position (e.g., \*[c#] and \*[g#]), and 2) consonantal contact restrictions that ban the co-occurrence of certain heterosyllabic consonants (e.g., \*[k.m] and \*[l.n]) in syllable boundaries due to various phonological processes that repair these sequences on the surface (e.g., /k.m/ → [ŋ.m] and /l.n/ → [l.l]). Kabak and Idsardi examined Korean, which has a more diverse array of consonantal sequences, but they note that Korean speakers (like the Japanese speakers in Dupoux et al.) still insert epenthetic vowels into syllables that do not match their L1, such as in loan words (e.g., [k<sup>h</sup>ʉ.ri.sʉ.ma.sʉ] = 'Christmas'). Kabak and Idsardi claim that research by Dupoux et al. provides clear evidence for the role of perception in explaining production errors. However, Kabak and Idsardi raised the question of whether perceptual epenthesis for Japanese speakers stems from the co-occurrence of certain consonants or from the fact that Japanese syllable structure is predominantly CV. Thus, Kabak and Idsardi provide a more fine-grained analysis of the syllable structures and boundaries under study. Consequently, they claim that given a L1 illicit \*[C<sub>1</sub>.C<sub>2</sub>] sequence, vowel epenthesis could occur in the following two phonological contexts:

- Consonantal Contact Hypothesis* – Korean listeners will apply perceptual epenthesis to all consonantal sequences that are illicit in Korean (e.g., \*[C<sub>1</sub>.C<sub>2</sub>] where C<sub>1</sub> and C<sub>2</sub> are licit consonants, but the combination is illicit as in \*[k.m]).
- Coda Condition Hypothesis* – Korean listeners will apply perceptual epenthesis only when there is a syllable structure violation concerning the coda consonant (e.g., \*[C<sub>1</sub>.C<sub>2</sub>] where C<sub>1</sub> is an illicit coda).

They employed an AX discrimination task in which they paired two nonce words, one that contained an illicit sequence of consonants and another that repaired the violation through vowel epenthesis (e.g., [p<sup>h</sup>akma] vs. [p<sup>h</sup>ak<sup>h</sup>uma]) or through L1 phonological rules (e.g., [p<sup>h</sup>akma] vs. [p<sup>h</sup>anma] in which stops become nasals in front of nasals through a Korean nasalization rule). They found that their subjects could significantly distinguish between the pairs that had been repaired by epenthesis (e.g. [p<sup>h</sup>akma] vs. [p<sup>h</sup>ak<sup>h</sup>uma]). That is, their results revealed that their subjects could successfully perceive certain clusters that never occur in Korean, which is not supportive of the first hypothesis (*Consonantal Contact*). Consequently, because several of the clusters under study were perceived at differing rates, they found more evidence for the second hypothesis (*Coda Condition*) such that specific syllable constraints, rather than the more general contact violation, influence the perception of consonant clusters that do not occur in the L1. This result is similar to Brown's (1998) conclusion that syllable constraints are independent of segmental constraints. Nevertheless, Kabak and Idsardi concluded that the epenthesis effect was related to illicit sequences that incur a syllable structure violation and that a more fine-grained analysis of syllable structure violations could explain mis-perception of clusters that do not occur in the L1.

Thus, just as Brown (1998) and Dupoux et al. (1999) concluded, Kabak and Idsardi (2007) also support the claim that when novel phonological input (such as in L2 input) is incorrectly parsed via the L1, the L2 learner can perceive the input incorrectly. Further, it is important to note that the above studies attribute variable perception of the L2 to syllable constraint differences between the L1 and L2 that lead to variability in performance, specifically in the aural perception of novel input. These accounts of aural misperception within the syllable reveal that for L2 learners who may have a limited L1 syllable structure (e.g., Mandarin

speakers), perceiving and producing complex consonant clusters as in the onset and codas of English syllables may present challenges, and these challenges could interfere with the production of the complex clusters required for the production of inflectional morphology. This potential area of L1 syllable constraint interference is discussed below.

Utilizing an Optimality Theoretic (OT) approach, Davidson (2005, 2006a, 2006b) examined the effects of L1 phonotactics on L2 production of non-native consonantal sequences. Davidson looked at the effects of phonotactics and coordination of non-native segmental sequences to claim that accuracy rates of non-native sequences are variable according to a reordering of constraints in the OT interlanguage grammar that is phonetically influenced and, therefore, not fully realized within the deeper levels of the L2 phonology. That said, she hypothesizes that the subjects under study were able to apply their L1 feature information from existing legal sequences in the phonology to novel sequences in other illicit syllable positions in the experimental items. She also claims that epenthetic schwa in non-native sequences (i.e., English speakers epenthesis of schwa between the [z] and [g] of *z<sup>2</sup>gano* due to the illicit phonotactics of the word-initial cluster) may not be realized within the phonology, but it may be inserted at the “articulatory” level due to “gestural mistiming” (i.e., lack of appropriate consonant coordination). Therefore, she claims that this epenthesized schwa is not fully realized as a full vowel segment. Davidson refers to this epenthetic schwa as a “transitional schwa.” Using an OT approach, which claims that all universal constraints are available within the grammar and must simply be ordered in a language-specific way to output any grammar, she claims that non-native speakers must simply reorder the constraints to output the non-native grammar in a target-like manner. Thus, in her study, she posits that the L2 phonotactic sequences of pseudo-Czech consonant clusters (i.e., nonce words) that are illicit to English speakers can be

produced if specific ALIGN (*Consonants*) and DEP constraints are ranked appropriately in the grammar. In non-native grammars, Davidson also posits what she calls “floating constraints” that account for variability in non-native production. She claims that non-native speakers employ varying constraint rankings in their interlanguage grammar that facilitate syllable repairs to maximize the recoverability of the cluster sequence because, as she claims, “when non-native speakers are aware of the intended phonemes of the word they are attempting to produce, they prefer to preserve as much information as possible” (p. 18). Davidson’s approach within the OT framework is that the non-native grammar does not lack the “representation” (i.e., the phonological and phonotactic constraints), but it will attempt to variably reorder constraints to output the native-like sequences according to the native phonotactic constraints. And as Davidson posits, the cluster repairs leading to the schwa epenthesis (or insertion of a “transitional vowel”) occur in variation with native-like cluster production, thus, revealing that the non-native speaker can apply non-native phonotactics to new segmental sequences leading to output that is native-like but still variably produced.

Davidson’s approach reveals that non-native output can be variable due to differences in phonotactic constraints, and as she claims, the repair strategies may be reflective of this variability, but by implication in the OT approach, these repair strategies are not evidence of a lack of representation within the L2 learner’s grammar. Additionally, she claims that variability is proportional to the strength of ranking within the L2 learner’s new grammar and that the interlanguage does vary and can move towards native-like fluency.

In an attempt to determine how phonotactic constraints are encoded in native speaker English, Goldrick (2004) used non-word sequences of CVC syllables to determine if native speakers of English would first exhibit the phonotactic constraints of English in these non-words

and, second, would these constraints be encoded at the *syllable* (i.e., segmental) level or at the *phonemic* (i.e., feature) level. For example, English disallows the phoneme /ŋ/ in syllable-initial position. Goldrick asks whether the processing system encodes this onset condition only at the syllable level (i.e., \*/ŋ/ syllable-initial) or does it also encode a restriction at the feature level (i.e., \*#[dorsal, nasal]). Replicating research by Dell, Reed, Adams, and Meyer (2000), Goldrick first found that when native English speakers were introduced to novel (or nonce) CVC syllables, the subjects' English phonotactic constraints exhibited a significant influence on the CVC sequences and showed an increase of error rates in production of the novel sequences that were illicit in English. However, Goldrick also found that when the syllable-initial and syllable-final segments were biased toward onset or coda position in the target nonce items (i.e., the target segments occurred 25% more times in onset or coda), feature constraints for those targets also showed an effect on the error rates although it was weaker than the segmental effects. Goldrick, therefore, concludes that phonotactic constraints are encoded both at the processing and representational levels. This conclusion is an important distinction when examining the effects of L1 phonotactic constraints on perception and production of a L2 because it first underscores the possibility of transfer of these constraints between the L1 and L2 grammar, and hence, it could be an area of interference leading to variability of perception and production at the syllable level. And second, it reveals that syllable constraints may act independently of phonological features, leading to an analysis at the conceptual or phonetic level. In the case of this dissertation, as will be detailed below, L1 (i.e., Mandarin) phonotactic constraints are proposed to have a significant effect on the L2 (i.e., English) syllable structure, specifically negatively affecting the word-final consonant clusters that include the target inflections under study.

## 2.5 Variable Perception and the Effects on Performance

Some research has examined the effects of perceptibility of L2 input in relation to overall L2 performance. In all of the following studies, the researchers examine the relative perceptibility (i.e., acoustic perceptibility) of specific syllable or segmental sequences and reveal that acoustic perceptibility of input can have an effect on overall L2 performance.

In a study of English-speaking learners of Spanish, Barcroft and VanPatten (1997) examined the effects that location within an utterance, the effects that acoustic stress, and the effects that boundedness (i.e., bound or unbound morphemes) have on the perceptibility of certain grammatical forms. They first define *acoustic salience* as any property that makes linguistic features more or less perceptible to a listener. The authors begin their motivation for the study by making reference to a theory of certain operating principles of L1 language processing (e.g., Peters, 1985), which claims that children learning their L1 have a tendency to segment and store input units as single syllables with particular salient properties such as *fas* for *faucet* or *daw* for *towel*. The authors also state that an additional operating principle is at work, which is that children learning the L1 tend to segment and extract those syllables that are stressed. Barcroft and VanPatten also reference research on foreigner talk (e.g., Hatch, 1983) to show that information that is at the beginning or end of an utterance is more salient to L2 learners than other non-initial or non-final position within utterances. Finally, the authors hypothesized that differences in boundedness between verb endings vs. unbound pronouns in Spanish (i.e., morphologically bound suffixes vs. pronouns) can have an effect on the relative acoustic perceptibility of the bound vs. unbound morphemes. The subjects in Barcroft and VanPatten's study (i.e., English-speaking beginning learners of Spanish) were asked to complete an elicited imitation task in which they heard and then repeated target L2 sentences which

contained target items in different positions and with different stress patterns within the target utterances. The subjects were scored on their accuracy to correctly repeat the target item within the utterance. Of the three conditions (i.e., location within the utterance, stress/unstress, and boundedness), the authors found the following two main results: 1) stressed items were more acoustically salient than unstressed items; 2) utterance-initial items were more acoustically salient than medial or final items. They found no effect of boundedness of the target items for their subjects—that is, there was no significant difference in accuracy of bound or unbound morphemes. The authors conclude that beginning L2 learners are likely more sensitive to utterance-initial position and that they must develop sensitivity to utterance-final (and then utterance-medial) position later in L2 development. However, as the authors claim, their subjects had very low Spanish proficiency, which may have caused the subjects to overly focus on the beginning of the sentences. Nevertheless, the authors claim that location of target items within a stimulus is an important consideration when designing performance tasks and that processing of L2 input can be mediated by surface conditions that may affect acoustic perceptibility of target items such as stress and location within a stimulus, and this relative perceptibility can have an effect on production. This result has implications for this dissertation, which investigates the possible effects of salience on L2 inflectional variability in production.

Klein et al. (2004) and Solt et al. (2004)<sup>6</sup> examined the connection between perceptual salience and written production. Following what Klein et al. termed the Perceptual Salience Hypothesis (PSH), their study examines the degree to which perceptual salience of a past tense English inflection *-ed* in terms of syllabicity of the final allomorph (i.e., non-syllabic [d] vs. syllabic [Vd]) contributes to its variability in performance. Klein et al. conducted a performance task in which subjects listened to target sentences that contained target inflected verbs and had to

---

<sup>6</sup> Klein et al. (2004) is an extended analysis of Solt et al. (2004), which was a poster presentation at BUCLD 28.

write these target L2 verbs with their obligatory inflections into the written cloze sentences. The authors' results revealed that perceptual salience (i.e., acoustic perceptibility) of the past inflection significantly affected the L2 learners' accuracy with regard to the suppliance of the target inflections on the target verbs, revealing that syllabic [Vd] was more accurately perceived and produced than non-syllabic [d].

Additionally, using the same subjects and production data, Solt et al. (2004) also conducted a pure perception task, which required subjects to listen to identical sentences that sometimes contained an inflected target verb (e.g., *walked*) and sometimes contained an uninflected target verb (e.g., *walk*). This aural perception task required subjects to classify each sentence pair as the 'same' or 'different.' The results of this perception task revealed the same result as the written production task, which is that the syllabic allomorph (i.e., [Vd]) was perceived more accurately than the non-syllabic allomorphs (i.e., [t] and [d]). Importantly, when examining the effects of the 'same' vs. 'different' items, the authors found that regardless of proficiency level, the subjects were significantly less able to perceive the missing inflection in the different items when the allomorph was non-syllabic [t] or [d]. Reflecting on the perception and production task findings together, Solt et al. propose that the subjects' "inability to perceive the regular past tense across its allomorphic variants...is a barrier to producing [the past tense – *ed*] morpheme in a target-like manner." The suggestion that there may be a direct connection between perception and production regarding L2 inflections is one of the main questions that is investigated by this dissertation.

In a study of the perception and production of American English (AE) vowels by native Mandarin speakers, Jia, Strange, Wu, Collado, and Guan (2006) found that oral productive accuracy could be linked to aural perceptual accuracy specifically with respect to vowel

contrasts. Jia et al. examined the perception and production of English vowels by Mandarin speakers at different ages and lengths of English native speaker contact. Three groups were examined: one group in China, who had little or no English native-speaker contact, and two groups in the US, one with a greater length of time living in the US than the other. The authors first measured their subjects' abilities to perceive English vowel contrasts in an AXB discrimination task in which nonce word pairs were compared to a third nonsense word that contained the target vowel. Participants had to decide which of the A or B stimuli sounded like the middle X target item. In a production task, subjects then imitated the same target nonce words in an oral elicitation task that required them to listen to and then repeat the target nonce words. Because the focus of their research was on vowel contrasts, none of the words violated English or Mandarin phonotactics such that they were all open syllables of CV. On the perception task, all subjects (even those who were in China and had very minimal exposure to native English speakers) scored above chance at a level of 70% accuracy in all contrasts. However, as might be expected, the subjects in China had significantly lower accuracy than the subjects who were in the US and who had greater exposure to the L2. Similarly on the production task, subjects in China had significantly lower accuracy rates than the two groups in the US. When Jia et al. compared the perception and production accuracy of the entire group, they found that there were significant positive correlations between perception and production especially for the group in China. They also found that subjects with younger Ages of Arrival to the US (AoA) revealed a significant effect due to AoA particularly for the younger subjects who had been in the US the longest compared to the new arrivals. Notably, because they found a positive correlation between perception and production at the individual and group levels, with more accurate perception than production for some vowel pairs, they claim that these results

indicate that “production lagged behind perception,” and at least at the individual level, “better perception performance significantly predicted better production performance” (p. 1128). The claim that perception precedes production in L2 acquisition is not new, as others in the literature have made this claim (e.g., in FLA, Wode, 1995, 1997; in SLA, Flege, 1991, 1995). However, Jia et al. note that a positive correlation between perception and production indicating a causal role between perception and production is controversial. Nevertheless, their findings that perception of certain L2 segments can be strongly correlated to those same segments in production is a significant finding that may suggest a relationship between the perception and production of other elements within the grammar such as L2 inflections. This dissertation examines this relationship with regard to L2 inflectional variability.

From the above studies, there appears to be a connection between perception and production in terms of perception and production of L2 segments (e.g., Jia et al., 2006) and in terms of salience (i.e., acoustic perceptibility) of L2 inflections (e.g., Klein et al., 2004) and the effects of perceptual salience on performance (e.g., Barcroft & VanPatten, 1997; Solt et al., 2004). Two main questions that are in need further examination, as suggested in Solt et al., are as follows: 1) Is variable production of L2 inflections L1-related or more universal? 2) What specifically is the relationship between the perception and production of L2 inflections? The above research strongly suggests that there are likely additional causes outside of deficits in morphosyntactic representation that can lead to variable performance of L2 inflections. Following the above research on performance factors, this dissertation attempts to shed light on this relationship.

## 2.6 General Performance-Based Accounts and the Effects on Performance

Finally, several researchers have proposed that L2 inflectional variability can be directly linked to deficiencies in performance and that this variability may be completely unrelated or only indirectly attributable to grammatical access. As previously suggested by Epstein, Flynn, and Martohardjono (1996) and Klein and Martohardjono (1999), general processing mechanisms could be responsible for non-target production of the L2. The following research provides accounts of performance effects that can be linked to factors such as processing load, working memory, and other factors that lead to a reduction in processing resources, which as this dissertation will investigate is a viable cause of L2 inflectional variability.

Building on the theory of phonological short-term memory (pSTM) within the theory of verbal working memory introduced by Baddeley (1986), Jacquemot, Dupoux, Decouche, and Bachoud-Lévi (2006) analyze data from two case studies of aphasic patients who have working-memory impairment due to reduced storage in the phonological buffer, which is theorized to temporarily store phonological input in order to check and resolve ambiguities as well as reexamine phonological input as needed. In their study, Jacquemot et al. found that their subjects showed excellent processing of single-word tasks, but the subjects showed significant impairment in their processing of target words within sentences. Although they examined L1 speakers with cognitive impairments and the authors' research questions are not directly related to the main questions under study within this dissertation, the findings of Jacquemot et al. are important to understand the nature of verbal working memory and how it may affect word processing vs. sentence processing prior to lexico-semantic or syntactic access. Extending this research on pSTM in neuroimaging studies, Jacquemot and Scott (2006) find that pSTM performance is influenced by the length of stimuli such that sequences of short words are

recalled better than sequences of long words. The authors posit that there is a direct connection between speech production and perception at the phonological level and that this connection must be maintained in processing for successful parsing to occur. The implication of this pSTM research to this dissertation is that working memory mediates processing of language at the phonological level (cf. Baddeley, 1986) and, thus, has the potential to be affected by processing overload irrespective of grammatical representation. The following research examines processing overload effects due to memory-load interference.

In a study of sentence processing in terms of memory load, Gordon, Hendrick, and Levine (2002) attempt to determine the nature of capacity limits on working memory resources used for language comprehension, as well as to determine if syntactic processing uses the same memory resources of more general cognitive processes. In a memory-load task, the subjects had to remember a list of “load” words while they were simultaneously engaged in a sentence comprehension task of subject-extracted or object-extracted cleft sentences. The authors hypothesized that if the load words and the extracted NPs were matched in terms of names or descriptions (e.g., John/John or doctor/doctor), the match should impair sentence comprehension, and this effect should be worse on the object-extracted sentences, which are claimed to be more challenging to interpret (cf. Gibson, 1998). Following similar research on memory-load effects in sentence comprehension (e.g., Caplan & Waters, 1999), Gordon, Hendrick, and Levine found that 1) sentences with greater syntactic complexity are more susceptible to memory-load effects, and 2) because processing load and processing of the syntactic structure utilize the same memory resources, the authors posit that memory interference can have a large effect on sentence comprehension. Again, as with the Jacquemot et al. and Jacquemot and Scott studies, this dissertation is not investigating the effects of online processing load effects of this type.

However, the results of these studies on memory-load interference suggest that processing of language can be affected by factors outside of grammatical representation and processing overload effects must be considered in research on variability of language performance.

In a recent study that examines the variability of functional morphology (i.e., article production), Trenkic (2007) explores the debate over what causes this variability in production. Referencing the various accounts (i.e., representational and mapping), the author posits that only when asymmetries in production due to different L1 influences can be predicted will morphological variability be better understood. Of the different major proposals of inflectional variability (i.e., RDH, PTH, FAH, MSIH), Trenkic proposes that only the PTH of Goad, White, and Steele has predictive power in terms of how languages differ in the way they prosodify functional material, but she claims that this account has problems when applied to English articles. Trenkic applies the PTH to Serbian speakers learning English to show that the PTH fails to account for all asymmetries in L2 production of functional morphology. Focusing on the comparison of L2 learners' English article production in Art+Adj+N vs. Art+N sequences, Trenkic begins with an accounting of adjunction of Serbian and English articles and finds that Serbian does have the appropriate representation to prosodify English articles without adjunction to the PWD—she notes that the PTH should not apply in this case because the L1 has the appropriate prosodic representation. Trenkic also notes that Serbian does not have definite articles, and therefore, it does not realize the feature [Def] in the grammar and cannot transfer this feature into the L2, English.

Using two tasks, one a spoken communicative and the other a written translation task, Trenkic found that in the spoken communication task, the subjects had nearly double the amount of omission of English articles in Art+Adj+N contexts compared to Art+N contexts. The author

claims that the PTH is unable to account for this significant difference in article omission in the contexts with an intervening Adj. In the written translation task, she again found that contexts with the intervening Adj showed significantly higher English article omission rates. In addition, she found that the subjects with higher L2 proficiency levels had significantly reduced omission rates compared to the lower proficiency level subjects. The author concludes that the PTH cannot account for these results because for one, the PTH makes predictions of spoken production only, and second, the L1 contains the requisite prosodic structure to produce the appropriate articles in both of the obligatory contexts, yet only the intervening Adj contexts produce significant omission of the articles. Thus, the PTH fails to predict this asymmetry in Serbian learners of English. Trenkic goes on to explain that the RDH also cannot explain the asymmetry in production across the different proficiency levels because the higher proficiency level subjects in her study produced articles are near-native levels—as she posits, the RDH should predict that article production would be nearly non-existent for both groups because of the lack of the [Def] feature in the L1. The author also dismisses the MSIH as *post hoc* and not predicative of the asymmetry in her own data.

Trenkic goes on to explain the L2 article production asymmetry of her subjects in terms of how language is conceptualized and produced from pre-verbal message to phonetic output. She explains that because the feature [Def] is not specified within the L1 syntax, the production of articles would have to happen at the conceptual and lexical level such that a speaker would have to first encode articles as *Adj-like* forms in the lexicon so that the syntax would identify the article as an Adj-like object. In addition, Trenkic claims that without the article specification in a subject's grammar, the L2 speaker who is aware that the L2 requires articles would likely monitor the usage of the articles in the L2 at the pragmatic level, which would require additional

processing resources or “attentional” resources as the author calls them, especially at the lower proficiency levels until the process becomes automatic at the higher proficiency levels. In cases in which there is an intervening Adj (i.e., Art+Adj+N), Trenkic claims that the computational requirements would be higher due to the more complex structure, increasing or even exceeding the processing or attentional resources, which in the case of her subjects led to more article omissions even when her subjects had higher proficiency levels. Because of the asymmetries that Trenkic observed even within production of the same functional morpheme in different surface contexts, the author concludes that variable production of L2 functional material should not be used to make absolute judgments about grammatical knowledge and that analyses of L2 learners should focus on the asymmetries in production to better understand L2 learners’ acquisition.

In a summary analysis of four experiments that measure L2 learners’ problems with L2 inflectional morphology, Hopp (2009) finds a common outcome, which is that persistent problems with L2 variability in surface inflections can be systematically attributed to L1 transfer and limitations in L2 processing efficiency. Additionally, he claims that the findings are not consistent with arguments that claim the existence of a critical period for representation in L2 acquisition, as in the case of the RDH and PTH. Hopp posits that researchers must be careful of assigning a general critical period for all inflectional variability, and he claims that differences in inflectional variability depending on the L1 and on the proficiency of the L2 “suggest that not all late L2 learners may suffer the same problem with inflection” (p. 902). Importantly, Hopp also claims that differences in task demands and inflection type result in differences in inflectional variability such that the variability more likely represents processing load effects as claimed by others (e.g., Trenkic, 2007) even at advanced L2 proficiency levels. In his analysis, expanding on the conclusion put forth by Prévost and White (2000), Hopp first reveals that although L2

learners often omit inflections, when L2 learners supply L2 inflections, they do so systematically and in a target-like fashion; thus, omission alone is not evidence of a representational deficit. In terms of L1 effects on the L2, Hopp suggests that tasks that attempt to separate evidence of representation from processing of inflections should be the focus of current research on L2 inflectional variability.

Hopp examines results from four experiments that combine off-line and on-line data from the same subjects with different L1 backgrounds who were advanced to native-like L2 learners of German. In the first experiment, off-line tasks designed to determine the subjects' knowledge of the German case inflection and word order variation revealed that the advanced (or as the author terms "lower-level") subjects make grammatical judgments that reflect L1 influences leading to lower accuracy, but he also showed that these L1 influences were cancelled out as proficiency level increased to the "near-native" level. Thus, he concludes that L2 inflection and word order are attainable for adult near-native L2 learners of various L1 backgrounds even when the L1 does not instantiate the L2 morphosyntactic variations. In the second experiment, Hopp examined whether the subjects could apply their off-line advanced knowledge of L2 morphosyntax to on-line processing in real-time comprehension. The author conducted a self-paced reading task that tested for the effects of case-marking and subject-verb agreement on sentence processing. In this task, like others of its type (e.g., moving window), a slowdown in reading (i.e., increased reading time) on a given utterance segment compared to the same segment in other conditions indicates increased processing effort. Hopp found that the near-native subjects responded like the native speakers and were able to utilize case-marking and subject-verb agreement to disambiguate the syntactic ambiguities in the target segments of the utterances. The lower-level "advanced" subjects did not show this same measured slow-down,

which as Hopp claims is reflective of proficiency-level effects on processing. He also notes that although the near-native subjects were able to disambiguate the target utterances like natives, the L2 learners' processing times were still considerably longer than the natives. Thus, the third experiment was designed to investigate if target use of inflection depends on speed by imposing the same task demands on the natives and L2 learners. Again, the author uses case and subject-verb agreement for targeted grammaticality judgments, but this task is now administered under time pressure to determine how processing of inflection for L2 learners responds under increased processing load caused by increased processing speed. Hopp found that native performance levels on case and subject-verb agreement is possible for near-native L2 learners. However, he also found that depending on the L1 (i.e., English and Dutch), the effects of speed revealed considerable reduction of processing of case assignment in the L2, German. Hopp surmises that this outcome could either be an effect of a representational deficit due to specific L1 interference effects, or on the other hand, he claims that if increased processing load can be shown to affect native speakers' performance reducing accuracy in the same way that processing load decreases accuracy of case assignment for L2 learners, this result would confirm processing load effects as a cause of performance deficiencies. Hence, experiment four was designed to test parallels between native and near-native L2 speakers' processing under increased task demands. Using the native speakers of German (i.e., the control subjects) from the other three experiments and the same speeded reading materials (i.e., "moving window") from the third experiment, Hopp designed speeded grammaticality judgment tasks that tested the natives' processing at five different presentation speeds, each decreased incrementally from the original 250ms per word down to 71ms per word. He then compared the accuracy judgments of the native speakers to the accuracy of the near-native speakers in experiment three. Interestingly, Hopp found the same

pattern of accuracy on subject-verb agreement as well as the same pattern of reduced accuracy on case assignment for the native speakers at the highest speed when compared to the near-native speakers. The results are confirmed to be significant and reveal that at higher task demands, the native speakers showed parallel processing deficiencies compared to the near-native L2 learners on the same task items. Overall, Hopp makes two important contributions to the argument of representational deficit vs. performance deficit effects on inflectional variability. One is that even though L1 effects can be minimized as L2 learners increase in L2 language proficiency, the L1 effects continue to effect accuracy of the L2 under greater processing load even at advanced levels of L2 proficiency; and two is that reduced accuracy of L2 inflections can be directly attributable to performance deficiencies (i.e., under increased processing load) and that these results are incompatible with a general critical period affecting L2 morphosyntactic representation.

Martohardjono, Valian, and Klein (2012) provide further evidence in favor of performance factors over morphosyntactic representational deficits as the cause of L2 inflectional errors. Their study examines whether beginning Chinese L2 learners of English have tense instantiated in their L2 grammars, and if so, what performance-related factors might be the cause of inflectional errors. The authors begin with the assumption that tense should be accessible in the L2 regardless of the L1<sup>7</sup> in both comprehension and production, which they base on the notion that adult L2 learners have full access to UG. The authors argue that L2 learners, depending on the task they are asked to perform, differentially access L2 grammatical

---

<sup>7</sup> This accessibility claim, known as the *Full-Access to UG* hypothesis, is not the focus of this dissertation. Nevertheless, access to UG for late L2 learners is assumed (cf. Flynn & Martohardjono, 1994) and will be discussed in more detail in Chapter 3.

knowledge<sup>8</sup> and that spontaneous production data is not an adequate source alone to explain errors in inflection. Their study tested Chinese L2 learner's comprehension and production of *tensed* items such as past regular (*-ed*), 3sg agreement (*-s*), copula, aux *be*, and modal auxiliaries (*will* and *did*). Martohardjono et al. used slightly modified instruments from Valian (2006) that were used to test knowledge of tense in 2, 3, and 4-year old L1 learners of English so that the authors could make comparisons in the results of the young L1 learners from Valian's study and their L2 adult subjects. The comprehension task was designed to reveal access to the tense feature within the learners' grammar by asking the subjects to distinguish between present and past tense when presented with scenarios showing different actions and then asking the subjects to choose if the action happened in the past, present, or future. In this first comprehension task, Martohardjono et al. found that the type of tense carrier (i.e., copula, progressive, auxiliary) had an effect on the accuracy of their comprehension. They also found frequency effects in which the subjects were biased towards present copula *is* more often than *was*, which the authors attribute to the fact that present copula *is* is more frequent in English usage. When compared to the child L1 learners of English in Valian's study, the authors found the same frequency effect.

Martohardjono et al. also conducted an oral elicited imitation task that was designed to determine if the L2 learners could produce the same tensed morphemes that were measured in the comprehension task. The authors used an oral elicited imitation task to mitigate performance factors such as message creation and discourse effects that would arise in a spontaneous production task, while still testing processing effects at the sentence level. They first propose that learners (i.e., L1 or L2) should not be able to repeat structures that are not instantiated in their grammars. The authors also admit that this task cannot reveal whether the learners assign formal

---

<sup>8</sup> This dissertation supports this claim. As explained below, this researcher also claims that access to L2 grammatical knowledge is affected by processing and by the proficiency level of the L2 learner as claimed by others (e.g., Hopp, 2009).

features to the target items, but it can still reveal the oral production rates of L2 morphology to make comparisons to native production rates and can reveal whether morphological variability patterns after L1 phonological effects or processing effects due to item type. The authors also included syllabic vs. non-syllabic past and present allomorphs (i.e., syllabic [Vd] and [Vz] vs. non-syllabic [t] and [s]) to make comparisons of salience effects such as those found in Klein et al. (2004) and Solt et al. (2004). The authors found no significant effect of salience between the syllabic and non-syllabic forms, but overall, the bound morphemes were the most challenging for the L2 learners. The authors surmise that this effect could be related to the lower saliency of bound morphemes in the input compared to the unbound tensed elements. Like the comprehension task, the authors again found a frequency effect for the present tense copula *is*, which was produced in favor of the less frequent past copula *was*. In addition, the authors found that the L2 learners were significantly more successful at producing the frame sentences accurately overall when the tensed target items under study were in the present tense rather than in the past tense. Martohardjono et al. attribute this effect to processing, as they and others have claimed that the present tense provides less cost in terms of processing load even for L1 learners (cf. Valian, 2006). Contrary to Hawkins and Lizska (2003), they conclude that their Chinese L1 subjects did show strong evidence of tense in their L2 grammars even though tense is not instantiated in the L1. Thus, they concluded the following: 1) the representations of the target forms exist in the L2 grammar of their Chinese speakers and L1 morphosyntactic transfer effects are not supported; 2) the subjects' variable production was related to frequency and to processing overload caused by particular tense markers that affected the accurate production of the sentences as a whole; and 3) the pattern of errors is comparable to the patterns found in child L1 learners (cf. Valian, 2006). Consequently, Martohardjono et al. conclude that even at a beginning

level of L2 proficiency, L2 learners have access to L2 features that are not instantiated in their L1 and that errors in production of L2 morphology do not signify a lack of knowledge of tense and agreement in the grammar. For Martohardjono et al., processing factors involved in accessing the grammar (not a lack of representations in the grammar) are responsible for their subjects' errors.

## 2.7 Summary of the Literature

Much of the current research that investigates the causes of L2 learner inflectional errors outside of the representational deficit accounts suggests that performance issues must be considered as an important factor in our understanding of the acquisition of L2 inflections. Accordingly, attempting to describe a L2 learner's morphosyntactic representation through spontaneous oral performance tasks alone (e.g., Hawkins & Liszka, 2003; Goad, White, & Steele, 2003; Lardiere, 1998a, 1998b, 2003) is clearly insufficient. When a L2 learner exhibits morphosyntactic competence or knowledge of L2 grammatical representations at a near-native level, accounts that continue to claim a deficit in the syntax due to a critical period become problematic because of the obvious inherent contradiction within this type of argument. As several researchers have proposed (e.g., Davidson, 2005, 2006a, 2006b; Hopp, 2009; Martohardjono et al., 2012; Trenkic, 2007), L2 surface inflectional variability is not, in itself, evidence of lack of L2 representation, especially when examining spontaneous production data. If a representational mismatch is claimed to exist between the L1 and L2 at the level of the syntax or phonology, and this mismatch is claimed to create a *permanent* deficit within the L2 grammar, it should be expected that the mismatch would not affect the L2 inflectional perception and production in random, unsystematic patterns, but it should lead to high rates of errors of

particular surface inflections that are not instantiated in the L1 and are, therefore, unattainable—that is, we should see an across-the-board systematic effect on L2 inflections.

As this dissertation will illustrate, the factors that cause variability in L2 learner inflectional errors are less likely a cause of permanent representational deficits as hypothesized by the RDH and PTH, but they are more likely caused by performance factors that, in some cases, may be independent of the grammar. Additionally, the arguments for mapping deficits (e.g., MSIH and FAH) can provide some insight into the variability of surface inflections, but as Trenkic (2007) has argued, these approaches are decidedly *post hoc* because they cannot *predict* L2 morphological variability in a systematic way, but they can only suggest that a lack of surface inflections will occur. Trenkic's claim is presumably because these mapping approaches can only predict omission will occur generally, but they cannot predict on which inflections the omissions will occur. She suggests that attempts to explain L2 inflectional variability should be able to predict more systematically where the inflectional variability will occur. Additionally, all of the representational deficit and mapping deficit approaches rely exclusively on spontaneous production, as proof for representational deficits, which as several researchers have proposed, is arguably not reliable. As Hopp (2009) has clearly shown, what is needed are approaches that can systematically predict the surface asymmetries of the same L2 inflections under different performance conditions (i.e., lesser and greater processing load) across different modes of performance (i.e., perception and production) and across different proficiency levels who may be approaching ultimate attainment of the L2 (i.e., advanced to near-native speakers). In light of this recent research on language performance, this dissertation systematically examines the variability of L2 surface inflections by exploring the connection between perception and production across different modes, under different processing conditions, and at

different levels of advanced L2 proficiency in an attempt to reconcile these different performance accounts. The results of this study will also provide further supporting evidence that L2 inflectional errors are not caused by deficits in L2 morphosyntactic or prosodic competence in cases in which the L2 learner can demonstrate native-like knowledge of L2 grammatical inflections. Chapter 3 summarizes the theoretical rationale that underpins this dissertation and provides a summary of the various reasons for variability of L2 surface inflections to be investigated by this researcher.

### Chapter 3: Theoretical Background and Research Questions

This dissertation examines the possible causes that lead to the breakdown in performance resulting in variability of L2 surface inflections. However, in order to discuss performance, the model of competence and performance that underlies this research and how it relates to the overall study in this dissertation must be clarified. The following sections briefly examine the nature of both competence and performance in order to position this study within a generativist framework. Building on Chomsky's (1965) original proposal of the division between *Competence* and *Performance*, this chapter provides a rationale for delineating underlying competence from performance and examines the performance factors that can have a direct effect on surface inflectional variability. That said, this dissertation does not attempt to examine Chomsky's definitions in detail or provide evidence of Universal Grammar (UG), as these definitions and theories are not under investigation within this study.

#### 3.1 A Model of Competence and Performance

In Figure 1 below, the model<sup>9</sup> assumes that competence and the corresponding grammars (e.g., L1-First language, IL-Interlanguage, L2-Second language, etc.) underlie an L2 learner's perception and production of grammatical elements (e.g., grammatical morphemes and phonemes) such that accurate perception and production of L2 inflections relies to a great extent on availability of L2 grammatical representation.

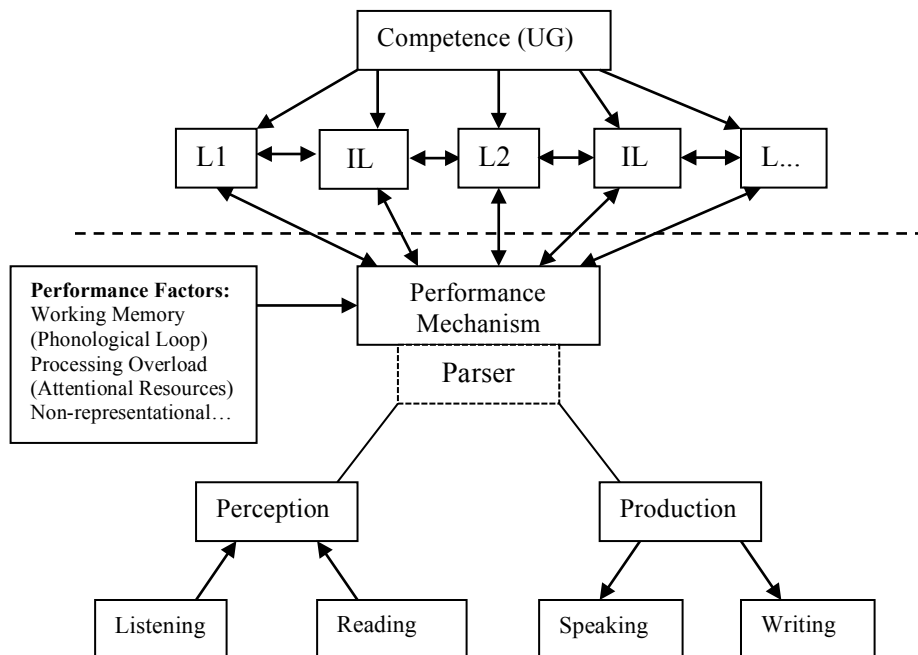
Thus, although UG is assumed to govern the process of grammar construction within the model, once the grammars are developed such that the L2 learner has grammatical representation

---

<sup>9</sup> This theoretical model of competence and performance was developed by this researcher and is modeled after a similar one developed by Klein and Martohardjono (1999).

available for parsing input and output, the parser (or processor) as part of the processing mechanism can then access the grammar to perceive or produce the grammatical surface forms effectively<sup>10</sup>. The following two sections describe the model below and explain where the possible breakdowns can occur in the process of acquiring and producing L2 inflections.

**Figure 1: Competence/Performance Model**



### 3.1.1 Parsing L2 Inflection: A Brief Summary

This dissertation is not concerned with testing the nature of the parser, per se. Instead, it attempts to determine the factors that intercede between the representation and successful perception and production of L2 inflections. Nevertheless, to reduce confusion about how parsing is reflected in the process by which an inflection is perceived and produced, this researcher briefly returns to the model (Figure 1). Again, this dissertation does not make claims

<sup>10</sup> The nature of the language performance mechanism and its component parsing mechanism is not under study in this dissertation. As such, the cognitive nature of the parser is beyond the scope of this dissertation and is only discussed to distinguish between underlying representation and the factors of performance that are not part of the competence.

regarding the nature of specific competence or performance mechanisms—it merely assumes that performance is separate from competence following Chomsky’s original proposal and represents this division as such in the theoretical model. However, this dissertation does address the factors that affect performance and, thus, offers an explanation of parsing between competence and performance as it is assumed in this model.

A successful parse would be one in which the parser within the performance mechanism would successfully access the underlying grammar in order to match either the input via perception to the underlying grammar or the output via production to a target surface form. For example, in the case of production of an English regular past tense verb such as *stopped*, the item is retrieved from the lexicon in two parts (e.g., *stop + ed*) as it has been described by some in the literature (e.g., Levelt, 1989; Levelt, Roelofs, & Meyer, 1999; see also Trenkic, 2007). Similar models of *word + morpheme* retrieval have recently been supported in brain imaging research in which word meaning and morphology are processed differently (e.g., Bozic et al, 2007; Sahin et al., 2009). Although the brain imaging approaches are beyond the scope of this dissertation, these studies claim that a real verb such as *stop* is first retrieved as a *lemma* from the lexicon, and then the morpheme for regular past tense *-ed* is retrieved as a separate functional morpheme which is then combined within the phonology to produce the surface past tense form. Accordingly, in the case of perception of a past regular verb (e.g., *stopped*), the L2 learner must not only be able to perceive the stem of the verb and map its meaning to the lexicon, but also the L2 learner must perceive and map the bound past tense morpheme separately to the grammatical feature [+past] within the syntax. Thus, because the content word (i.e., verb) likely takes precedence over the inflectional morphology within the parser (see Trenkic, 2007 for a discussion of L2 learners’ monitoring pragmatic/conceptual content vs. inflection), the inflection is likely more susceptible

to mis-perception or mis-production even when the grammatical representation has been shown to be available to the L2 learner. The important consideration within this model of processing (i.e., parsing) is that several factors could directly intercede at the performance stage to reduce processing accuracy of the inflections as is represented in Figure 1 (e.g., Performance Factors).

Thus, even if a L2 learner has successfully acquired the grammatical representation needed to perceive and produce L2 inflections, L2 learners can still exhibit variability of inflections (as various studies examining factors outside of representational deficits have shown). As such, it is critical to first determine if the L2 shows evidence of the appropriate morphosyntactic representation to produce L2 target inflections. If availability of L2 morphosyntax is adduced, we can then claim that other factors such as interference from L1 phonotactic constraints at the syllable level (e.g., Brown, 1998; Davidson, 2005, 2006a, 2006b; Dupoux et al., 1999; Goldrick, 2004; Kabak & Idsardi, 2007; Peperkamp, 2004), acoustic perceptibility (e.g., Barcroft & VanPatten, 1997; Jia et al., 2006; Klein et al., 2004; Solt et al., 2004), performance factors related to working memory (e.g., Gordon et al., 2002; Jacquemot et al., 2006; Jacquemot & Scott, 2006), or processing load that reduces processing resources leading to increased error rates (e.g., Epstein et al., 1996; Hopp, 2009; Klein & Martohardjono, 1999; Martohardjono et al., 2012; Trenkic, 2007) interfere with the successful perception and production of the target inflections. Consequently, assuming that L2 morphosyntactic representations *are* intact within the L2 learner's grammar, we are left with the question of *how these phonotactic, salience, and performance factors can reduce access to grammatical representation and subsequently lead to mis-perception and mis-production of L2 grammatical inflections.*

### 3.1.2 Determining Grammatical Competence or Knowledge

Before examining the possible non-morphosyntactic effects on L2 inflectional variability, it is critical to determine that a learner's L2 morphosyntactic knowledge is intact. However, to determine if the L2 learner has grammatical knowledge of the target inflections, we must first consider what constitutes lack of grammatical knowledge. For example, if acquisition of the requisite grammatical representation has not occurred due either to a) insufficient L2 input or b) a critical period that disallows access to UG and L1 transfer of target L2 features from the L1, the L2 inflection (e.g., for past regular *-ed* and 3<sup>rd</sup> singular *-s*) would, thus, be unavailable<sup>11</sup> in the L2 learner's grammar. Consequently, in either case, any attempt to produce the L2 would lead to production of L2 inflections that is highly variable or completely omitted across *all* modes of performance including written grammaticality judgment tasks and across all inflectional surface variations (e.g., allomorphs [t], [d], [Vd] and [s], [z], and [Vz]). However, when L2 grammatical representations are attested (i.e., the L2 learners reveal target-like *knowledge* of L2 grammatical representations via a grammaticality judgment task), acquisition of the target inflection can be assumed.

It is important to note that within this definition of competence, the term *grammatical knowledge* has been a point of debate. Some researchers have suggested that the L1 knowledge that a child acquires is somehow different from the L2 knowledge that is acquired by an adult. Specifically, some recent research (cf. Ellis, 2005, 2008; Ullman, 2004) has attempted to deconstruct Chomsky's notion of competence into separate grammars that are accessed under different conditions. Ellis, for example, has claimed that there is an *implicit* grammar of which the learner is said to have unconscious knowledge and an *explicit* grammar of which the learner

---

<sup>11</sup> This term "unavailable" here does not mean un-acquirable as is claimed by some researchers. Within a *Full Access to UG* approach, this researcher simply means that it has yet to be acquired by the L2 learner.

is said to have conscious knowledge such that he can verbalize the *explicit* grammar in a conscious manner. In Ellis' recent research (and in some of his followers' work, notably DeKeyser, 2003; Paradis, 2004, among others), the attempt has been to determine to what extent tests of grammatical knowledge are accessing the *implicit* (i.e., intuitive, automatic) grammar versus the *explicit* (i.e., conscious, controlled) grammar. In his attempt to reveal these two different types of grammar and the differences in access to them, Ellis claims that un-timed grammaticality judgment tasks, for example, allow L2 learners to access their explicit (i.e., "conscious") knowledge of grammar while other types of timed production tasks such as unplanned oral communicative tasks provide a measure of L2 learners' implicit (i.e., "unconscious") knowledge. However, Ellis admits, "Even if task conditions that inclined learners to use one type of knowledge in preference to the other could be identified, it would be impossible to construct tasks that would provide pure measures of the two types of knowledge" (2005, p. 153).<sup>12</sup> Ellis concludes that learners will likely draw on whatever resources they have available.

As Ellis points out, it may be unlikely that empirical testing can show if L2 learners have different *types* of internal mental grammars from an analysis of language production (cf. Martohardjono, 1998). Irrespective of this point, the primary concern in this dissertation is with the question of whether the L2 learner *is* or *is not* accessing the grammar as a whole. We leave aside Ellis' distinction between internal and external grammars while adopting the notion that grammaticality judgments reflect grammatical knowledge (*implicit* or *explicit*). As such, this dissertation is not concerned with testing the nature of *Competence* or *knowledge* and how

---

<sup>12</sup> In fact, some authors (cf. Clahsen & Felser, 2006; Hahne, Muller, & Clahsen, 2006) provide a counter-explanation that accounts for the L2 learner's apparent lack of access to the implicit grammatical knowledge. In contrast, they claim that differences in grammatical processing are related to the depth of syntactic processing of sentential representations such that it is more a question of processing complex sentence strings than it is different types of access to different parts of the grammar.

competence/knowledge is ordered within the mind/brain. Accordingly, it is presupposed that there is only one type of internal mental grammar that informs the performance of a L2 learner.

### 3.2 Results of a Pilot Study of Competence and Performance

Before detailing the mechanisms that affect performance in the model, this researcher reports the results of a pilot study that further motivate a performance-based analysis. An untimed written Grammaticality Judgment Task (GJT) was used to indicate baseline knowledge of L2 inflectional morphemes, and an Aural Sentence Perception Task (ASPT) and a timed semi-spontaneous Oral Production Task (SPEAK test) were used to compare differences in subjects' accuracy of perception and production of inflections in different modes of performance (see Appendix A for data results and Appendix B for test items). Ten L1 Mandarin-speaking advanced<sup>13</sup> speakers of English were tested and scored at a mean 96% correct on the written GJT (see Appendix A2) indicating target-like knowledge of the L2 grammatical inflections under study. However, these same subjects scored at a mean 82.7% correct on the ASPT (see Appendix A3) with each participant scoring considerably lower on several of the ASPT items than on the GJT items, which first suggests that an advanced knowledge of L2 grammatical inflections does not automatically equal an equally high level of perception of those same inflections. Secondly, the subjects showed considerably lower perception accuracy of regular past tense *-ed* items that contained the allomorph [t] (58% correct) when compared to items that contained allomorph [Vd] (95% correct), which is reflective of previous research into the effects of acoustic perceptibility of past allomorphs (cf. Klein et al., 2004; Solt et al., 2004). That said, the subjects

---

<sup>13</sup> The subjects in the pilot were advanced graduate students (mean age 23) studying at an American university for at least one year and had already achieved the required TOEFL scores needed to be admitted to the university without English remediation requirements—that is, they all scored over 80 on the Internet-Based TOEFL with above university-target listening, speaking, writing, and reading subscores for unconditional admission to their academic programs.

were still able to accurately perceive (i.e., comprehend) L2 inflections well above chance overall even when items were presented only in aural form.

The results of the semi-spontaneous oral SPEAK task revealed that the same subjects scored at a mean 60.3% correct (see Appendix A4) with each participant scoring much lower in 3<sup>rd</sup> singular and past regular inflectional accuracy in the SPEAK items than on GJT and ASPT tasks. Because the SPEAK task consisted of spontaneous speech samples, it did not reflect the same items and cannot be compared directly to the GJT and ASPT tasks. Nevertheless, even when a L2 learner reveals target-like knowledge of L2 inflections, obvious increased performance pressures due to the timed spontaneous nature of this speaking task reduced accuracy of L2 inflections considerably. The results of this pilot study are suggestive of previous findings reported above that advanced knowledge of grammar does not automatically lead to advanced levels of perception (i.e., comprehension) and production (e.g., Hopp, 2009).

Overall, this pilot study shows that knowledge of grammatical inflections does not entail an equal level of perception of the same inflections; and a high level of perception of inflections does not entail an equal level of production of the same inflections. The results also suggest that in order to systematically study the causes of inflectional variability in L2 learners, the tasks must be systematically controlled in terms of item type, length, complexity, and salience, and they must account for the effects of processing pressures such as those observed in semi-spontaneous speaking tasks (e.g., SPEAK task).

Thus, this dissertation builds upon this pilot to include the following considerations: 1) the possible performance factors that cause the variable omission of inflections; 2) the inflections that are most susceptible to mis-perception and mis-production; and 3) the degree to which other

factors such as L1 phonotactics and proficiency level of the subjects affect the perception and production of L2 inflections.

### 3.3 Possible Performance “Deficits” Assuming Underlying Knowledge

Returning to our model in Figure 1, this dissertation assumes that competence underlies performance, and that once grammatical competence has been acquired, certain types of systematic grammaticality judgment tasks can be used as a gauge of grammatical knowledge since syntactic competence is not directly observable through the surface morphology (cf. Martohardjono, 1998). However, as the pilot study revealed along with several other performance accounts of L2 inflectional variability, target-like accuracy of L2 inflections does not automatically entail that the L2 learner’s accuracy rates will be the same across all modes of performance. In fact, it is these asymmetries that Trenkic (2007) claims are evidence against strict representational deficit accounts, and it is these asymmetries in perception and production of L2 inflections that must be closely examined to reveal the reasons for variable suppliance of L2 inflections. The following sections detail the various possible processing breakdowns proposed above and how these factors are might provide evidence for a performance account of inflectional variability.

#### 3.3.1 L1 Phonotactic Interference

One possible factor that could interfere with a successful parse is L1 phonotactic constraints such that the coda clusters of the past regular *-ed* or 3<sup>rd</sup> singular *-s* in English may represent illicit word-final clusters in the L1. In his discussion of transfer in second language phonology, Major (2008) states, “One of the reasons that L2 learners cannot perceive of

differences in the L2 is that...they transfer their L1 perceptual systems when hearing the L2s” (p. 75). In other words, as some research has confirmed (cf. Brown, 1998; Davidson, 2005, 2006a, 2006b; Dupoux et al., 1999; Goldrick, 2004; Kabak & Idsardi, 2007), perception of the L2 could be significantly affected by the L1 such that successful parsing of the L2 could depend on the extent to which the perceptual mechanism is being informed by the L1 phonotactics.

Accordingly, for L2 learners whose L1 phonotactic constraints disallow consonant clusters in the syllable coda (e.g., Mandarin Chinese speaking subjects), English inflections on regular verbs such as *stopped* and *walks* that are already more susceptible to mis-parsing (per the processing model described above) should be problematic in both perception and production. As such, lower accuracy of perception and production of word-final consonants and consonant clusters that are disallowed in the L1 syllable coda should be observed within the aural and oral performance modes, resulting in the possible deletion of some of the target allomorphs such as [t], [d], [s], and [z], especially when the L1 phonotactic constraints disallow syllable-final obstruents. In other words, this researcher claims that omission of target inflections could be a result of L1 interference at the syllable-level due to a mismatch in L1 and L2 phonotactics, rather than deficits at the morphosyntactic level. Additionally, considering Davidson’s (2005, 2006a, 2006b) claims regarding L2 syllable repairs, it could also be the case that L2 inflections are not deleted in any significant way, but that the word-final consonant clusters that contain L2 inflections could be repaired according to L1 syllable constraints. As discussed in Chapter 4 and 5 below, the design of the tasks and the analysis of the results account for these possible word-final cluster repairs.

### 3.3.2 Perceptual Saliency of Inflections

Additionally, it is also possible that the L2 learner perceives the verb stem while simultaneously not regularly perceiving the target morpheme when the bound morpheme is less perceptually salient—that is, the bound morpheme could be less acoustically robust in the input and, therefore, lead to variable accuracy in performance. Let us take the example we already mentioned of the regular past tense, which is represented by three surface phonetic forms or allomorphs: the less salient non-syllabic [t] as in *walked* and [d] as in *caused* and the more salient syllabic form [Vd] as in *wanted*. Of course, these bound morphemes are less salient (i.e., less acoustically robust) than unbound tense carriers such as the copula, but even within the various allomorphs of the past morpheme *-ed*, there are degrees of saliency that likely affect variable processing of the inflection. It is possible that the two less salient non-syllabic forms are not regularly perceived by L2 learners (cf. Klein et al., 2004; Solt et al., 2004). As such, the non-syllabic word-final allomorphs that close the non-syllabic coda consonant clusters [Ct], [Cd], [Cs], and [Cz] should be less salient and, therefore, less accurately perceived than the syllabic inflections that create the final syllables [CVd] and [CVz].

Extending this notion of acoustic perceptibility, it may also be expected to see a difference in accuracy between codas ending in the obstruents [t] and [d] and the obstruents [s] and [z] due to a higher degree of perceptual saliency afforded by sibilants, which have been shown to have acoustically more energy leading to a higher degree of sonority than the stops (see Ladefoged, 2006 for a discussion of the difference in sonority regarding sibilants vs. stops). Finally, it may also be expected that there would be an effect of voicing leading to different rates of accuracy between the allomorphs [t] and [d] and between [s] and [z] with the voiced [d] and

[z] being more accurate due to their higher degree of acoustic robustness in the voiced inflections (again, see Ladefoged, 2006).

Additionally, if the effects of syllabicity and sonority are observed in perception on specific allomorphs (e.g., [Vd] more accurate than [t] and [d]) or with regard to the relative sonority of different codas (e.g., [s] more accurate than [t]), these effects may in fact mirror a reduction of accuracy in their production in the word-final target consonant clusters. For example, if non-syllabic allomorphs are less accurately perceived than syllabic allomorphs, then it is likely that these same non-syllabic allomorphs will also reveal reduced accuracy in production. With regard to sonority, reduced accuracy in perception of less sonorous codas [t] and [d] may also be mirrored in reduced accuracy of production of these same codas. This dissertation is not specifically concerned with the underlying phonological or derivational principles that may explain the acquisition of these contrasts—it is, however, concerned with the patterns by which these differences are exhibited with regard to their specific effects on the target codas, as these effects may be a possible reason for surface variability of L2 inflections. These possible effects will be examined within this study.

### 3.3.3 Processing Load Effects on Performance

The third possible factor that could cause variable performance is increased processing load or processing pressure due to increased task demands, which can lead to a reduction in attentional resources that affect the processing of a L2 learner's input and output. Processing load effects on performance are expected when increased levels of processing are required within an item or task, following evidence in the literature previously described. Higher processing load could reduce access to L2 grammatical representation leading to more variable performance on

certain task items. As such, considering the attested effects of memory load on performance, this dissertation will investigate the claim that L2 learner's accuracy of inflections decrease when perceiving and producing longer, more complex sentences compared with their accuracy when perceiving and producing shorter, less complex sentences within the same performance mode.

#### 3.3.4 L2 Proficiency Level and Overall Performance

Finally, assuming *Full Access to UG* for L2 learners, this researcher presupposes that under the right conditions, ultimate attainment of L2 inflection should be possible (following Hopp, 2009). As such, L2 learners at an advanced and especially near-native L2 proficiency level should exhibit native-like inflectional accuracy under the right conditions, even when the L2 grammatical features are not instantiated in the L1 grammar. Considering the possible phonotactic and performance effects on inflectional variability outlined thus far, this researcher would also expect to see these effects be reduced as a L2 learner reaches an advanced to near-native level in the L2. Regarding L1 phonotactic interference, this result is posited by Davidson's (2005, 2006a, 2006b) OT approach such that under the right input conditions, native-like production of the target word-final consonant clusters would be observed. In fact, she observes this proficiency-level effect in her subjects. As such, this researcher assumes that perception of the L2 also increases with proficiency level, and consequently, the effects of salience on word-final consonants should decrease as L2 perception becomes more native-like, which would reveal increased accuracy of perception of L2 inflection and target word-final consonants. Finally, regarding performance effects such as processing overload that leads to reduced attentional resources and increased inflectional variability as observed in Hopp (2009), it is again expected that processing effects would be mitigated as L2 proficiency level increases to the near-

native level. Thus, testing subjects at an advanced to near-native proficiency of L2 inflection should reveal the same results—that is, the mitigation of phonotactic and performance factors on L2 inflection and illicit word-final consonant clusters as the L2 learners reach a near-native level of L2 proficiency.

### 3.4 A Comparison of Mandarin and English: Possible Interference Effects

In order to test the possible effects of the L1 on the L2 in terms of the factors that have been described thus far, this dissertation focuses on two languages that differ significantly in the representation of their phonology and morphology and whose learners have already been studied to some extent, as described in Chapter 2: Mandarin Chinese L1 speakers learning English as a L2, specifically adult Standard Mandarin speakers from China. In order to illustrate the similarities and differences of the languages under study and their relation to the purpose of this study, some of the differences between the two languages and the potential interference effects that could occur between the L1 and the L2 are described below.

#### 3.4.1 Morphological and Morphosyntactic Differences

Mandarin Chinese does not represent tense<sup>14</sup> in the grammar (cf. Lin, 2005). Mandarin also does not use overt number/person agreement inflectional morphemes such as *-s* or *-es* as is the case with English 3sg present tense subject/verb agreement. In general, Mandarin has only derivational morphemes and aspectual morphemes to represent agreement and temporality in the verb. Verbs can also be two (sometimes three) bound roots that are combined into compounds with the head verbal root on the left and the resultative root on the right as in the following:

---

<sup>14</sup> Lin (2005) claims that there is no tense node (TP) available in Chinese, but Chinese uses an aspect node (AspP) to represent temporal constructions.

- (6) zuò-wán  
*do-complete*  
“to finish”

Example (6) shows the most common verbal root+root form in Chinese and is the result of combining two roots that can be combined with a multitude of other roots to create many different characters. Verbal roots can also be combined with aspectual morphemes such as the perfective suffix *-le*, which indicates among other things<sup>15</sup> the aspectual relation of completeness as follows in (7):

- (7) tā chī-le liǎng-wǎn fàn  
*s/he eat-PFV two-CL rice*  
“He ate two bowls of rice.”

Chinese also does not attach 3sg agreement inflection to verbs, but instead, the agreement is understood by the context as in the following example (8) containing the preposition *tā*, which conveys the *number* (neutral *gender*) of the subject without overt agreement on the verb:

- (8) tā màn-màn-de pǎo  
*s/he slowly-CL run*  
“S/he runs slowly.”

Many of the meaning relationships that are understood between subjects and predicates are accomplished either by syntactic juxtaposition of specific characters or by the limited aspectual or derivational morphemes. Thus, learning English bound morphology for Mandarin speakers could be challenging since Mandarin morphology exhibits several differences from English morphology. In fact, there can be no positive transfer of L1 morphology to English in terms of specific grammatical features such as tense and 3sg agreement. The only potential positive transfer from the L1 to the L2 for such learners could be the fact that Mandarin has bound

---

<sup>15</sup> Sun (2006) notes that this marker can also be used in a series to denote the number of its occurrence and with future tense. It can also be used to represent an entirely different root with a different meaning (e.g., *liǎo* – “to finish”).

morphemes, so attaching morphemes to English verbal roots is a familiar morphosyntactic principle within their L1 grammars. Nevertheless, because Mandarin does not represent tense, we investigate evidence of this potential area of negative transfer and attempt to show that even though Mandarin does not instantiate tense in the grammar nor does it inflect for 3sg present tense subject-verb agreement, these differences in the grammar do not inhibit the representation of L2 inflection for the Chinese subjects in this study.

### 3.4.2 Phonological and Phonotactic Differences

Mandarin Chinese restricts stops /t/ and /d/<sup>16</sup> and fricatives /s/ and /z/ from the syllable coda position within its syllable structure. In fact, in addition to vowels [i, u, o], Mandarin only allows nasals /n/ and /ŋ/ and variably /z/<sup>17</sup> in the syllable coda (cf. Sun, 2006). Additionally, Mandarin syllable structure disallows consonant clusters in onset and coda, so English syllable-final consonant clusters that include past tense or 3sg present tense endings are prohibited by Mandarin L1 syllable structure constraints (cf. Li & Thompson, 1987). For example, the three allomorphs of the past tense in English and their corresponding coda clusters (i.e., [Ct], [Cd], and [CVd]) create syllable structures that are illicit according to Mandarin phonotactic constraints.

The design of this study systematically examines these effects of L1 phonotactic constraints on the L2 inflections in both perception and production. Thus, it is possible that if Mandarin L1 speakers cannot easily perceive or produce the consonants and consonant clusters of English codas that contain target inflections, then lower accuracy rates of those target

---

<sup>16</sup> Actually, the voiced stop /d/ is represented in *pinyin* phonetically as unaspirated [t] or orthographic ‘d’ vs. the aspirated [t<sup>h</sup>] or orthographic ‘t’. In fact, there is no voicing distinction in the phonological inventory of Mandarin (cf. Sun, 2006). As such, the distinction between ‘s’ and ‘z’ in *pinyin* is also represented by a phonetic distinction [s] and unaspirated [ts].

<sup>17</sup> This voiced retroflex fricative is variable due to the many varieties of Mandarin speakers in China. Nevertheless, stops and fricatives are restricted in the syllable coda for standard Mandarin speakers.

inflections could be attributable to phonotactic mis-parsing, leading to variable perception and production of the target inflections within the syllable codas.

### 3.4.3 Bound Morphemes in Mandarin and English: A Comparison of Saliency

If we compare English and Mandarin surface inflectional morphology directly in terms of saliency (i.e., acoustic robustness) of bound verbal morphemes, Mandarin, for example, has bound syntactic verbal morphemes (e.g., Aspectual morphemes *-le*, *-guo*, and *-zhe*) that are relatively more acoustically robust than English verbal morphemes, especially the non-syllabic English allomorphs [t], [d], [s], and [z]. The bound morphemes in Mandarin are always syllables of at least CV in structure, and although they are not assigned a lexical tone, they are full syllables and even full syntactic words.<sup>18</sup> English, on the other hand, has bound inflectional morphemes for past tense (*-ed*) and 3sg agreement (*-s*) that are limited in their application as follows: a) They do not appear on certain verbs such as irregular past or 3<sup>rd</sup> plural agreement verbs, respectively; b) they are non-syllabic in many instances; and c) they do not receive any stress. Thus, the perceptibility of most English surface inflections by Mandarin speakers may be variable due to the fact that many English inflections are highly variable in the input and are acoustically less robust in comparison with the more consistently applied and more acoustically robust bound inflectional morphology of Mandarin.

---

<sup>18</sup> Packard (2000) describes the aspectual morphemes in Mandarin as having full lexical status such that they can stand alone and can fill syntactical nodes unlike the English bound morphemes, which are only features that attach to lexical items in the syntax.

Another potential challenge for these L2 learners is the weak perceptual salience (i.e., acoustic robustness) of non-syllabic, regular past tense inflectional morphemes in English speech. Again, according to Klein et al. (2004) and Solt et al., (2004), perceptual salience<sup>19</sup> of the final past morpheme [t] as in *walked* or [d] as in *closed* was significantly lower than for the more salient syllabic past morpheme [Vd] as in *waited*. Although not investigated by Klein et al. (2004) and Solt et al., (2004), the perceptual salience of 3sg inflections may also be variable for Mandarin speakers such that the non-syllabic final [s] as in *walks* and final [z] as in *hangs* may be significantly less salient than the syllabic [Vz] allomorph in *causes*.

Thus, the L1 phonotactic interference effect is likely compounded by the fact that the regular past tense and 3<sup>rd</sup> singular agreement morphemes in English are variably represented in the surface structures. For example, English represents past regular *-ed* in the surface structures as the following: add [t] if the verb ends in a voiceless consonant; add [d] if the verb ends in a vowel or voiced consonant; and add [Vd] if the verb ends in an alveolar stop /t/ or /d/. Thus, in order to test for the possible performance effects of this variability, the study in this dissertation includes items that represent the various surface structures of English syllable codas to determine if the varying structures are affected differentially by the L1 phonotactics and/or salience of the inflections. Consequently, because L1 phonotactic constraints disallow all of the consonant clusters that are necessary for perceiving and producing English inflections, if we were to see a significant difference in the perception and production accuracy of syllabic vs. non-syllabic inflections (e.g., [Vd] vs. [d] or [t]), it is likely attributable to the perceptual salience of the inflection type.

---

<sup>19</sup> Note that Klein et al. (2004) state in their *Perceptual Salience Hypothesis* that “a second language learner will perceive and produce a syllabic grammatical suffix more accurately than a non-syllabic grammatical suffix because a syllable is more perceptually salient than a consonant (or cluster of consonants).”

### 3.4.4 L1 Mandarin Speakers' Perception and Production of L2 English

What do these facts about the L1 and L2 tell us about Mandarin speakers' performance with regard to English inflections? Mandarin a) has no T(ense) node in the syntax; b) does not inflect for tense or agreement; c) has L1 phonotactic constraints that disallow the types of codas that are necessary for the perception and production of English surface inflections; d) and has consistently more salient syllabic bound morphemes than the variable syllabic and non-syllabic inflectional morphology of English. Furthermore, as was noted, the tense and agreement features required in the L2 must be available in the competence for successful parsing of the L2 because these features are not transferable from the L1.

Thus, this researcher deduces that the English inflections and English syllable structure for Mandarin speakers cannot be transferred from the L1 but claims that acquisition of L2 inflection and syllable structure is possible for adult L2 learners. Thus, if L2 inflection is shown to be target-like, but production of word-final target consonant clusters is still variable, it is claimed that performance factors and L1 phonotactic constraints may still be affecting accuracy of surface inflections. It is also possible that the subjects' L2 perceptual mechanism is being informed by their L1 as some researchers have suggested (cf. Brown, 1998; Dupoux et al., 1999; Jia et al., 2006; Kabak & Idsardi, 2007; Major, 2008) causing reduced accuracy of perception of target inflections as well. Finally, perceptual salience of target inflections in the input could be low as compared to Mandarin bound morphemes, resulting in variable perception and production of English bound verbal morphemes even when the L2 learners' knowledge of the inflections can be attested. In fact, this researcher claims that even after Mandarin speakers acquire the knowledge of L2 inflection such that they produce these inflections at ceiling on grammaticality judgment tasks, these intervening factors may continue to cause variable perception and

production of English inflections due to reduced access to the grammar within the aural and oral modes of performance due to the described performance and processing factors.

Mandarin speakers' knowledge of English inflections has been attested in several studies. Building on this claim that English L2 inflection is acquirable by L1 Mandarin speakers, this study reveals that even when an advanced knowledge of target inflections is attested, there are interceding factors that can affect the L2 learners' ability to perceive and produce English inflections, and these factors are not part of the syntactic or prosodic representation. This study systematically examines these factors.

### 3.5 Research Questions and Hypotheses

The purpose of this dissertation is to shed light on a performance-based account of L2 learner inflectional errors by systematically investigating the relationship between perception and production and the potential factors that lead to variable production of surface inflectional endings outside of L2 morphosyntax. In fact, in representational deficit accounts, namely the Representational Deficit Hypothesis and the Prosodic Transfer Hypothesis, those authors claim that features of the L2 grammar are unattainable by adult learners if their L1 grammars do not instantiate these features. However, it seems clear from the research to date that examining L2 inflectional errors in spontaneous production does not provide sufficient evidence for absence of morphosyntactic or prosodic representations in the L2. We have also seen that even for L2 learners who cannot positively transfer grammatical features such as tense from the L1 to the L2 (i.e., because the features are not instantiated in the L1), these features can still become available in the L2 grammar.

Again, it is proposed that L2 error patterns in the perception and production of inflectional morphemes are not a question of a permanent deficit within the L2 competence, but that variable perception and production of inflectional endings is likely related to performance factors and other factors that reduce access to the L2 grammar rather than to the availability of L2 grammatical representations. Specifically, this dissertation examines the claim that performance can be affected by 1) L1 phonotactic interference effects; 2) salience of the inflections (i.e., acoustic robustness of the input); 3) processing load effects on performance such that certain task items should require relatively higher attentional resources within the L2 learner's processing mechanism; and 4) proficiency level such that when L2 learners reach a near-native level of L2 inflection, the effects of the above factors will be significantly mitigated.

### 3.5.1 Research Questions

Considering the possible reasons for L2 inflectional variability, the general research questions from Chapter 1 are revisited and are restated with more specificity as follows:

- 1) First, what role do L1 phonotactic constraints play in the variability of perception and production of English inflections, and is the mismatch between L1 and L2 phonotactics a possible cause for the well-attested omission of inflections in spontaneous L2 speech?
- 2) Second, what is the effect of salience (i.e., *acoustic perceptibility*) of English inflectional allomorphs on the perception and production of the corresponding L2 inflections? Specifically, are syllabic inflections more salient than non-syllabic inflections (e.g., [Vd] vs. [d] in pairs like *wanted* vs. *caused*; [Vz] vs. [z] in pairs like *causes* vs. *reads*) and will more sonorous word-final segments

(e.g., [s] and [z]) be more accurately perceived and produced than the less sonorant segments (e.g., [t] and [d])?

- 3) Third, what non-representational processing factors might account for inflectional variability? For example, would omission of inflections still be a common error when task demands are considerably reduced, as they would be in a controlled production task where lexical selection is predetermined? And would length and complexity of utterances have an effect on the perception and production of inflections in these controlled tasks?
- 4) And finally, if L2 inflection is not subject to a critical period and is attainable as Hopp (2009) claims, what is the effect of proficiency level on perception and production in terms of the following factors?
  - a. Considering Davidson's OT account, if all possible constraints are available in the grammar, but variability is proportional to the strength of ranking within a L2 learner's new grammar, how would L1 phonotactics affect the perception and production of illicit word-final L2 consonant clusters as L2 proficiency increases?
  - b. And in terms of acoustic perceptibility of word-final L2 consonant clusters, assuming L2 learners' perception of these word-final clusters also increases as they become more proficient in the L2, would this increased perceptibility correlate with increased production of the same target clusters in the L2?

### 3.5.2 Hypotheses

The following hypotheses are posed to answer the above research questions.

1. L1 phonotactic constraints should negatively affect the perception and production of L2 inflections such that when a syllable coda violates the L1 phonotactic constraints, performance accuracy for that coda should be significantly reduced in both perception and production. Accordingly, the reduction in performance for any codas that contain target inflections should be exhibited via a repair of the word-final cluster, resulting in deletion of one or more consonants in the cluster or in schwa epenthesis before or after the word-final consonant that reflect the L1 phonotactic constraints.
2. Lower salience (i.e., *acoustic perceptibility*) of L2 inflections due to differing degrees of syllabicity should have a negative effect on performance such that reduced accuracy should be observed in both perception and production for the less salient (i.e., non-syllabic) allomorphs ([t], [d], [s], and [z]) than for the more salient (i.e., syllabic) allomorphic variants ([Vd], [Vz]). Accuracy is also expected to be lower for stops ([t] and [d]) when comparing accuracy rates of the stops ([t] and [d]) to accuracy rates of the fricatives ([s] and [z]), and lower on the non-syllabic, unvoiced inflections ([t] and [s]) compared to their non-syllabic, voiced counterparts ([z] and [d], respectively).
3. Accuracy rates of target L2 inflections and word-final consonant clusters should be higher on items that require lower degrees of processing resources (e.g., shorter, less complex sentences) than on items that require higher

degrees of processing resources (e.g., longer, more complex sentences) within the same task.

4. Following Hopp (2009), native-like attainment of L2 inflection is possible, but certain non-morphosyntactic factors still have an effect on performance.
  - a. Thus, just as production of L2 inflection can reach near-native levels, the higher proficiency level subjects in this study should produce the L2 inflections and word-final consonant clusters at or near ceiling when processing resources are not diminished, revealing that L2 phonotactic constraints can be acquired (following Davidson, 2005, 2006a, 2006b) even when they are not instantiated within the L1 grammatical representation.
  - b. Further, assuming near-native perception of L2 inflection and word-final consonant clusters is also possible, the higher proficiency level subjects in this study should perceive the target L2 inflections and word-final consonant clusters at or near ceiling, again, when processing resources are not diminished. As such, accuracy rates in perception and production should correlate such that accuracy rates in perception of specific allomorphs should correlate to similar accuracy rates of production of those same allomorphs; and as L2 learners' proficiency levels approach ceiling, non-significant differences in accuracy rates across the aural and oral performance modes should be observed.

## Chapter 4: Methodology

### 4.1 Subjects

The subjects (n=60) were native Mandarin Chinese speakers living in Beijing, China. The initial subject pool consisted of two groups of 30 subjects each as determined by their English proficiency levels via their college majors and grade level and via results on the College English Test (CET), Band 4, which all Chinese university non-English major students are required to take yearly and the Test for English Majors (TEM), Band 4, which is required for English majors to advance to junior year in their undergraduate English studies. The purpose of the CET-4 test is to provide a measure of the general English proficiency of undergraduate students in China to ensure that Chinese undergraduates reach the required English levels specified in the National College English Teaching Syllabuses (NCETS). The CET-4 is mandatory for all university students in China who are not English majors. Proficiency in English is also a prerequisite for a bachelor's degree in China (Zheng & Cheng, 2008). The CET-4 measures English proficiency in listening, reading, writing, and speaking with specific tasks that are relevant to this study, including cloze tests that require subjects to listen for and accurately write the target inflected verbal items such as those used as target items in this dissertation (e.g., *cleaned*, *counted*, *checked*). This task is scored on accuracy of the suppliance of the target items. The CET also includes timed semi-spontaneous spoken picture description tasks in which the subjects must describe a picture sequence in the past tense, for example. Analytic scoring rubrics for the semi-spontaneous speaking prompts include specific criteria including accuracy of form (i.e., grammatical form), as well as cohesion and coherence of the discourse. The TEM-4 also tests the above language modes via similar tasks, but it is considered a much more challenging test

because it is designed for English majors who are required to graduate with much higher English proficiency scores because they will be translators and English teachers or go on for graduate studies in advanced communications or English Literature.

The two groups were separated into one group of 30 non-English-major intermediate to advanced speakers (e.g., freshmen who have taken the CET-4<sup>20</sup>) and the other group of 30 English-major advanced speakers (e.g., juniors who have passed the TEM-4 with a score of 70 or above, which is the required cutoff score to be considered for graduation as English majors). Note also that students who take the TEM-4 have already exceeded the top scores on the CET-4. In the final analysis, twenty subjects had to be excluded due to non-completion of the tasks in this study, resulting in a final group of 40 subjects (19 lower; 21 higher). All of the subjects were college-aged students (mean age=19; range=17-21) who had been taught English by Chinese non-native English speakers via grammar instruction, vocabulary memorization, reading aloud, reading comprehension, oral recitation training, listening comprehension, aural cloze tasks, and some short extemporaneous writing in China.

On post-testing questionnaires, the lower group (e.g., freshmen) reported having studied English for a period of 8-9 years while the higher group (e.g., juniors) had studied English for a period of 10-11 years, with the juniors having had two more years of daily intensive English instruction in college than the freshmen.<sup>21</sup> The post-testing questionnaires also revealed that every subject had little or no English native-speaker contact, as all of their instructors were non-native speakers. However, the higher group had a much higher level of English due to the fact

---

<sup>20</sup> The CET-4 scores were unavailable for these lower subjects. However, all students at Chinese universities must pass English entrance exams that require them to perform in English at a level that is acceptable to be able to listen, read, write, and speak in English at a college level, as the English classes that all students are required to take are taught mostly in English. Thus, the lower level subjects were determined to be at English level that is sufficient to understand and speak college-level English because they had already taken the CET-4 at least one time.

<sup>21</sup> In fact, the non-English major freshmen receive only 4-5 hours of English instruction per week because English is not their majors. Conversely, the junior English majors receive 15+ hours of intensive English instruction per week in college due to the fact that English is their chosen major.

that they were all English majors studying for careers requiring advanced English proficiency. Nevertheless, all reported on the questionnaire that they had no more than one week of prior English native-speaker contact, and none had traveled outside of China. Considering the above background of these subjects, they were chosen for two main reasons: 1) the amount and type of English instruction were much easier to control as they have all had standardized English instruction from age 10 or 11 (per the Chinese governments' English teaching standards) with little to no native English speaker contact; and 2) the distinction in proficiency level between the two groups was easily determined by their grade level (i.e., freshmen vs. juniors) and by their majors (i.e., non-English majors vs. English majors). The clear distinction between proficiency levels was critical to test proficiency level effects on aural perception and oral production of target L2 inflection even though it was presupposed that all subjects would have an advanced knowledge of the target inflections due to the similar nature and length of academic classroom-based English instruction that they had all received previously.

#### 4.2 Materials and Procedures

Figure 2 summarizes each task by type and its intended purpose. The procedure and rationale for each task is explained below.

Mode	Task	Purpose
A. Morphosyntactic knowledge	1) Written Grammaticality Screening Tasks (WGS)	Knowledge of grammatical inflections – used as a screening task to rule out subjects that did not have advanced knowledge of target inflections.
B. Experiment 1: Aural Perception	2) Aural Word Perception Task (AWP)	a) Effect of L1 phonotactic constraints on perception of L2 codas in isolation. b) Effect of syllabicity and sonority on perception of L2 codas in isolation.
	3) Aural Sentence Perception Task (ASP)	a) Effect of L1 phonotactic constraints on perception of L2 codas within inflected items. b) Effect of syllabicity and sonority on

		perception of L2 codas within inflected items. c) Effect of processing load on perception of L2 codas between longer, more complex and shorter, less complex sentences.
C. Experiment 2: Oral Production	4) Oral Word Production Task (OWP)	a) Effect of L1 phonotactic constraints on production of L2 codas in isolation. b) Effect of syllabicity and sonority on production of L2 codas in isolation.
	5) Oral Sentence Production Task (OSP)	a) Effect of L1 phonotactic constraints on production of L2 codas within inflected items. a) Effect of syllabicity and sonority on production of L2 codas within inflected items. b) Effect of processing load on production of L2 codas between longer, more complex and shorter, less complex sentences.

This dissertation investigated the relationship between perception and production by testing the subjects' ability to perceive and produce specific inflections of past tense and 3sg agreement. Specifically, the study tested the effects of L1 phonotactic constraints on the L2 codas to determine to what extent the L1 phonotactic constraints affect both perception and production of the target codas in the listening and speaking modes. The study also tested to what extent salience of the input and processing pressure due to task demands affect performance. To determine the relationship between perception and production and to determine the effects that L1 phonotactic constraints, salience, and processing pressure have on the performance of specific inflections, the researcher conducted the following experiments.

#### 4.2.1 Written Grammaticality Screening Task

An un-timed Written Grammaticality Screening (WGS) Task was administered to all of the subjects to determine their knowledge of regular past tense and 3sg subject/verb agreement and to determine if any subjects should be excluded from the study. There were two parts: 1) in the first part, they chose between inflected and uninflected verbs in 24 target sentence; and 2)

second, they had to conjugate and write the verbs in another 24 target sentences. Both tasks included copula and non-verbal filler/distracter items such as the following: *Last year, the woman used to walk in the park by \_\_\_\_\_ . (herself / himself)* Because the study was concerned with the question of whether the subjects have the knowledge of the inflections in their grammars (and not a specific *knowledge type* as described above), the task was un-timed. These two grammar-screening tasks required a written response for each item before the subjects could continue. The subjects were also not able to go back to prior questions.

#### 4.2.1.1 WGS, Part 1

The following example is from WGS, Part 1:

(9) *The boy always (asks / ask) for a cookie after dinner.*

In Part 1, the verb pairs were varied for the six inflection-types under study: i.e., regular past tense, unvoiced [t] (*walk/walked*); regular past tense, voiced [d] (*caused/cause*); present tense, 3<sup>rd</sup> person singular, unvoiced [s] (*ask/asks*); present tense, 3<sup>rd</sup> person singular, voiced [z] (*reads/read*); syllabic past tense [Vd] (*wait/waited*), and syllabic present tense [Vz] (*fixes/fix*). Four tokens were given for each type. Presentation of inflected and uninflected verbs were pseudo-randomized and varied systematically between short sentences (8-11 syllables) and long sentences (12-18 syllables). Short sentences contained simple NPs in subject position (e.g., the boy), and long sentences contained complex NPs in the same position (e.g., the boy with the big red bike). There were 12 short sentences and 12 long sentences. To avoid possible phonological interference due to word-boundary blending (e.g., *Yesterday, he talk\_\_ to....* [tɔk.tu]), every target verb was followed by a word with a vowel onset such as *Yesterday, he talked on the phone*. Tasks were developed to allow for long and short items and simple and complex subjects

to appear at random for every participant, so no participant encountered the same order of items in any task. Participants were asked to silently read sentences on a computer screen, choose between two verbs (inflected and uninflected), and type their answer before moving on to the next item, which appeared upon pressing the spacebar.

#### 4.2.1.2 WGS, Part 2

WGS, Part 2 consisted of simple sentences in which subjects were required to conjugate infinitive verb forms, as shown in example (10) below:

(10) *Yesterday, a big dog in the park \_\_\_\_\_ a young boy on a bike. (to chase)*

The design followed that of WGS-1. Verbs were varied for the same six inflection-types; and 12 short and 12 long sentences were included, and as described above in WGS-1, items were randomized and controlled for phonological interference effects. WGS-2 again required subjects to read sentences on a computer screen, but in this task, the subjects saw the infinitive form of a verb (e.g., to chase) under the target sentence and were required to conjugate the verb to correctly match the sentence and then type their answer. Both tasks contained an equal number of filler/distracter items such as copula and modal auxiliaries as well.

The purpose of the WGS tasks was to reveal the subjects' underlying knowledge of 3sg agreement and regular past tense in English. It was expected that all of the subjects would have an advanced knowledge of the target inflections due to their long exposure to grammar-translation instructional methods in China. However, because the study relied on the subjects having an advanced knowledge of the grammar, this grammatical screening task was used to exclude any participant who did not have advanced knowledge of the L2 inflections. None of the

subjects were excluded due to lack of advanced knowledge of the target grammatical forms as is shown in the results (see Chapter 5).

#### 4.2.2 Experiment 1: Aural Perception Tasks

Experiment 1 sought to measure subjects' aural perception of inflections and to determine the effects of L1 phonotactic interference on target codas (e.g., [Vd], [t], [d], [Vz], [s], [z]) as well as to determine whether there were differential effects of perceptual salience (i.e., acoustic perceptibility) between syllabic and non-syllabic inflections. Two tasks were included: a word-based (Aural Word Perception or AWP) task and a sentence-based (Aural Sentence Perception or ASP) task. The ASP task also measured the effects of sentence length on perceptual accuracy in order to examine if higher processing pressure or processing load would significantly reduce accuracy of the target inflections.

##### 4.2.2.1 Aural Word Perception Task

The Aural Word Perception (AWP) task was administered first. This task was an AX word-pair discrimination task that contained 24 pairs of target inflected real verbs (12 same and 12 different – e.g., *walked/walked* and *walked/walk*). The verb pairs were varied for the six inflection-types under study: i.e., regular past tense, unvoiced [t] (*smoke/smoked*); regular past tense, voiced [d] (*called/call*); present tense, 3<sup>rd</sup> person singular, unvoiced [s] (*look/looks*); present tense, 3<sup>rd</sup> person singular, voiced [z] (*holds/hold*); syllabic past tense [Vd] (*wait/waited*); and syllabic present tense [Vz] (*use/uses*). Four tokens were given for each type. Following prior research (e.g., Dupoux et al, 1999; Kabak & Idsardi, 2007; Goldrick, 2004; Davidson, 2005, 2006a, 2006b; Jia et al., 2006), nonce word pairs (e.g., *duzed/duzed*; *duzed/duze*) were included

that match English phonotactics and mirror the L2 inflections under study to determine if L1 phonotactic effects or syllabicity effects would also surface on non-lexicalized items that contained target coda clusters (cf. Solt et al., 2004). Word pairs were recorded via Audacity by a male native speaker of American English.

There were two types of stimuli: *same* word pairs (e.g., *walks/walks*) and *different* word pairs (e.g., *walks/walk*). The *same* pairs were generated by copying both items from one recorded file, and the *different* pairs were generated recording one word (e.g., *walks*), and then using Audacity, the final inflection (e.g., [s]) was removed from speech stream and the *walk* was spliced to a the original *walks* to create the AX pair. The spliced stimuli were closely matched in their amplitude and length. The possibility that the *different* items contained co-articulated acoustic cues that would inform the subjects that an item was incomplete such as the lack of pitch dropping at the end of the spliced word was considered. However, because the study was focused on whether the subjects would simply perceive the inflection or not, this effect was regarded as potentially negligible because this researcher was not conducting a specific acoustic level analysis. Nevertheless, if there was an effect, higher accuracy rates would have been expected for the truncated *different* items compared to the *same* items due to the leftover acoustic clues caused by truncation (e.g., vowel duration and pitch that would increase acoustic differences); however, accuracy rates were higher on the *same* items allowing this researcher to conclude that truncation on different items had no apparent effect on accuracy. Accuracy rates were calculated by the number of correct decisions (either *same* or *different*) over the total number of target items in that coda type to obtain averages for each item type.

The Inter-stimulus Interval (ISI) between each of the items in the AX pairs was controlled. The word pairs contained an ISI of 500 ms (*ms* = milliseconds). A good deal of

research in the area of speech processing (cf. Pisoni & Tash, 1974; Werker & Logan, 1985; Mody, Studdert-Kennedy, & Brady, 1997; Demany & Semal, 2005) has determined that tasks that measure aural speech processing (i.e., perception of speech) must control for the time intervals between items especially for AX discrimination tasks due to memory limitations during speech processing. These tasks are commonly used for determining the perception of discrete phonological segments or features that comprise a phonological segment. Thus, because this researcher was concerned with determining the L1 phonotactic effects on the L2 learners' perception of inflections (i.e., phonetic level differences), which have been determined to be accessed within a 500 ms interval, it is important to note the choice of ISI and the rationale for the ISI in this task. The research noted above cites ISIs of durations anywhere between 50 ms to 1500 ms with debates over whether subjects are accessing phonemic or phonetic categories or whether they are using auditory clues such as the duration of the speech signal. However, the purpose in this task was not to test subjects' *rate* of perception of the inflections, but it was simply to test that they are either perceiving or not perceiving the target codas. Thus, it was important that the subjects had sufficient time to process each item in the pairs. However, the time constraint also provided enough processing pressure on the subjects' parsers to mimic the speed of natural speech and human speech processing. An ISI of 250 ms is considered a minimum appropriate length of ISI for speech perception (i.e., auditory and phonetic level differences) for researchers who have employed the AX discrimination task for speech perception experiments, and an ISI of 500 ms has been determined to reveal no significant difference in phonetic-level processing when employing the AX discrimination task (cf. Werker & Logan, 1985).<sup>22</sup> Thus, this researcher decided to use a 500 ms ISI duration. The time interval

---

<sup>22</sup> Recent research in cognitive neuroscience (Sahin et al., 2009) has revealed that there are specific intervals of processing time for lexical, grammatical, and phonological functions within Broca's Area and that the processing of

between the different target word pairs was self-paced in order to reduce the chances of additional processing load effects due to time pressure and working memory overload between the task pairs. The subjects were not able to go back to any items after these had already been presented to them.

Subjects heard each pair once and were instructed to choose between the “same” or “different” keys that were clearly marked on the keyboard. Participants were told that some of the pairs matched exactly and some were different and that they must decide which they heard and then choose the appropriate key on the keyboard (i.e., A for same or L for different). The keyboard was also marked with *same* and *different* accordingly. Additionally, while they were listening to each word pair, the subjects saw the visual prompt, “Choose ‘A’ for same or ‘L’ for different,” on the computer screen to remind them of their choices. These directions were also written in Mandarin and were described orally to them in Mandarin. The subjects completed four sample items prior to the task to ensure that they understood the task procedure. Once they chose “same” or “different,” a screen would appear that directed them to press the spacebar to continue. Their choices were automatically recorded into a text file that was saved to a main computer.

The purpose of this task was to determine if the subjects would accurately perceive the consonant clusters under study. If the subjects reveal high rates of accuracy on the *same* and *different* pairs as a whole, but accuracy rates vary depending on the coda type, it would be assumed that L1 phonotactic constraints are not the cause of reduced accuracy across the board. However, it is possible that L1 phonotactic constraints will reveal asymmetric effects on specific L2 inflectional allomorphs resulting in a coda-specific effect that causes deletion or schwa

---

these different parts of the grammar proceed in a specific order with lexical processing occurring at 200 ms. Thus, an ISI of 250 ms is supported in other areas of cognitive research as well.

epenthesis only on some of the codas. Error patterns may also reveal effects due to varying degree of salience of a coda—in other words, salience effects (i.e., syllabicity or sonority differences) could reveal a specific effect on different coda types. For example, if the subjects perceive a significantly higher number of the allomorph [Vd] over [d], this result would provide evidence for their ability to perceive the more acoustically perceptible syllabic allomorph. Additionally, accuracy on the *same* items (e.g., walked/walked) may be higher than accuracy on the *different* items. The rationale for the claim of higher accuracy of *same* pairs over *different* pairs in this task is as follows: If the subjects' perception is weak for the codas, it would follow that differences in the codas as in [wɔkt] vs. [wɔk\_] in the *different* items would not be perceptible, and these differences would go unnoticed by the subjects, resulting in subjects choosing *same* more often when the items are actually *different*, thus, increasing the error rates on the *different* items. Solt et al. (2004) also found this effect in which their participants were less able to consistently detect the absence of the non-syllabic allomorphs, revealing relatively higher accuracy of *same* items when compared to the *different* items. For this reason, *same* and *different* items were separated in the analysis.

#### 4.2.2.2 Aural Sentence Perception Task

The Aural Sentence Perception (ASP) task was designed to measure not only the effects of L1 phonotactic constraints and salience on subjects' aural perception of inflections, but this task also included equal items of shorter, less complex and longer, more complex sentences to determine the effects of increased processing pressure on subjects' accuracy of target inflections due to sentence length. This task was a refined and extended version of the perception task that

was employed by Solt et al. (2004). The task contained 24 pairs with inflected real verbs as seen below in examples (11) and (12):

(11) Same:

*Yesterday, the boy walked on the sidewalk/Yesterday, the boy walked on the sidewalk.*

(12) Different:

*Yesterday, the boy walked on the sidewalk/Yesterday, the boy walk on the sidewalk.*

Verb pairs were varied for the same six inflection-types under study. Pairs were counter-balanced equally between *same* and *different* and, for the *different* pairs, between inflected first and inflected second (e.g., stopped/stop and stop/stopped). The task also contained 24 nonce pairs (e.g., *Last month, the girl feamp a new dress./Last month, the girl feamped a new dress.*). As in the AWP task, the nonce word pairs were included to determine if L1 phonotactic effects or syllabicity effects would surface on non-lexicalized items that, nevertheless, contained the target L2 inflections. As in the WGS tasks, items were balanced between long and short sentences with simple and complex subjects, and each verb was followed by a vowel to avoid word-boundary consonant blending interference effects. The task also included 16 copula pairs and 24 non-violating pairs as filler/distracters. The results of these fillers/distracters are not reported, as they are not relevant to the study.

As in the AX word perception task, in this task, subjects heard each pair of sentences once and were instructed to choose the “same” or “different” keys that were clearly marked on the keyboard. The subjects hit the spacebar to prompt each new stimulus. Once they chose “same” or “different,” a screen would appear that directed them to press the spacebar to continue. Their choices were automatically recorded into a text file that was saved to the main lab computer.

Syntactic complexity was controlled so that no complements contained relative clauses to minimize processing load errors due to embedded relative clauses (cf. Epstein et al., 1996; Gibson, 1991). Similar to the AWP task, the ISI was controlled to 500 ms between each sentence in the pairs.<sup>23</sup> In between each sentence pair after subjects made their choice (A for *same* or L for *different*), the interval of time to prepare for the next stimulus was self-paced (i.e., subjects tapped the spacebar to continue to the next item). After they tapped the spacebar, the next pair played aurally, and they again saw the prompt, “Choose ‘A’ for same or ‘L’ for different” on the screen. The keyboard was marked accordingly with “same” and “different” to be sure that there were no mis-typing errors. Again, the directions were written and spoken in Mandarin, and the subjects were not able to go back to any items that had already been presented.

Accuracy rates were calculated by the subjects’ total correct choices of each target verb over the total of verbs in that category to obtain a percentage of accuracy for each item type. The analysis separates *same* from *different* accuracy rates in order to determine if same pairs or different pairs were more accurately perceived.

Following the same design in the AWP task, all of the sentences were recorded in Audacity by a male American English speaker prior to the task to control for uniform volume and rate of delivery. The sentence pairs were presented aurally on headphones. Each sentence pair was played only once. To control the sentence pairs for uniformity, the sentence pairs were generated by first recording a frame sentence and then by splicing in the two different verbs recorded in separate sentences into the original frame sentence. For example, for the *different* items, the following was recorded: (a) *She always leaves a bag on the plane;* and (b) *She always leave a bag on the plane.* Using Audacity, the uninflected verb *leave* was then spliced into a

---

<sup>23</sup> Because there is no agreed upon ISI in the literature for sentence tasks that require grammaticality judgments (cf. Gass & Mackey, 2007) and because a short ISI between the items (i.e., 500ms) will force the participants to focus on the perceptual acoustic and phonetic differences between the items, the ISIs were deliberately kept short.

copy of the first frame sentence so that the sentences were exactly the same length, amplitude, etc. with the only difference being in the verb which was recorded naturally in the second frame sentence to retain the natural acoustic features of the vowel (e.g., duration and pitch) within the target items. And unlike the truncated forms in the AWP task, having to control for truncation was unnecessary because the uninflected verbs were not being truncated but recorded naturally in each sentence context. With careful splicing, the spliced frame sentence had no perceptible co-articulation on either side of spliced verb because there was enough silence between the verb and the surrounding material to avoid overlap at the word boundaries.

This task judged the subjects' ability to perceive the presence (or absence) of target verbal inflections and L2 consonant clusters in the aural processing mode. Thus, similar to the AWP task, this task was designed to reveal differences in aural perception of the inflections due to L1 phonotactic interference, salience effects, and processing load effects that were claimed to cause lower accuracy on the longer, more complex sentences than on the shorter, less complex sentences. Again, similar to the AWP task, if L1 phonotactic interference has a significant general effect on perception of the target codas, an across the board effect on accuracy rates is expected such that any item that violates L1 phonotactics should reveal reduced accuracy in the form of deletion of one or more of the consonants that comprise the target L2 inflection. However, if L1 phonotactic interference does not reveal an across the board effect, but variable accuracy is still observed, it is possible that L1 phonotactic effects reduce accuracy of specific L2 inflections due to the phonetic quality of the coda. In addition, syllabicity effects mediating perception of different allomorphs are expected, resulting in asymmetric accuracy rates with non-syllabic L2 inflections being less accurately perceived than syllabic L2 inflections. Again, as

described in the AWP task, it is expected that *different* items will be less accurate overall than *same* items.

#### 4.2.3 Experiment 2: Oral Production Tasks

Experiment 2 was designed to measure the subjects' ability to produce the target inflections in isolated words, as well as within sentences in order to investigate L1 phonotactic effects and processing load effects on oral production. As in Experiment 1, two tasks were included: a word-based task and a sentence-based task.

##### 4.2.3.1 Oral Word Production Task

The Oral Word Production (OWP) task was administered first. The task included four verb items for each target allomorph under investigation (i.e., [Vd], [t], [d], [Vz], [s], [z]) to mirror the design of the word perception tasks. Four tokens were given for each type. Items appeared at random for every participant, so no participant encountered the same order of items in any task. Nonce items were also included to determine the effects of L1 phonotactic interference and syllabicity on non-lexicalized items in production. In addition, monomorphemic, non-verbal items (e.g., *mitt*, *mid*, *crag*, *crack*) were included in this task to determine if L1 phonotactic interference would have an across the board effect on *any* coda that is illicit in the L1, not just items that contain L2 inflection. These monomorphemic items were divided into 4 categories: 1) *voiced stops* (e.g., *mid*), *unvoiced stops* (e.g., *mitt*), *voiced fricatives* (e.g., *prize*), and *unvoiced fricatives* (e.g., *price*).<sup>24</sup> Non-violating items and copula items were also included as filler/distracters; these “filler” data are not reported, as they are not relevant to the study.

---

<sup>24</sup> Fricatives also included voiced and unvoiced affricates such as *surge* and *pitch*.

For the OWP task, subjects were instructed to hit the spacebar to load each stimulus word. Each single-word stimulus appeared on the computer screen for 500ms, and the subjects were instructed to say the target word that they saw on the screen, which was digitally recorded and later analyzed via Praat. The duration of each stimulus item on the screen was at least two times the length of the longest word in the task spoken by a native speaker (as determined by an informal pilot test). That said, the short duration of the items on the screen was deliberate to induce the subjects to process the target items under processing pressure that mimics natural language production. Additionally, to avoid conflating both aural and oral processing in one task, the target word only appeared visually and was not heard. After each word disappeared from the screen, the subjects were prompted on the screen to hit the spacebar to continue to the next item when they were ready. Again, as in the design of the aural perception tasks, because the focus of the oral tasks was on the accurate production of the target words and not on speed of the overall task, self-pacing between items was deliberate to reduce unnecessary processing load effects caused by memory overload interference between the task items.

If L1 phonotactic interference shows an effect in this task, it would be expected that accuracy of production of the target items including the nonce and monomorphemic items would be significantly reduced such that items that contain illicit codas would be reduced via deletion of one or more of the word-final consonants or schwa epenthesis within the cluster or before or after the word-final coda in the cluster. In fact, there could be syllable repair effects such as those suggested in Peperkamp (2004) and Davidson (2005, 2006a, 2006b) that preserve the syllable-final consonants while producing changes to the cluster itself such as schwa epenthesis. The effect of L1 phonotactic interference could also be modulated by the type of coda, resulting in reduced accuracy on only some of the coda types. In addition, an analysis of the effects of

syllabicity may reveal specific effects on non-syllabic items that may be less accurately produced than syllabic items. In addition, comparisons of accuracy between subjects' perception and production of the same types of codas across the aural word and oral word tasks may reveal a correlation that suggests a relationship between perception and production of L2 inflection following Jia et al. (2006).

#### *4.2.3.2 Oral Sentence Production Task*

The Oral Sentence Production (OSP) task targeted the subjects' ability to produce target inflections within sentences to determine the effects of L1 phonotactics, syllabicity, and increased processing pressure on accuracy. This task also can provide evidence of productivity of the subjects' knowledge of L2 inflections such that a high rate of accuracy on these items could reveal advanced knowledge of L2 morphosyntax. The task contained four items of each inflection under investigation as follows in example (13):

(13) Last year, the athlete \_\_\_\_\_ all the time. (to rest).

Additionally, this task included 24 nonce items following the same design of the other aural and oral tasks. As in the WGS and ASP tasks, items were balanced between long and short sentences with simple and complex subjects, and each verb was followed by a vowel to avoid word-boundary blending interference. Items appeared at random for every participant, so no participant encountered the same order of items in any task. Again, non-violating items and copula items were also included as filler/distracters; these "filler" data are not reported, as they are not relevant to the study.

As in the other tasks, the procedure of this task was explained in written and spoken Mandarin and required the subjects to hit the spacebar to load each stimulus. After hitting the

spacebar, a priming sentence appeared on the screen for four seconds, and then the target “frame” sentence appeared on the screen for four seconds. The priming sentence contained the target verb in an infinitive or uninflected form after a model or auxiliary verb (e.g., *The young athlete will rest before the big game tomorrow.*). Then the target frame sentence automatically appeared on the next screen with the infinitive form of the verb in parentheses (*to rest*). This task was based on a similar production task design utilized by Solt et al. (2004), Klein et al. (2004), and Szupica-Pyrzanowski (2009) in which subjects were given a background prompt sentence that contained the target verb (e.g., *The boy used to walk in the park.*) and then the task provided a target frame sentence which required the subjects to produce (in the case of this dissertation, orally produce) the target verb with the target inflection within an entire sentence such as *Yesterday, the boy \_\_\_\_\_ (to walk) in the park.* However, the task in this study differed slightly from the previous studies such that this researcher only provided the background prompt and target frame sentences visually and did not present them aurally, and the subjects were required to produce the entire target sentence. The aural prompt was removed from this task to avoid conflating the two performance modes within the same task. The subjects were instructed to say the target sentence with the correct verb inserted into the sentence. This 4-second time frame for presentation of the prime and target sentences was based on an informal pilot in which native English speakers read the sentences and were timed. The length of time for a native speaker to say the longest sentence in the task was no more than four seconds, so the time that the sentences appeared on the screen for every sentence including the short sentences was set at a maximum of four seconds. This time pressure was again added to put pressure on their processing resources as would be the case in natural language production with the expectation that this added processing pressure would cause a reduction in accuracy for the longer sentences. However, again because

the focus of the task was for the subjects to produce the entire target sentence, enough time was given to do so. After the subjects finished saying the target sentence, they pressed the spacebar to prompt the next visual stimulus—again, self-pacing between target stimuli was provided to reduce unnecessary processing pressure between items. Participants' responses were digitally recorded and were later analyzed via Praat.

The main purpose of this task was to determine the effects of L1 phonotactic constraints on the oral production of target L2 inflections within sentences. If L1 phonotactic interference shows an across the board effect, high rates of deletion or repair via schwa epenthesis of one or more of the word-final consonants is expected on all of the target items that reflect L1 phonotactic interference in the syllable-final consonant clusters including the nonce items. As suggested in Peperkamp (2004) and Davidson (2005, 2006a, 2006b), L1 syllable effects could induce deletion of the word-final consonants or preservation of the syllable-final consonants via schwa epenthesis to break up the word-final cluster. The subjects' accuracy with regard to the different sentence lengths (i.e., 8-11 vs. 12-18) was also tested to determine if sentence length would have a significant effect on processing load within this task, and as hypothesized, higher accuracy rates on the shorter, less complex sentences than on the longer, more complex sentences are expected. As with the other tasks, if an across the board reduction in accuracy of the target codas due to L1 phonotactic interference is not observed, it is possible that variability of inflection is mediated by the coda type or by varying degrees of syllabicity of inflections, and as claimed, it would be expected to see higher accuracy rates on syllabic inflections vs. less salient, non-syllabic inflections. As with the oral word production task, comparisons of syllabicity between subjects' perception and production of the same types of codas may also

reveal a correlation that suggests a relationship between perception and production of L2 inflection following Jia et al. (2006).

#### 4.2.4 Cross-Modal Design of Target Items

In order to make comparisons across tasks with regard to the syllable-final consonant cluster types in terms of perception and production, the rhyme and codas of all target items were equally matched across tasks for the following: 1) the number of consonants in the final cluster (i.e., -VC, -CC, CCC), and 2) the type of consonants in the cluster and the vowel preceding cluster (e.g., *-olds* as in *fold*s, *scold*s, *hold*s, *mo*ld)s), as well as in the nonce items (e.g., *bazold*s, *ganold*s, *zabold*s, *kavold*s). Although the items were presented pseudo-randomly in every administration of the task (i.e., every subject received a different order of items within each task to avoid item-fatigue effects in which the last items of every task may show higher rates of error), all tasks were presented in the same order for each subject (see task order below in section 4.2.5). Thus, this researcher wanted to avoid lexical-item frequency effects that might have been caused by the increased familiarity of the same item across the tasks—that is, a repeatedly used target verb might become more familiar to the subjects as the tasks progressed causing accuracy on that item to increase due to lexical frequency effects. As such, the target syllable-final consonant clusters under investigation were controlled, while the onsets differed, in order to compare phonotactic, saliency, and performance effects even though the same lexical items were not reused across tasks. In addition, the nonce items provided an additional layer of protection against frequency effects because the nonce items would not be retrievable from the lexicon. As such, the above design was developed to avoid effects of lexical item familiarity, while still providing comparable target codas across the tasks.

#### 4.2.5 General Procedure

The aural tasks were administered first, then the oral tasks, and finally the written grammaticality screener tasks. All directions for each task were presented in written form in Mandarin and then read aloud in Mandarin to the subjects before they began the tasks. The two written grammaticality screener tasks were administered last to avoid increasing subjects' awareness of the target inflections under investigation. The subjects were also given four practice items at the beginning of each task to ensure that they understood the task requirements before beginning the actual task items. After the tasks were completed, the written grammaticality screener task results were to be used to exclude any subjects that did not have an advanced knowledge of the inflections under study. After a review of the results, none of the subjects was excluded due to their near-ceiling accuracy on the written grammaticality screening tasks. However, as stated previously, some subjects were excluded due to lack of completion of tasks or due to poor oral recordings related to unanticipated technical problems with some of the subjects' recording devices.

## Chapter 5: Results

### 5.1 Experiment 1: Aural Word and Sentence Perception Tasks

In Experiment 1, the Aural Word Perception (AWP) and Aural Sentence Perception (ASP) tasks were administered to determine the effect of the subjects' L1 phonotactic constraints, syllabicity of L2 target codas, and processing load effects on perception of target L2 codas. The data were analyzed for subjects' accuracy in identifying target items in an AX *same/different* discrimination paradigm.

#### 5.1.1 Aural Word Task Results – Real and Nonce Items

Table 1 shows the mean number correct and percentage correct for both groups of subjects on the AWP task for *same* real (i.e., lexicalized) word pairs only.

**Table 1: Aural Word Perception – Same Items (Real)**

Coda Type		Lower Proficiency (n=19)		Higher Proficiency (n=21)	
		Mean	SD	Mean	SD
[Vd] (n=2)	Mean	1.95	98%	1.90	95%
	SD	0.23		0.30	
[t] (n=2)	Mean	1.90	95%	1.57	79%
	SD	0.32		0.60	
[d] (n=2)	Mean	1.95	98%	2.00	100%
	SD	0.23		0.00	
[Vz] (n=2)	Mean	2.00	100%	1.90	95%
	SD	0.00		0.30	
[s] (n=2)	Mean	2.00	100%	1.95	98%
	SD	0.00		0.22	
[z] (n=2)	Mean	1.84	92%	1.90	95%
	SD	0.37		0.30	

In order to determine the effects of the two proficiency levels (lower and higher) and six target codas (e.g., [Vd], [t], [d], [Vz], [s], [z]) on subjects' perceptual accuracy on the *same* real word

pairs, a mixed two-way repeated measures ANOVA was conducted, with target codas in real items as within subject factors and proficiency level as the between-subject factor.

With the groups separated, no main effect of coda type,  $F(5, 34)=2.39, p=.058$ , or proficiency level,  $F(1, 38)=2.42, p=.13$ , was found. There was also no significant interaction between proficiency level and coda type for the *same* word pairs,  $F(5, 34)= 1.71, p=.16$ . In fact, with the exception of the 79% on coda [t] for the higher group, the near-ceiling<sup>25</sup> accuracy levels for both groups on this task revealed no across the board effect of L1 phonotactics or any effect of proficiency level on accuracy of the target items. Planned, post-hoc comparisons between voiced, syllabic codas [Vd] and [Vz] and their voiced, non-syllabic counterparts [d] and [z], respectively, revealed no significant differences between the different coda types.

Table 2 shows the mean number correct and percentage correct for both groups of subjects on the AWP task for *different* real word pairs only.

**Table 2: Aural Word Perception – Different Items (Real)**

Coda Type		Lower Proficiency (n=19)		Higher Proficiency (n=21)	
		Mean	SD	Mean	SD
[Vd] (n=2)	Mean	2.00	100%	1.95	98%
	SD	0.00		0.22	
[t] (n=2)	Mean	1.74	87%	1.95	98%
	SD	0.56		0.22	
[d] (n=2)	Mean	1.84	92%	2.00	100%
	SD	0.50		0.00	
[Vz] (n=2)	Mean	2.00	100%	1.90	95%
	SD	0.00		0.44	
[s] (n=2)	Mean	1.95	98%	1.95	98%
	SD	0.23		0.22	
[z] (n=2)	Mean	1.95	98%	2.00	100%
	SD	0.23		0.00	

<sup>25</sup> In many cases, the ANOVAs reported for the word perception and production tasks should be interpreted with caution in those conditions where performance was consistently at ceiling. In those few tables that show results consistently over 90%, the threat to validity is understood. Nevertheless, the results that were at ceiling were not used in any discussion as proof of significant variance.

A mixed two-way repeated measures ANOVA with coda-type as within-subject and proficiency level as between-subject factors revealed the following: No main effect for coda type on perceptual accuracy,  $F(5, 34)=1.8, p=.14$ . There was also no main effect for proficiency level,  $F(1, 38)=.86, p=.36$ , and there was no interaction between proficiency level and coda type,  $F(5, 34)=.70, p=.63$ .

Planned, post hoc comparisons between syllabic and non-syllabic codas revealed no significant difference between the comparable codas (i.e., [Vd]/[d], [t]/[d], [Vz]/[z], [s]/[z]). However, once again, when comparing the means of the different codas, it appears that unvoiced [t] has the lowest accuracy rate in the word perception task.

Table 3 shows the mean number correct and percentage correct for both groups of subjects on the AWP task for the *same* nonce word pairs.

**Table 3: Aural Word Perception – Same Items (Nonce)**

Coda Type		Lower Proficiency (n=19)		Higher Proficiency (n=21)		
			Mean	SD	Mean	SD
[Vd] (n=2)	Mean	1.95	96%	1.90	95%	
	SD	0.23		0.45		
[t] (n=2)	Mean	1.95	96%	1.90	95%	
	SD	0.23		0.31		
[d] (n=2)	Mean	1.84	92%	1.95	96%	
	SD	0.37		0.22		
[Vz] (n=2)	Mean	1.90	95%	2.00	100%	
	SD	0.32		0.00		
[s] (n=2)	Mean	1.90	95%	1.90	95%	
	SD	0.46		0.31		
[z] (n=2)	Mean	1.90	95%	1.90	95%	
	SD	0.46		0.31		

Again, the subjects' accuracy rates are near ceiling similar to the AWP real *same* items and did not show any significant main or interaction effects.

However, in an analysis of the *different* nonce word pairs in the AWP task, a significant effect of coda type was observed. Table 4 shows the mean number correct and percentage correct for both groups of subjects on the AWP task for the *different* nonce word pairs.

**Table 4: Aural Word Perception – Different Items (Nonce)**

Coda Type		Lower Proficiency (n=19)		Higher Proficiency (n=21)	
[Vd] (n=2)	Mean	1.95	96%	1.90	95%
	SD	0.23		0.45	
[t] (n=2)	Mean	2.00	100%	2.00	100%
	SD	0.00		0.31	
[d] (n=2)	Mean	1.58	79%	1.75	88%
	SD	0.61		0.22	
[Vz] (n=2)	Mean	1.95	96%	1.95	96%
	SD	0.23		0.00	
[s] (n=2)	Mean	1.84	92%	1.85	93%
	SD	0.37		0.31	
[z] (n=2)	Mean	1.84	92%	1.90	95%
	SD	0.37		0.31	

A mixed two-way repeated measures ANOVA with coda-type as within-subject and proficiency level as between-subject factors revealed a main effect for coda type on perceptual accuracy,  $F(5, 33)=5.91, p=001$ . However, the main effect for proficiency level did not reach significance,  $F(1, 37)=.49, p=.49$ , and there was also no significant interaction between proficiency level and coda type,  $F(5, 33)=.57, p=.72$ . Again, the accuracy of both proficiency levels on these nonce items were as equally high as on the real word items.

Planned, post hoc comparisons between syllabic ([Vz]/[Vd]) and non-syllabic ([t]/[d]/[s]/[z]) codas on nonce items revealed a significant difference only between the non-syllabic [t] ( $M=2.00, SD=.00$ ) and non-syllabic [d] ( $M=1.66, SD=.85$ ),  $p=.005$  with [t] codas being significantly more accurately perceived than [d] codas. This result was not in accordance with the real *different* items in this task, in which [t] was the least perceptible allomorph. As is shown below in the sentence perception results, the same pattern occurs – that is, in real *different* items, [d] is less perceptible than [t], but in the nonce *different* items, [t] is less perceptible than

[d]. The real and nonce items were equally balanced in terms of coda type [Cct] vs. [CCd] and [Ct] vs. [Cd] across real and nonce items and in terms of sentence length across these different codas, so the effect is not likely a reflection of task design. The potential causes will be discussed in Chapter 6.

No other comparisons between syllabic and non-syllabic items within the nonce items reached significance. Overall, within aural perception at the word level, the two proficiency level groups scored nearly at ceiling on all of the codas in both real and nonce items suggesting that aural perception of the word-final target coda consonant clusters, at least at the word level, do not appear to be affected by L1 phonotactic interference or syllabicity effects across the board.

### 5.1.2 Aural Sentence Task Results – Real and Nonce Items

Table 5 shows the mean number correct and percentage correct for both groups of subjects on the ASP task for *same* real inflected pairs within sentences.

**Table 5: Aural Sentence Perception - Same Items (Real)**

INFL Type		Lower Proficiency (n=19)		Higher Proficiency (n=21)	
[Vd] (n=2)	<i>Mean</i>	1.68	84%	1.57	79%
	<i>SD</i>	0.48		0.60	
[t] (n=2)	<i>Mean</i>	1.68	84%	1.57	79%
	<i>SD</i>	0.58		0.60	
[d] (n=2)	<i>Mean</i>	1.89	95%	1.62	81%
	<i>SD</i>	0.32		0.59	
[Vz] (n=2)	<i>Mean</i>	1.74	87%	1.81	91%
	<i>SD</i>	0.56		0.40	
[s] (n=2)	<i>Mean</i>	1.79	90%	1.71	86%
	<i>SD</i>	0.42		0.64	
[z] (n=2)	<i>Mean</i>	1.37	69%	1.48	75%
	<i>SD</i>	0.76		0.52	

A mixed two-way repeated measures ANOVA with inflection-type as within-subject and proficiency level as between-subject factors revealed the following: No main effect for inflection type on perceptual accuracy,  $F(5, 34)=1.96, p=.11$ . The main effect for proficiency level also did

not reach significance,  $F(1, 38)=.52, p=.48$ . There was also no significant interaction between proficiency level and inflection type,  $F(5, 34)=.76, p=.59$ . Overall, generally high accuracy rates were observed across the board for the *same* items in the sentence task items.

However, planned, post hoc comparisons between syllabic ([Vz]/[Vd]) and non-syllabic ([t]/[d]/[s]/[z]) allomorphs revealed a significant difference between the comparable allomorphs [Vz] ( $M=1.78, SD=.48$ ) and non-syllabic [z] ( $M=1.43, SD=.64$ ),  $p<.05$  with [Vz] allomorphs being significantly more accurately perceived than [z] allomorphs. Conversely, when comparing the means of the syllabic [Vd] and non-syllabic [d], the syllabic [Vd] appeared less perceptible than [d] contra to the hypothesized syllabicity effect, but this difference did not reach significance, so the difference was considered negligible.

Table 6 shows the mean number correct and percentage correct for both groups of subjects on the ASP task for *different* real inflected pairs within sentences. The results reveal an obvious effect of inflection type and proficiency level on perception.

**Table 6: Aural Sentence Perception - Different Items (Real)**

INFL Type		Lower Proficiency (n=19)		Higher Proficiency (n=21)	
		<i>Mean</i>			
[Vd] (n=2)	<i>Mean</i>	1.42	71%	1.62	81%
	<i>SD</i>	0.69		0.50	
[t] (n=2)	<i>Mean</i>	0.32	16%	0.95	48%
	<i>SD</i>	0.48		0.67	
[d] (n=2)	<i>Mean</i>	0.74	37%	1.24	62%
	<i>SD</i>	0.81		0.44	
[Vz] (n=2)	<i>Mean</i>	1.63	82%	1.81	92%
	<i>SD</i>	0.60		0.40	
[s] (n=2)	<i>Mean</i>	1.68	84%	1.71	86%
	<i>SD</i>	0.58		0.56	
[z] (n=2)	<i>Mean</i>	1.16	58%	1.62	81%
	<i>SD</i>	0.69		0.59	

A mixed two-way repeated measures ANOVA with inflection-type as within-subject and proficiency level as between-subject factors revealed a significant difference for the *different*

pairs within sentences. There was a main effect for inflection type on perceptual accuracy,  $F(5, 34)=16.16, p<.001$ . The main effect for proficiency level also reached significance,  $F(1, 38)=12.71, p=.001$  with the higher level proficiency group revealing a much higher level of perception on several of the allomorphs. There was no interaction between proficiency level and inflection type,  $F(5, 34)=1.49, p=.22$ .

Planned, post hoc comparisons between syllabic ([Vz]/[Vd]) and non-syllabic ([t]/[d]/[s]/[z]) allomorphs revealed a significant difference between the comparable allomorphs [Vd] ( $M=1.53, SD=.60$ ) and non-syllabic [d] ( $M=1.00, SD=.68$ ),  $p<.05$  with [Vd] allomorphs being significantly more accurately perceived than [d] allomorphs. Additionally, although no other comparisons reached significance, the mean accuracy of the non-syllabic [z] and especially non-syllabic [t] reveal that these two allomorphs within the *different* items are the least accurate overall.

Table 7 shows the mean number correct and percentage correct for both groups of subjects on the ASP task for nonce *same* target inflection pairs within sentences.

**Table 7: Aural Sentence Perception – Same Items (Nonce)**

INFL Type		Lower Proficiency (n=19)		Higher Proficiency (n=21)	
		Mean		Mean	
[Vd] (n=2)	Mean	1.84	92%	1.80	90%
	SD	0.50		0.41	
[t] (n=2)	Mean	1.68	84%	1.40	70%
	SD	0.67		0.68	
[d] (n=2)	Mean	1.84	92%	1.80	90%
	SD	0.37		0.41	
[Vz] (n=2)	Mean	1.74	87%	1.85	93%
	SD	0.56		0.37	
[s] (n=2)	Mean	2.00	100%	1.75	88%
	SD	0.00		0.55	
[z] (n=2)	Mean	1.74	87%	1.60	80%
	SD	0.45		0.50	

A mixed two-way repeated measures ANOVA with inflection-type as within-subject and proficiency level as between-subject factors revealed no significant main or interaction effects. In fact, there were generally high accuracy rates for the *same* nonce items in the sentence task.

Planned, post hoc comparisons between syllabic ([Vz]/[Vd]) and non-syllabic ([t]/[d]/[s]/[z]) allomorphs also revealed no significant interactions. Again, the real *same* items revealed no across the board effects of L1 phonotactics and syllabicity and showed no significant effects on accuracy of the target L2 inflections or final consonant clusters.

Table 8 shows the mean number correct and percentage correct for both groups of subjects on the ASP task for *different* nonce target items within sentences. There is an obvious reduction in accuracy of some target inflections for both the lower and higher proficiency level groups. Although there does not appear to be an across the board effect of L1 phonotactic interference, overall, the results reveal an obvious effect of inflection type and proficiency level on perception of some specific items types suggesting that L1 phonotactics and salience caused an asymmetrical reduction of accuracy depending on the inflection type.

**Table 8: Aural Sentence Perception –Different Items (Nonce)**

INFL Type		Lower Proficiency (n=19)		Higher Proficiency (n=21)	
		Mean		Mean	
[Vd] (n=2)	Mean	1.63	82%	1.45	73%
	SD	0.50		0.60	
[t] (n=2)	Mean	1.37	69%	1.50	75%
	SD	0.68		0.51	
[d] (n=2)	Mean	0.47	24%	0.90	45%
	SD	0.70		0.72	
[Vz] (n=2)	Mean	1.63	82%	1.65	83%
	SD	0.50		0.49	
[s] (n=2)	Mean	1.42	71%	1.70	85%
	SD	0.77		0.66	
[z] (n=2)	Mean	1.58	79%	1.65	83%
	SD	0.61		0.49	

The subjects performed more poorly on the *different* items (particularly on the non-syllabic stops [t] and [d]) than on the *same* items, revealing that these subjects' perception of the differences in the stops was particularly challenging.

A mixed two-way repeated measures ANOVA with inflection-type as within-subject and proficiency level as between-subject factors revealed a main effect for inflection type on perceptual accuracy,  $F(5, 33)=13.43, p<.001$ . However, the main effect for proficiency level did not reach significance,  $F(1, 37)=2.71, p=.16$ , and there was no significant interaction between proficiency level and inflection type,  $F(5, 34)=.77, p=.57$ .

Planned, post hoc comparisons between syllabic ([Vz]/[Vd]) and non-syllabic ([t]/[d]/[s]/[z]) allomorphs revealed a significant difference in accuracy rates between the comparable allomorphs [Vd] ( $M=1.54, SD=.09$ ) and non-syllabic [d] ( $M=.69, SD=.11$ ),  $p<.001$  with [Vd] allomorphs being significantly more accurately perceived than [d] allomorphs. Additionally, a significant difference was observed between the non-syllabic [t] ( $M=1.43, SD=.09$ ) and non-syllabic [d] ( $M=.69, SD=.11$ ),  $p=.001$  with [t] allomorphs being significantly more accurately perceived than [d] allomorphs. Although no other comparisons reached significance, items that contained the [t] and [d] allomorphs showed a lower accuracy rates than the items that contained the [s] and [z] allomorphs, suggesting that the non-syllabic stops are less acoustically perceptible than the non-syllabic sibilants. However, these comparisons did not reach significance.

Overall, the higher level of acoustic perceptibility is noticeable in the differences between the mean accuracy rates between the sibilants vs. the stops, as the non-syllabic stops reveal a much lower level of accuracy in perception than the non-syllabic sibilants. This effect is likely due to the more acoustically salient nature of sibilants that have higher sonority than other

consonants (cf. Ladefoged, 2006). This difference between accuracy of stops and sibilants is an important result and will be discussed in more detail in Chapter 6.

### 5.1.3 Sentence Length Effects in Aural Sentence Perception

Additionally, the effect of sentence length on aural perception was determined by comparing the subjects' accuracy rates in short vs. long sentences. Table 9 shows the mean number correct and percentage correct on the ASP task in terms of sentence length for real target items.

**Table 9: Aural Sentence Perception - Sentence Length Effects (Lexical)**

Sentence Length	Proficiency Level	Mean	SD	% Correct
Long (n=12)	Lower Group (n=19)	8.32	1.49	69%
	Higher Group (n=21)	8.62	1.60	72%
	Total (n=40)	8.48	1.54	71%
Short (n=12)	Lower Group (n=19)	8.79	1.84	73%
	Higher Group (n=21)	10.10	1.34	84%
	Total (n=40)	9.48	1.71	79%

A mixed two-way repeated measures ANOVA with sentence length (short/long) as within-subject and proficiency level as between-subject factors was conducted. A significant main effect for sentence length,  $F(1, 38)=11.26, p<.001$  was found, as well as a significant main effect for proficiency level,  $F(1, 38)=3.96, p<.05$ . No significant interaction was found between sentence length and proficiency level,  $F(1, 38)=2.98, p=.09$ .

The results of the nonce items in the ASP task also showed an effect of sentence length on accuracy of L2 inflection. Table 10 shows the mean number correct and percentage correct on the ASP task in terms of sentence length for nonce target items.

**Table 10: Aural Sentence Perception - Sentence Length Effects (Nonce)**

Sentence Length	Proficiency Level	Mean	SD	% Correct
Long (n=12)	Lower Group (n=19)	8.74	1.19	73%
	Higher Group (n=21)	8.25	1.74	69%
	Total (n=40)	8.49	1.50	71%
Short (n=12)	Lower Group (n=19)	10.21	1.13	85%
	Higher Group (n=21)	10.80	1.36	90%
	Total (n=40)	10.51	1.27	88%

A mixed two-way repeated measures ANOVA with sentence length (short/long) as the within-subject factor and proficiency level as between-subject factor was conducted. A main effect for sentence length was found,  $F(1, 37)=58.55, p<.001$ , revealing that processing pressure due to longer, more complex sentences caused lower accuracy of L2 inflection on nonce items. However, no main effect for proficiency level was found,  $F(1, 37)=.02, p=.89$ . There was a significant interaction between sentence length and proficiency level,  $F(1, 38)=4.19, p<.05$ , revealing that the groups did not show the same effects of sentence length although this effect was marginal. Nevertheless, sentence length clearly had an effect on accuracy of perception of the target nonce items.

Overall, in Experiment 1, there does not appear to be any across the board effects of L1 phonotactic interference in terms of aural perception of the target L2 inflections and word-final consonant clusters. However, L1 phonotactic effects were evident with regard to some target codas at the word level, but particularly with some allomorphs at the sentence level, specifically depending on the phonetic nature of the items (i.e., stops vs. sibilants) as will be discussed in Chapter 6. Accuracy rates were also affected by processing pressure, as evident in the longer, more complex sentences, and reduced accuracy is clearly present in both the real items, as well as in the nonce items. There were also effects of syllabicity in some of the planned post-hoc comparisons; however, these effects did not affect accuracy across the board. The significance of these results will be discussed in more detail in Chapter 6.

## 5.2 Experiment 2: Oral Word and Sentence Production Tasks

In Experiment 2, the Oral Word Production (OWP) and Oral Sentence Production (OSP) tasks were administered to determine the effect of the subjects' L1 phonotactic constraints and processing load on production of target L2 codas. Planned, post-hoc analyses to examine syllabicity effects were also conducted. The data were analyzed in two ways: 1) the accuracy of suppliance of the target codas was analyzed; and 2) the degree to which the subjects used coda repair strategies such as schwa epenthesis to break up word-final consonant clusters was also analyzed. In addition, devoicing of word-final target codas as a repair strategy was analyzed as well. The first analysis below examines deletion of target codas (i.e., accuracy of suppliance). Each deletion of a target coda was determined via an item-by-item spectrographic analysis using Praat to ensure a systematic analysis of syllable-final coda cluster production including an accounting of the total consonants produced, as well as an analysis of the types of coda consonant cluster repair (i.e., schwa epenthesis and devoicing) that was produced by each subject for each target item.

### 5.2.1 Analysis of Suppliance of Real Items in the Oral Word Task

Table 11 shows the mean correct accuracy scores of suppliance of the target consonant clusters in real items for both groups of subjects on the OWP task. First, the effects of L1 phonotactic constraints were examined by determining the overall suppliance of the target codas. As is shown, the subjects had few omissions overall in the single word task and produced the target syllable-final clusters at or near ceiling with few deletions to any part of the coda.

**Table 11: Oral Word Production - Target Coda Suppliance (Real)**

Coda Type		Lower Proficiency (n=19)		Higher Proficiency (n=21)	
[Vd] (n=4)	<i>Mean</i>	3.84	96%	3.95	99%
	<i>SD</i>	0.37		0.22	
[t] (n=4)	<i>Mean</i>	3.95	99%	3.90	98%
	<i>SD</i>	0.23		0.31	
[d] (n=4)	<i>Mean</i>	3.95	99%	3.95	99%
	<i>SD</i>	0.23		0.22	
[Vz] (n=4)	<i>Mean</i>	4.00	100%	3.75	94%
	<i>SD</i>	0.00		0.72	
[s] (n=4)	<i>Mean</i>	3.89	97%	3.95	99%
	<i>SD</i>	0.32		0.22	
[z] (n=4)	<i>Mean</i>	3.68	92%	3.85	96%
	<i>SD</i>	0.75		0.37	

A mixed two-way repeated measures ANOVA with six coda-types as within-subject and two proficiency levels as between-subject factors revealed the following: There was no main effect of proficiency level,  $F(1, 37)=.01, p=.92$  or coda type,  $F(5, 33)=.70, p=.63$  on accuracy. There was also no interaction between proficiency level and coda type,  $F(5, 33)=.92, p=.48$ , suggesting that at the word-level for these subjects, L1 phonotactic constraints did not appear to have an effect on deletion of the target codas. Planned, post-hoc analyses were conducted regarding the effects of syllabicity, but again, no significant difference between the means was observed.

### 5.2.2 Analysis of Suppliance of Nonce Items in the Oral Word Task

Table 12 shows the mean correct accuracy scores of suppliance of the target consonant clusters in nonce items in the OWP task. Overall, the production of word-final target codas is nearly the same as production of the same target codas in real items, although the effects of coda type do reveal significant effects unlike the analysis of real items in the OWP task.

**Table 12: Oral Word Production - Target Coda Suppliance (Nonce)**

Coda Type		Lower Proficiency (n=19)		Higher Proficiency (n=21)	
[Vd] (n=4)	<i>Mean</i>	3.84	96%	3.95	99%
	<i>SD</i>	0.37		0.22	
[t] (n=4)	<i>Mean</i>	4.00	100%	3.80	95%
	<i>SD</i>	0.00		0.41	
[d] (n=4)	<i>Mean</i>	3.84	96%	3.90	98%
	<i>SD</i>	0.50		0.45	
[Vz] (n=4)	<i>Mean</i>	3.42	86%	3.30	83%
	<i>SD</i>	0.77		1.03	
[s] (n=4)	<i>Mean</i>	4.00	100%	4.00	100%
	<i>SD</i>	0.00		0.00	
[z] (n=4)	<i>Mean</i>	3.52	88%	3.55	89%
	<i>SD</i>	0.61		0.69	

A mixed two-way repeated measures ANOVA with six coda-types as within-subject and two proficiency levels as between-subject factors revealed a main effect of coda type on accuracy of production,  $F(5, 33)=10.97, p<.001$ , but no main effect of proficiency level on accuracy,  $F(1, 37)=.11, p=.74$ . There were also no interaction between proficiency level and coda type,  $F(5, 33)=1.90, p=.34$ . Planned, post-hoc analyses were conducted regarding the effects of syllabicity, but because of the obvious high suppliance rates, no significant differences between the means were observed. Although a significant effect of coda type on suppliance of nonce word items was observed, the high rates of suppliance (sometimes at ceiling) of the target codas on nonce items also do not support an L1 phonotactic interference account with regard to suppliance of target codas.

### 5.2.3 Analysis of Suppliance of Monomorphemic Items in the Oral Word Task

An analysis of monomorphemic items within the OWP task also revealed similar patterns of target coda suppliance across the categories. The monomorphemic items included four types of word-final codas that were designed to mirror the inflections under study. The four types were

as follows: 1) *voiced stops* (e.g., *mid*), *unvoiced stops* (e.g., *mitt*), *voiced fricatives* (e.g., *prize*), and *unvoiced fricatives* (e.g., *price*). Table 13 shows the suppliance of word-final codas in the monomorphemic items.

**Table 13: Oral Word Production – Monomorphemic Suppliance**

Coda Type		Lower Proficiency (n=19)		Higher Proficiency (n=21)	
		Unvoiced Stops (n=6)	<i>Mean</i>	6.00	100%
	<i>SD</i>	0.00		0.00	
Voiced Stops (n=6)	<i>Mean</i>	5.95	99%	6.00	100%
	<i>SD</i>	0.23		0.00	
Unvoiced Fricatives (n=6)	<i>Mean</i>	6.00	100%	6.00	100%
	<i>SD</i>	0.00		0.00	
Voiced Fricatives (n=6)	<i>Mean</i>	5.95	99%	6.00	100%
		0.23		0.00	

A mixed two-way repeated measures ANOVA revealed no significant main or interaction effects. The high rate of suppliance of word-final target codas in monomorphemic items mirrors the high accuracy of production of the target codas in the inflected items within the oral word task, and thus, there appears to be no across the board effects of L1 phonotactics on suppliance of codas in the oral word production task. However, as will be shown below in Section 5.2.3, there are definitive systematic repairs on both the oral word real and monomorphemic items that suggest a clear effect of L1 phonotactic interference.

#### 5.2.4 Analysis of Suppliance of Real Items in the Oral Sentence Task

Table 14 below shows the mean rate and percentage of suppliance for both groups of subjects on the OSP task. Rates of suppliance were lowest on the syllabic [Vz] items, but overall, the suppliance rates were still quite high suggesting that there is a lack of evidence of across the board L1 phonotactic constraints on oral production. In addition, these results provide further support for the claim that these subjects have grammatical competence for L2 inflection as

indicated by the high accuracy rates especially in the past regular allomorphs. The fairly high rate of deletion of the [Vz] allomorph was attributed

**Table 14: Oral Sentence Production - INFL Suppliance (Real)**

INFL Type		Lower Proficiency (n=19)		Higher Proficiency (n=21)	
[Vd] (n=4)	<i>Mean</i>	3.68	92%	3.60	90%
	<i>SD</i>	0.58		0.60	
[t] (n=4)	<i>Mean</i>	3.74	94%	3.70	93%
	<i>SD</i>	0.56		0.57	
[d] (n=4)	<i>Mean</i>	3.84	96%	3.75	94%
	<i>SD</i>	0.37		0.44	
[Vz] (n=4)	<i>Mean</i>	2.79	70%	2.80	70%
	<i>SD</i>	1.03		1.06	
[s] (n=4)	<i>Mean</i>	3.32	83%	3.45	86%
	<i>SD</i>	1.11		0.69	
[z] (n=4)	<i>Mean</i>	3.37	84%	3.20	80%
	<i>SD</i>	0.60		1.11	

A mixed two-way repeated measures ANOVA with inflection-type as within-subject and proficiency level as between-subject factors revealed the following: There was no main effect of proficiency level on suppliance rates,  $F(1, 37)=.06, p=.80$ . However, there was a main effect of inflection type on suppliance rates,  $F(5, 33)=7.57, p<.001$ . There was no interaction between proficiency level and inflection type,  $F(5, 33)=.21, p=.95$ . The 3<sup>rd</sup> singular allomorphs were produced with significantly less accuracy than the past regular allomorphs in contrast to the patterns observed in the ASP task in which non-syllabic past allomorphs were less accurate. This contrast will be discussed later.

Because there was a main effect of inflection type, a planned, post-hoc comparison of syllabic and non-syllabic and voiced and unvoiced inflections was conducted to determine if acoustic perceptibility had an effect on production of target inflections, and this analysis revealed only a significant difference between syllabic [Vz] ( $M=2.79, SD=.17$ ) and non-syllabic [z] ( $M=3.28, SD=.14$ ),  $p<.05$  with [Vz] allomorphs being significantly less accurately produced than

[z] allomorphs. This result runs counter to the hypothesized effect in which performance with syllabic allomorphs was expected to be more accurate than performance with non-syllabic allomorphs.

### 5.2.5 Sentence Length Effects on Suppliance of L2 Inflection in Real Verbs

An analysis of the effect of sentence length on accuracy rates in the OSP task was also conducted to determine if processing pressure effects due to increased sentence length resulted in higher rates of deletion of target inflections. Table 15 shows the means and percentage of suppliance rates of real items on the OSP task in terms of sentence length effects.

**Table 15: Oral Sentence Production - Sentence Length Effects (Real)**

<b>Sentence Length</b>	<b>Proficiency Level</b>	<b>Mean</b>	<b>SD</b>	<b>% Correct</b>
Long (n=12)	Lower Group (n=19)	10.26	1.62	86%
	Higher Group (n=21)	10.25	1.29	85%
	Total (n=40)	10.25	1.44	85%
Short (n=12)	Lower Group (n=19)	10.47	1.54	87%
	Higher Group (n=21)	10.25	1.94	85%
	Total (n=40)	10.35	1.73	86%

A mixed two-way repeated measures ANOVA with sentence length (short/long) as within-subject and proficiency level as between-subject factors revealed the following: There was no main effect of proficiency level on deletion of inflections,  $F(1, 37)=.06, p=.80$  or of sentence length on deletion of inflections,  $F(1, 37)=.23, p<.64$ , indicating that longer sentences containing real verbs did not result in significant deletion of target inflections. There was also no significant interaction between proficiency level and sentence length in terms of the deletion of inflections,  $F(1, 37)=.23, p=.64$ . These results indicate that in oral sentence production, neither sentence length nor proficiency level had a significant effect on suppliance of L2 inflections in real verbs.

### 5.2.5 Analysis of Suppliance of Nonce Items in the Oral Sentence Task

The effect of the subjects' L1 phonotactic constraints on L2 production of target codas within sentences with nonce items was also analyzed. Table 16 shows the accuracy of L2 inflection suppliance in nonce items on the OSP task.

**Table 16: Oral Sentence Production - INFL Suppliance (Nonce)**

INFL Type		Lower Proficiency (n=19)		Higher Proficiency (n=21)	
			Mean		
[Vd] (n=4)	Mean	3.42	86%	3.60	90%
	SD	0.61		0.60	
[t] (n=4)	Mean	3.68	92%	3.95	99%
	SD	0.48		0.22	
[d] (n=4)	Mean	3.84	96%	3.70	93%
	SD	0.37		0.57	
[Vz] (n=4)	Mean	3.00	75%	2.85	71%
	SD	1.25		1.04	
[s] (n=4)	Mean	3.42	86%	3.45	86%
	SD	0.84		0.83	
[z] (n=4)	Mean	3.32	83%	3.25	82%
	SD	1.00		0.91	

Rates of suppliance were the lowest on the syllabic [Vz] items, but overall, the suppliance rates were still quite high suggesting that there is a lack of evidence of across the board L1 phonotactic interference on oral production. Additionally, the high rates of suppliance reveal that productivity of L2 inflection for these subjects is advanced and that any reduction of suppliance is mediated by the inflection type rather than by an across the board representational deficit.

A mixed two-way repeated measures ANOVA with inflection-type as within-subject and proficiency level as between-subject factors revealed a main effect of inflection type on suppliance rates,  $F(5, 33)=6.44, p<.001$ , but there was no main effect of proficiency level on suppliance rates,  $F(1, 37)=.01, p=.91$ . There were also no interaction between proficiency level and inflection type,  $F(5, 33)=1.29, p=.29$ . Only the 3<sup>rd</sup> singular allomorph [Vz] was produced

with significantly less accuracy than the other allomorphs, mirroring the result of the lower suppliance of [Vz] in the real sentence production items.

Planned, post-hoc comparisons of syllabic and non-syllabic and voiced and unvoiced allomorphs revealed no significant differences between syllabic and non-syllabic allomorphs, and just as the results in the real sentence production items showed, syllabicity again had no effect on suppliance of the target inflections.

### 5.2.6 Sentence Length Effects on Suppliance of L2 Inflection in Nonce Verbs

Table 17 shows the effects of sentence length on the production of the target inflections in sentences containing nonce items.

**Table 17: Oral Sentence Production - Sentence Length Effects (Nonce)**

Sentence Length	Proficiency Level	Mean	SD	% Correct
Long (n=12)	Lower Group (n=19)	9.32	1.60	78%
	Higher Group (n=21)	9.85	1.66	82%
	Total (n=40)	9.59	1.63	80%
Short (n=12)	Lower Group (n=19)	11.36	1.89	87%
	Higher Group (n=21)	10.95	1.88	84%
	Total (n=40)	11.15	1.87	86%

A mixed two-way repeated measures ANOVA with sentence length (short/long) as within-subject and proficiency level as between-subject factors revealed a main effect of sentence length on deletion of L2 inflections,  $F(1, 37)=30.18, p<.001$ , indicating that longer sentences containing nonce verbs resulted in significant deletion of target inflections in support of the processing pressure hypothesis for production. There was no main effect of proficiency level,  $F(1, 37)=.01, p=.91$ , and there was also no interaction between proficiency level and sentence length in terms of the deletion of inflections,  $F(1, 37)=2.76, p=.11$ . Nonce items appear to require more processing resources leading to reduced accuracy because they cannot be recalled from the lexicon as would be the case with the real target verbs.

As will be discussed in Chapter 6, because production of target L2 inflections was near ceiling with regard to the past regular *-ed* inflectional allomorphs and at an advanced level with regard to 3<sup>rd</sup> singular inflectional allomorphs, this result provides further evidence against representational deficit accounts. Furthermore, although there was no effect of sentence length on suppliance of target inflections with regard to the real items, there was a main effect of sentence length within nonce target items, suggesting that processing pressure did have an effect on performance, particularly in items that were not lexicalized. As will be discussed in Chapter 6, nonce items appear to have caused a higher degree of processing pressure because lexical recall is unavailable for these items, which is suggested by this researcher to increase processing pressure, thus reducing the available resources to successfully parse the nonce items. In fact, in the analysis of the nonce target codas, there were significance main effects of inflection type within both the OWP and OSP tasks and with regard to sentence length on the nonce items, which suggest a greater processing effect. Syllabicity did not show significant effects across the target inflections in oral production.

### 5.3 Experiment 2: Analysis of Repair Strategies in the Oral Tasks

Although suppliance of real and nonce target items within oral production was found to be advanced, the subjects produced the target inflections using various repair strategies that had significant effects on production of some target codas in both the real and nonce target items. It was expected that items that contain illicit codas would be reduced either via deletion of one or more of the word-final consonants or via schwa epenthesis within the cluster or before or after the word-final coda in the cluster. In fact, for these subjects, there appears to be more syllable repair effects such as those suggested in Peperkamp (2004) and Davidson (2005, 2006a, 2006b)

that preserve the syllable-final consonants while producing changes to the cluster itself such as schwa epenthesis. These repairs were also asymmetrically applied across the different types of allomorphs, suggesting that these repairs are not a result of underlying deficits, but that they were modulated by the different acoustic factors of the allomorphs. For example, there were more repairs on coda clusters that contained word-final [t] or [d] when compared to [s] and [z]. In the case of these subjects, syllable repair was exhibited by schwa epenthesis before and after the word-final consonants such as in the following examples: “caused” [kɔz.də] or [kɔ.zə.də]. Any repair of the word-final cluster via schwa epenthesis was considered an error, as it exhibited preservation of the allomorph, but with an accompanying re-syllabification of the cluster to better reflect the phonotactics of the L1.

As such, the effect of L1 phonotactic interference appears to be modulated by the type of coda, resulting in reduced accuracy on only some of the coda types. In addition, an analysis of the effects of syllabicity reveals specific effects on non-syllabic items that were less accurately produced than syllabic items. Finally, comparisons of accuracy between subjects’ perception and production of the same types of codas across the aural word and oral word tasks reveal a correlation that suggests a relationship between perception and production of L2 inflection following Jia et al. (2006). The subjects employed syllable-coda repair strategies in oral production such as the schwa epenthesis observed by Davidson (2005, 2006a, 2006b) before and after the target inflections. Additionally, a significant effect on word-final voiced consonants was the devoicing of the final segments, a result that is consistent with others (cf. Anderson, 1983; Broselow, 2004; Broselow, Chen, & Wang, 1998). Devoicing of word-final consonants were also sometimes combined with epenthesis repairs such as in the following example: “bleeds”

[bliy:t.sə]. In these cases, these items were counted as a dual repair and separated from the individual schwa epenthesis and individual devoicing repairs.

Overall, the subjects produced final codas/inflections with three separate variations: 1) final coda devoicing, 2) schwa epenthesis on either side of the word-final segment that resulted in simplification of the final coda consonant cluster into more simple syllables of CV, and 3) a combination of devoicing and schwa epenthesis. The following analysis reveals these systematic repairs across the real, nonce, and monomorphemic items.

### 5.3.1 Devoicing of Real Items in the Oral Word Task

Table 18 shows the percentage of devoicing of voiced codas within the OWP task real items with the proficiency levels separated in an analysis of a subset of the subjects (n=24). The subjects for this subset were chosen based on the quality of their recorded oral samples. Those subjects who had lower quality recordings were excluded in order to ensure a high quality sample for the more detailed Praat analysis. The subset resulted in 12 lower proficiency and 12 higher proficiency subjects. The [t] and [s] inflected items, as well as any individual items in which the target codas were deleted, were excluded from the devoicing numbers for obvious reasons.

**Table 18: Oral Word Production - Target Coda Devoicing (Real)**

<b>Proficiency Level</b>	<b>[Vd]</b>	<b>[d]</b>	<b>[Vz]</b>	<b>[z]</b>
Lower Group (n=12)	69%	47%	90%	75%
Higher Group (n=12)	68%	58%	64%	63%

A mixed two-way repeated measures ANOVA with coda-type as within-subject and proficiency level as between-subject factors revealed the following: There was a main effect of coda type on devoicing rates,  $F(4, 19)=60.75, p<.001$ . However, there was no main effect of proficiency level on devoicing rates,  $F(1, 22)=0.60, p=.45$ . There was, however, an interaction

effect between proficiency level and coda type,  $F(4, 19)=3.41, p<.05$ , indicating that final coda devoicing was a significant repair strategy for these subjects depending on the coda type regardless of proficiency level.

Planned, post-hoc analyses comparing syllabic and non-syllabic codas revealed no specific significant difference between syllabic [Vz] and non-syllabic [z] or between syllabic [Vd] and non-syllabic [d] codas suggesting that relative syllabicity of target codas had no effect on devoicing rates. Word-final consonant devoicing has been attested on several occasions within the literature (e.g., Anderson, 1983; Broselow, 2004; Broselow, Chen, & Wang, 1998, among others). Its relative frequency will be examined by this researcher later with regard to the perception and production of target codas to make claims about the larger connection between perception and production. This connection will be discussed further in section 6.3 below.

### 5.3.2 Devoicing of Nonce Items in the Oral Word Task

A similarly high level of devoicing was observed in the nonce items in the OWP task.

Table 19 shows the percentage of devoicing of voiced codas within the OWP task nonce items.

**Table 19: Oral Word Production - Target Coda Devoicing (Nonce)**

<b>Proficiency Level</b>	<b>[Vd]</b>	<b>[d]</b>	<b>[Vz]</b>	<b>[z]</b>
Lower Group (n=12)	61%	31%	79%	63%
Higher Group (n=12)	61%	42%	63%	52%

A mixed two-way repeated measures ANOVA with coda-type as within-subject and proficiency level as between-subject factors revealed the following: There was a main effect of coda type on devoicing rates,  $F(3, 20)=4.56, p=.01$ . However, there was no main effect of proficiency level on devoicing rates,  $F(1, 22)=0.24, p=.63$ . There was also no interaction between proficiency level and coda type,  $F(3, 20)=0.70, p=.57$ . Final coda devoicing was a significant factor in repair strategies in nonce items revealing the devoicing was a productive repair for these subjects.

Planned, post-hoc analyses comparing syllabic and non-syllabic codas revealed no specific significant difference between syllabic [Vz] and non-syllabic [z] or between syllabic [Vd] and non-syllabic [d] codas suggesting that relative syllabicity of target codas had no effect on devoicing rates.

### 5.3.3 Schwa Epenthesis of Real Items in the Oral Word Task

Table 20 reveals the percentages of coda cluster repair via schwa epenthesis before or after the target coda on the OWP task. In this analysis, items in which the target codas were deleted were excluded from the analysis, but the unvoiced codas are now included.

**Table 20: Oral Word Production - Target Coda Schwa Epenthesis (Real)**

<b>Proficiency Level</b>	<b>[Vd]</b>	<b>[t]</b>	<b>[d]</b>	<b>[Vz]</b>	<b>[s]</b>	<b>[z]</b>
Lower Group (n=12)	9%	26%	9%	0%	4%	11%
Higher Group (n=12)	6%	30%	15%	0%	0%	4%

A mixed two-way repeated measures ANOVA with coda-type as within-subject and proficiency level as between-subject factors revealed the following: There was a main effect of coda type on epenthesis rates,  $F(5, 18)=3.14, p<.05$ . However, there was no main effect of proficiency level on epenthesis rates,  $F(1, 22)=0.02, p=.88$ . There was also no interaction between proficiency level and coda type,  $F(5, 18)=0.35, p=.87$ . Repair of the coda cluster via schwa epenthesis at least for some of the target codas was a significant factor within the types of repair strategies for these subjects, indicating that although L1 phonotactics did not show an across the board effect of deletion of target codas, L1 phonotactic interference appeared to target specific coda types depending on their phonetic qualities.

Planned, post-hoc analyses comparing syllabic and non-syllabic codas revealed no specific significant difference between syllabic [Vz] and non-syllabic [z] or between syllabic

[Vd] and non-syllabic [d] codas. However, the subjects showed an obvious tendency to use schwa epenthesis repair on codas that contained the past allomorph [t] at a greater rate than the other coda types.

### 5.3.4 Schwa Epenthesis of Nonce Items in the Oral Word Task

A similar pattern of schwa epenthesis was observed in the nonce items in the OWP task. Table 21 shows the percentage of schwa epenthesis of target codas within the OWP task nonce items.

**Table 21: Oral Word Production - Target Coda Schwa Epenthesis (Nonce)**

<b>Proficiency Level</b>	<b>[Vd]</b>	<b>[t]</b>	<b>[d]</b>	<b>[Vz]</b>	<b>[s]</b>	<b>[z]</b>
Lower Group (n=12)	4%	27%	13%	0%	6%	2%
Higher Group (n=12)	8%	25%	13%	0%	0%	0%

A mixed two-way repeated measures ANOVA with coda-type as within-subject and proficiency level as between-subject factors revealed the following: There was a main effect of coda type on epenthesis rates,  $F(5, 18)=4.72, p=.01$ . However, there was no main effect of proficiency level on epenthesis rates,  $F(1, 22)=0.08, p=.78$ . There was also no interaction between proficiency level and coda type,  $F(5, 18)=0.58, p=.72$ . Final repair of the coda cluster via schwa epenthesis in the nonce items at least for some of the target codas was a significant factor within the types of repair strategies for these subjects, and the fact that these repairs are more productive on specific codas reveals that these repairs are mediated by specific phonetic qualities of the individual codas rather than an across the board effect.

As in the real items, there is a similar pattern to epenthesize schwa in the codas that contain stops compared to the codas that contain sibilants. And although planned, post-hoc analyses comparing syllabic and non-syllabic codas revealed no specific significant difference

between syllabic [Vz] and non-syllabic [z] or between syllabic [Vd] and non-syllabic [d] codas, the subjects showed an obvious tendency to use schwa epenthesis repair on codas that contained the past allomorph [t] and [d] at a greater rate than the other codas types. This pattern to epenthesize schwa at a higher rate (i.e., reducing accuracy rates of the word-final clusters in production) with regard to the [t] and [d] allomorphs is reminiscent of the reduced rates of perception of the [t] and [d] allomorphs in target items in Experiment 1. This pattern is examined further in correlations within Section 5.6.

### 5.3.5 Devoicing Plus Schwa Epenthesis of Real Items in the Oral Word Task

Table 22 reveals the percentage of items in which the target codas was both devoiced and the coda cluster was repaired via schwa epenthesis before or after the target codas. Again, items in which the codas were omitted and in which the codas were unvoiced were excluded from the analysis.

**Table 22: Oral Word Production - Target Coda Devoicing + Schwa Epenthesis (Real)**

Proficiency Level	[Vd]	[d]	[Vz]	[z]
Lower Group (n=12)	17%	43%	4%	15%
Higher Group (n=12)	15%	25%	0%	28%

A mixed two-way repeated measures ANOVA with coda-type as within-subject and proficiency level as between-subject factors revealed the following: There was a main effect of coda type on devoicing plus epenthesis repair rates,  $F(4, 19)=6.34, p<.05$ . However, there was no main effect of proficiency level on devoicing plus epenthesis repair rates,  $F(1, 22)=0.15, p=.70$ , and there was no interaction between proficiency level and coda type,  $F(4, 19)=1.09, p=.39$ . These results reveal that final coda devoicing plus epenthesis repair of the coda cluster was a significant repair strategy for these subjects.

Planned, post-hoc analyses comparing syllabic and non-syllabic codas revealed no specific significant difference between syllabic [Vz] and non-syllabic [z] or between syllabic [Vd] and non-syllabic [d] codas, although the rate of this combined repair strategy was obviously higher for the non-syllabic voiced [d] codas for these subjects.

### 5.3.6 Devoicing Plus Schwa Epenthesis of Nonce Items in the Oral Word Task

A similar pattern of schwa epenthesis plus devoicing was observed in the nonce items in the OWP task. Table 23 shows the percentage of schwa epenthesis plus devoicing of voiced codas within the OWP task nonce items.

**Table 23: Oral Word Production - Target Coda Devoicing + Schwa Epenthesis (Nonce)**

<b>Proficiency Level</b>	<b>[Vd]</b>	<b>[d]</b>	<b>[Vz]</b>	<b>[z]</b>
Lower Group (n=12)	21%	45%	0%	27%
Higher Group (n=12)	17%	46%	2%	31%

A mixed two-way repeated measures ANOVA with coda-type as within-subject and proficiency level as between-subject factors revealed the following: There was a main effect of coda type on devoicing plus epenthesis repair rates,  $F(3, 20)=18.59, p<.001$ . However, there was no main effect of proficiency level on devoicing plus epenthesis repair rates,  $F(1, 22)=0.02, p=.880$ . There was also no interaction between proficiency level and coda type,  $F(3, 20)=0.15, p=.93$ . Final coda devoicing plus epenthesis repair of the coda cluster was a significant repair strategy for these subjects in nonce items, which again reveals that this repair was productive to the target codas and was mediated by the coda type.

Planned, post-hoc analyses comparing syllabic and non-syllabic codas revealed a significant difference between [Vz] ( $M=2.79, SD=.17$ ) and non-syllabic [z] ( $M=3.28, SD=.14$ ),  $p<.05$  with [Vz] codas being significantly more accurately produced than [z] codas and between [Vd] ( $M=2.79, SD=.17$ ) and non-syllabic [d] ( $M=3.28, SD=.14$ ),  $p<.05$  with [Vd] codas being

significantly more accurately produced than [d] codas, indicating that syllabicity of the final coda had an effect on the degree of repair that occurred.

### 5.3.7 Schwa Epenthesis Totals for Real and Nonce Items in the Oral Word Task

Table 24 combines the total schwa epenthesis results (i.e., schwa epenthesis and schwa epenthesis+devoicing) of the real items in the OWP task. The combined results reveal that the percentage of schwa epenthesis is higher overall on the codas that contain the [Vd], [t], [d], and [z] allomorphs, suggesting that the stops are overall more prone to syllable repair. However, rates are still below chance for the most part, and there is a lack of across the board re-syllabification effects, especially with regard to codas that contain [Vz] and [s], indicating that L1 phonotactic interference did not exhibit an affect on all coda types for these subjects.

**Table 24: Oral Word Production - Target Coda Schwa Epenthesis Totals (Real)**

Proficiency Level	[Vd]	[t]	[d]	[Vz]	[s]	[z]
Lower Group (n=12)	26%	26%	52%	4%	4%	26%
Higher Group (n=12)	21%	30%	40%	0%	0%	32%

Table 25 combines the total schwa epenthesis results (i.e., schwa epenthesis and schwa epenthesis+devoicing) for the nonce items in the OWP task.

**Table 25: Oral Word Production - Target Coda Schwa Epenthesis Totals (Nonce)**

Proficiency Level	[Vd]	[t]	[d]	[Vz]	[s]	[z]
Lower Group (n=12)	25%	27%	58%	0%	6%	29%
Higher Group (n=12)	25%	25%	59%	2%	0%	31%

Just as in the real items, the combined results of the nonce schwa epenthesis totals reveal that the percentage of schwa epenthesis is higher overall on the codas that mirror the [Vd], [t], [d], and [z] allomorphs. However, because rates are mostly below chance (except for [d]), and

there is a lack of across the board re-syllabification effects especially with regard to codas that contain [Vz] and [s], this result does not support an L1 phonotactic interference effect for all coda types. That said, the rates of schwa epenthesis within the OWP task reveal interesting and common patterns that reflect an obvious coda-specific effect across both the real and nonce items. As will be elaborated further in Chapter 6, the repairs appear to be modulated by the phonetic quality of the codas.

#### 5.4 Experiment 2: Analysis of Repair Strategies in the Oral Sentence Task

An analysis of the results from same subset of subjects on the Oral Sentence Production (OSP) task also revealed similar repair patterns of devoicing and epenthesis. The subjects for this subset analysis were again chosen based on the quality of their recorded oral samples. Those subjects who had lower quality recordings were excluded in order to ensure a high quality sample for the more specific Praat analysis. The subset groups again resulted in the same 12 lower proficiency and 12 higher proficiency subjects.

##### 5.4.1 Devoicing of Real Items in the Oral Sentence Task

Table 26 shows the percentage of inflections in real items that were devoiced within the OSP task with the proficiency levels separated. The [t] and [s] inflected items as well as any individual items in which the inflection was deleted were again excluded from the devoicing numbers for obvious reasons.

**Table 26: Oral Sentence Production - INFL Devoicing (Real)**

Proficiency Level	[Vd]	[d]	[Vz]	[z]
Lower Group (n=12)	34%	32%	61%	56%
Higher Group (n=12)	14%	32%	25%	24%

A mixed two-way repeated measures ANOVA with inflection-type as within-subject and proficiency level as between-subject factors revealed the following: There was a main effect of inflection type on devoicing rates,  $F(4, 19)=13.77, p<.001$ . There was, however, no main effect of proficiency level on devoicing rates,  $F(1, 22)=0.28, p=.60$ , even though the lower level proficiency group clearly showed a higher degree of devoicing on the voiced inflections ( $mean = 46%$ ) relative to the higher level group ( $mean = 24%$ ). There was also no interaction between proficiency level and inflection type,  $F(4, 19)=0.20, p=.93$ . This result indicates that final coda devoicing was a significant factor in repair strategies for these subjects within the OSP task. Interestingly, the results of devoicing in the OSP task reveals that devoicing occurred to a lesser degree than in the OWP task (cf. Table 18 and 19). One reason for this reduction in devoicing within sentences may have been due to co-articulation that occurred between the word-final voiced consonant and the initial vowel in the following word. As a reminder, the design of the target sentences included a vowel following the target items to ensure that no word-boundary blending effects would occur between the word-final target consonant and a following consonant. However, it was observed in the Praat analysis that the voicing from the final consonant was frequently extended into the following vowel without a burst (i.e., word-final consonant release), reducing the incidences of the potential word-boundary environment to devoice the word-final consonant on the target verb.

Post-hoc analyses comparing syllabic and non-syllabic allomorphs revealed a specific significant difference between the mean of syllabic [Vd] ( $M=3.50, SD=.78$ ) and non-syllabic [d] ( $M=2.75, SD=1.03$ ),  $p=.01$ , and also between the mean of syllabic [Vz] ( $M=3.83, SD=.48$ ) and non-syllabic [z] ( $M=3.13, SD=.80$ ),  $p<.01$ . However, when examining the means of the syllabic and non-syllabic items, the differences are not supportive of the syllabicity hypothesis with

regard to devoicing rates because in some cases, the syllabic items show a higher degree of devoicing than the non-syllabic. And although there was no main effect of proficiency level, when comparing the means of the two groups, devoicing rates were obviously higher for the lower proficiency group (*mean* = 46%) compared to the higher group (*mean* = 24%), indicating that word-final devoicing occurred more frequently when the subjects' proficiency levels were lower.

#### 5.4.2 Devoicing of Nonce Items in the Oral Sentence Task

A pattern of devoicing was also observed in the nonce items in the OSP task. Table 27 shows the percentage of devoicing of voiced inflections within the OSP task nonce items.

**Table 27: Oral Sentence Production - INFL Devoicing (Nonce)**

Proficiency Level	[Vd]	[d]	[Vz]	[z]
Lower Group (n=12)	35%	42%	52%	35%
Higher Group (n=12)	21%	45%	16%	45%

A mixed two-way repeated measures ANOVA with inflection-type as within-subject and proficiency level as between-subject factors revealed the following: There was a main effect of inflection type on devoicing rates,  $F(5, 18)=15.70, p<.001$ . However, there was no main effect of proficiency level on devoicing rates,  $F(1, 22)=0.93, p=.35$ . There was an interaction between proficiency level and inflection type,  $F(5, 18)=3.39, p<.05$ , indicating that final coda devoicing at least for some of the allomorphs was a significant repair strategy for these subjects regardless of proficiency level.

Post-hoc analyses comparing syllabic and non-syllabic allomorphs revealed a specific significant difference between the means of syllabic [Vd] ( $M=2.88, SD=1.07$ ) and non-syllabic [d] ( $M=2.25, SD=1.19$ ),  $p=.004$ , but not between the means of syllabic [Vz] ( $M=2.63, SD=1.53$ ) and non-syllabic [z] ( $M=2.38, SD=1.31$ ),  $p=.45$ . In fact, as can be observed in the means and

percentages for the lower group, syllabic [Vz] was devoiced more than non-syllabic [z]. That said, the higher group appears to show a syllabicity effect on the sibilants when comparing the percentage of devoicing between syllabic [Vz] and non-syllabic [z].

#### 5.4.3 Schwa Epenthesis of Real Items in the Oral Sentence Task

Table 28 reveals the percentages of coda cluster repair via schwa epenthesis before or after the target inflection in real items. Again, items in which the inflections were omitted were excluded, but the unvoiced inflections are now included. Overall, there appears to be a clear pattern of asymmetry between the degree of repair between the allomorphs with stop codas and allomorphs with sibilant codas.

**Table 28: Oral Sentence Production - INFL Schwa Epenthesis (Real)**

Proficiency Level	[Vd]	[t]	[d]	[Vz]	[s]	[z]
Lower Group (n=12)	14%	39%	29%	0%	16%	2%
Higher Group (n=12)	18%	6%	25%	5%	14%	5%

A mixed two-way repeated measures ANOVA with inflection-type as within-subject and proficiency level as between-subject factors revealed the following: There was a main effect of inflection type on epenthesis repair rates,  $F(5, 18)=10.97, p<.001$ . However, there was no main effect of proficiency level on epenthesis repair rates,  $F(1, 22)=1.38, p=.25$ . There was an interaction between proficiency level and inflection type,  $F(5, 18)=6.33, p=.001$ , indicating that final repair of the coda cluster via epenthesis at least for some of the allomorphs was a significant factor in repair strategies for these subjects regardless of proficiency level. Post-hoc analyses comparing syllabic and non-syllabic allomorphs revealed no significant difference between syllabic [Vz] and non-syllabic [z] or between syllabic [Vd] and non-syllabic [d] allomorphs.

Overall, although proficiency level did not show a significant effect on the incidence of schwa epenthesis in oral sentence production, it is interesting to note the one allomorph in which the lower proficiency level showed a considerable increase in schwa epenthesis (i.e., [t]). This result suggests that even though L2 inflection is not subject to high rates of deletion in the lower level subjects or the higher level subjects, the lower level subjects do show a higher incidence of schwa epenthesis with regard to the least sonorous allomorph (i.e., [t]), and both lower and higher proficiency level groups show a higher incidence of schwa epenthesis on the least sonorous, non-syllabic allomorphs (i.e., [t] and [d]) when compared to the more sonorous, non-syllabic [s] and [z]. Interestingly, the results of the *different* items for both real and nonce items in the Aural Sentence Perception (ASP) task (cf. Tables 6 and 9) also reveal that the least sonorous, non-syllabic items [t] and [d] appear to be the least accurate. The relationship between perception and production is discussed further in the correlations in Section 5.3.6.

#### 5.4.4 Schwa Epenthesis of Nonce Items in the Oral Sentence Task

A similar pattern of schwa epenthesis was also observed in the nonce items in the OSP task. Table 29 shows the percentage of schwa epenthesis within the OSP task nonce items. There is a clear pattern of asymmetry in the rates of repair indicating that although word-final syllable repair has a clear effect on these subjects' oral production, there is not an equal across the board effect on all L2 inflection.

**Table 29: Oral Sentence Production - INFL Schwa Epenthesis (Nonce)**

Proficiency Level	[Vd]	[t]	[d]	[Vz]	[s]	[z]
Lower Group (n=12)	23%	44%	50%	6%	17%	40%
Higher Group (n=12)	33%	21%	27%	10%	6%	21%

A mixed two-way repeated measures ANOVA with inflection-type as within-subject and proficiency level as between-subject factors revealed the following: There was a main effect of inflection type on epenthesis repair rates,  $F(5, 18)=12.58, p<.001$ . However, there was no main effect of proficiency level on epenthesis repair rates,  $F(1, 22)=1.85, p=.19$ . There was an interaction between proficiency level and inflection type,  $F(5, 18)=3.28, p<.05$ , indicating again that final repair of the coda cluster via epenthesis at least for some of the allomorphs was a significant factor in repair strategies for these subjects regardless of proficiency level. In fact, the rates of repair of the non-syllabic [t], [d], and [z] appear higher for the lower proficiency group than for the higher proficiency group. Post-hoc analyses comparing syllabic and non-syllabic allomorphs revealed no significant difference between syllabic [Vz] and non-syllabic [z] or between syllabic [Vd] and non-syllabic [d] allomorphs.

#### 5.4.5 Schwa Epenthesis Plus Devoicing of Real Items in the Oral Sentence Task

Table 30 reveals the percentage of items in which the target inflection was repaired both via devoicing and schwa epenthesis in the real items. Again, items in which the inflections were omitted and in which the inflections were unvoiced were excluded.

**Table 30: Oral Sentence Production - INFL Devoicing + Schwa Epenthesis (Real)**

Proficiency Level	[Vd]	[d]	[Vz]	[z]
Lower Group (n=12)	13%	32%	4%	32%
Higher Group (n=12)	15%	32%	6%	28%

A mixed two-way repeated measures ANOVA with inflection-type as within-subject and proficiency level as between-subject factors revealed the following: There was a main effect of inflection type on devoicing plus epenthesis repair rates,  $F(4, 19)=17.95, p<.001$ . However, there was no main effect of proficiency level on devoicing plus epenthesis repair rates,  $F(1, 22)=0.15, p=.70$ , and there was no interaction between proficiency level and inflection type,  $F(4, 19)=1.09,$

$p=.39$ . Thus, final coda devoicing plus epenthesis repair of the coda cluster appears to be a significant factor in repair strategies for these subjects in the oral sentence task, even though proficiency level effects are not a significant factor in this repair.

Post-hoc analyses comparing syllabic and non-syllabic allomorphs revealed no specific significant difference between syllabic [Vz] and non-syllabic [z] or between syllabic [Vd] and non-syllabic [d] allomorphs. However, in examining the mean rates of schwa epenthesis, it should be noted that even though the comparisons did not reach significance, the syllabic items were repaired less frequently than the non-syllabic items. Thus, this result provides further support for the argument that relative syllabicity (i.e., syllabic vs. non-syllabic) has an effect on repairs of target L2 inflections.

#### 5.4.5 Schwa Epenthesis Plus Devoicing of Nonce Items in the Oral Sentence Task

A similar pattern of schwa epenthesis plus devoicing was also observed in the nonce items in the OSP task. Table 31 shows the percentage of schwa epenthesis plus devoicing within the OSP task nonce items.

**Table 31: Oral Sentence Production - INFL Devoicing + Schwa Epenthesis (Nonce)**

Proficiency Level	[Vd]	[d]	[Vz]	[z]
Lower Group (n=12)	15%	38%	4%	33%
Higher Group (n=12)	18%	23%	6%	19%

A mixed two-way repeated measures ANOVA with inflection-type as within-subject and proficiency level as between-subject factors revealed the following: There was a main effect of inflection type on devoicing plus epenthesis repair rates,  $F(4, 19)=15.44$ ,  $p<.001$ . However, there was no main effect of proficiency level on devoicing plus epenthesis repair rates,  $F(1, 22)=1.30$ ,  $p=.27$ , and there was no interaction between proficiency level and inflection type,  $F(4, 19)=1.54$ ,  $p=.23$ . Thus, final coda devoicing plus epenthesis repair of the coda cluster appears to be a

significant factor in repair strategies for these subjects in the oral sentence task, even though proficiency level effects are not a significant factor in this combined repair.

Post-hoc analyses comparing syllabic and non-syllabic allomorphs revealed no specific significant difference between syllabic [Vd] and non-syllabic [d] allomorphs. However, there was a significant difference between [Vz] ( $M=3.79$ ,  $SD=0.41$ ) and non-syllabic [z] ( $M=2.95$ ,  $SD=1.081$ ),  $p=.01$ . Also, even though the [Vd] to [d] comparison did not reach significance, it should be noted that the syllabic items were repaired less frequently than the non-syllabic items. Again, this result provides further support for the argument that relative syllabicity (i.e., syllabic vs. non-syllabic) has an effect on repairs of target L2 inflections.

The lack of across the board epenthesis (i.e., across all allomorphs) reveals that although schwa epenthesis was a factor for some of the allomorphs, evidence is again lacking for an argument of an across the board effect of L1 phonotactic interference on deletion of target codas. Nevertheless, there are clear effects of L1 phonotactics on specific codas, but the effects are asymmetrical and regularly target the less sonorous stops more than the sibilants. As will be discussed in more detail in Chapter 6, this result also indicates a lack of evidence for a prosodic deficit as would be the case under a strict PTH analysis of Goad et al. (2003), but in fact, the results may still support Davidson's OT account of "floating constraints" such that asymmetric application of schwa epenthesis in the output may reflect an interlanguage rule that is variably applied under different processing conditions.

### 5.5 Analysis of Repairs in Monomorphemic Items

The following results reveal the different repairs and their occurrence in the monomorphemic items that contain word-final codas that violate L1 phonotactics within the

OWP task. Comparisons are made between repairs in the monomorphemic items and the repairs in the real items to provide more support for the argument that the word-final repairs are due to syllable-level effects and are not likely due to morphosyntactic or prosodic deficits in the L2.

### 5.5.1 Devoicing in Monomorphemic Items

Table 32 reveals the percentage of devoicing in these monomorphemic items. For obvious reasons, only the voiced items are included. Again, the categories are separated into word-final stops and sibilants that mirror the target allomorphs under investigation.

**Table 32: Oral Word Production – Monomorphemic Devoicing**

Coda Type	Lower Group (n=19)	Higher Group (n=21)
Voiced Stops (n=6)	58%	60%
Voiced Fricatives (n=6)	65%	68%

A mixed two-way repeated measures ANOVA with word-final coda type as within-subject and proficiency level as between-subject factors revealed the following: There was no main effect of coda type on devoicing between the stops and fricatives,  $F(1, 22)=0.14, p=.72$ . There was also no main effect of proficiency level of devoicing between monomorphemic stops and fricatives  $F(1, 22)=0.59, p=.45$ , and there was no interaction between proficiency level and coda type,  $F(1, 22)=0.02, p=.90$ . Thus, although devoicing was observed to a fairly high degree in the monomorphemic items, there were no significant differences due to the coda type (i.e., stops vs. fricatives) or due to proficiency level. However, the degree of devoicing in these items are generally comparable to the devoicing rates in the real items in the OWP task, and as will be shown below, there were positive correlations between inflected items and monomorphemic items with regard to this repair.

### 5.5.2 Schwa Epenthesis in Monomorphemic Items

Table 33 reveals the percentage of schwa epenthesis in the monomorphemic items. Again, the categories are separated into codas with stops and sibilants that mirror the target allomorphs under investigation. In this analysis, all voiced and unvoiced coda types are included.

**Table 33: Oral Word Production – Monomorphemic Schwa Epenthesis**

<b>Coda Type</b>	<b>Lower Proficiency (n=19)</b>	<b>Higher Proficiency (n=21)</b>
Unvoiced Stops (n=6)	1%	0%
Voiced Stops (n=6)	8%	14%
Unvoiced Fricatives (n=6)	0%	0%
Voiced Fricatives (n=6)	10%	1%

A mixed two-way repeated measures ANOVA with word-final coda type as within-subject and proficiency level as between-subject factors revealed the following: There was a marginal effect of coda type on schwa epenthesis between the stops and fricatives,  $F(3, 20)=2.94, p=.058$ . There was no main effect of proficiency level on schwa epenthesis between monomorphemic stops and fricatives  $F(1, 22)=0.13, p=.72$ . However, there was an interaction between proficiency level and coda type,  $F(3, 20)=4.01, p<.05$ . Thus, although schwa epenthesis alone was observed at a fairly low rate in the monomorphemic items, there were observable differences due to the coda type (i.e., stops vs. fricatives), and this effect appeared to affect both proficiency levels.

### 5.5.3 Schwa Epenthesis Plus Devoicing in Monomorphemic Items

Table 34 reveals the percentage of devoicing plus schwa epenthesis in the monomorphemic items. Again, the categories are separated into codas that contain stops and sibilants that mirror the target allomorphs under investigation, and the unvoiced items are excluded.

**Table 34: Oral Word Production – Monomorphemic Epenthesis + Devoicing**

<b>Coda Type</b>	<b>Lower Proficiency (n=19)</b>	<b>Higher Proficiency (n=21)</b>
Voiced Stops (n=6)	25%	10%
Voiced Fricatives (n=6)	15%	7%

A mixed two-way repeated measures ANOVA with word-final coda type as within-subject and proficiency level as between-subject factors revealed the following: There was a main effect of coda type on schwa epenthesis plus devoicing between the stops and fricatives,  $F(2, 21)=16.21, p<.001$ . There was no main effect of proficiency level on schwa epenthesis plus devoicing between monomorphemic stops and fricatives  $F(1, 22)=1.56, p=.23$ , and there was no interaction between proficiency level and coda type,  $F(2, 21)=0.75, p=.49$ . Thus, schwa epenthesis plus devoicing combined revealed a greater effect on the voiced stops than on the fricatives, and although there was no main effect of proficiency level, there were obvious differences between the proficiency levels with the lower group producing monomorphemic items with higher rates of repair. Again, similar to the repairs in the real items in the OWP task, the schwa epenthesis plus devoicing rates in these monomorphemic items reveal that rates of devoicing plus epenthesis are higher on the voiced stops.

#### 5.5.4 Schwa Epenthesis Totals in Monomorphemic Items

Finally, Table 35 reveals the percentage of total schwa epenthesis in the monomorphemic items combined. Again, the categories are separated into codas that contain stops and sibilants that mirror the target allomorphs under investigation.

**Table 35: Oral Word Production – Total Monomorphemic Schwa Epenthesis**

<b>Coda Type</b>	<b>Lower Proficiency (n=19)</b>	<b>Higher Proficiency (n=21)</b>
Unvoiced Stops	26%	10%
Voiced Stops	8%	14%
Unvoiced Fricatives	0%	0%
Voiced Fricatives	25%	8%

The degree of schwa epenthesis in the real items and the monomorphemic items in the OWP task show similar patterns of schwa epenthesis. And although the rates of schwa epenthesis on the monomorphemic items are not as robust as the real items, there is a clear asymmetry between the stops and fricatives and between the lower group and higher group that suggest an effect of sonority of the coda. This result will be discussed further in Chapter 6 along with the results of the correlations between the repair rates of monomorphemic and real items.

#### 5.6 Correlations between Perception and Production and within Oral Production

Within the field of Second Language Acquisition, some have claimed that there is a link between aural perception and oral production such that aural perception leading to intake of language forms likely must occur before production of the language forms can occur (cf. Flege, 1995; Jia et al., 2006 for a discussion).<sup>26</sup> In order to determine the relationship between the perception of target inflections and their production, this researcher conducted correlations between the accuracy rates of target inflections within the perception tasks and production tasks. The objective was to see if accuracy in perception could be correlated to accuracy in overall production. These analyses below included only the real items, as nonce items were determined to reveal no additional information with regard to the relationship between perception and production. In other words, if it were not evident in the real items, then further analysis of nonce

---

<sup>26</sup> In an isolated study, Sheldon and Strange (1982) found contradictory evidence that suggested that in some cases, L2 learners were able to produce segments that they were unable to perceive. However, this study was examining phones in isolation, not L2 inflection, and has not been replicated.

items would not confirm the specific relationship in this study even though accuracy of nonce items was lower than real items. Comparisons were made between the accuracy rates of the different<sup>27</sup> items in the Aural Perception tasks to the following three separate accuracy rates within the Oral Production tasks:

- 1) **Different** items (e.g., “walk/walks”) in Perception to **Suppliance** in Production (e.g., “She walks on the sidewalk.”)
- 2) **Different, Voiced** items (e.g., bleeds/bleed) in Perception to **Devoicing** in Production (e.g., “bleeds” [bliyds])
- 3) **Different** items (e.g., walk/walks) in Perception to **Schwa Epenthesis** in Production (e.g., “walkse’ [wɔk.sə])

In the first comparison, as accuracy of perception increases, an increase in accuracy of production should be observed such that as perception increases, suppliance of target items should also reveal an increase in accuracy. A positive correlation in this first comparison would suggest that perception and production of L2 inflection increase together incrementally, thus, providing support for a positive relationship between these two aspects of performance with regard to L2 inflectional acquisition.

In the second and third comparisons, decreased accuracy in perception should reveal an increase in repairs. These analyses looked at reduced perception accuracy to reduced accuracy in target production of the codas—in other words, with regard to repairs, reduced accuracy in production was indicated by a higher rate of repairs, either devoicing or schwa epenthesis within the target coda clusters.

---

<sup>27</sup> Same items were excluded from the analysis because these items do not require the participants to distinguish between items with inflections and without inflections – it is understood that different items require the participants to perceive a missing inflection and, therefore, are more comparable to accessing the grammar during production.

### 5.6.1 Different Items in Aural Word Task to Suppliance in Oral Word Task

First, the accuracy of the different items on the Aural Word Perception (AWP) task were correlated with the suppliance of target codas on the Oral Word Production (OWP) task using a Pearson product-moment correlation coefficient to see if there was a positive correlation between the accuracy on these two tasks. There was no significant correlation between the accuracy of these two tasks:  $r=.13$ ,  $p=.42$  (Sig. 2-tailed).

### 5.6.2 Different, Voiced Items in Aural Word Task to Devoicing in Oral Word Task

Then the accuracy of the voiced, different items in the AWP task was correlated with the rates of devoicing of the voiced items in the OWP task to determine if perceptual accuracy would significantly correlate with production accuracy in terms of devoicing of the target codas. There was also no significant correlation between accuracy rates of the AWP voiced, different items and the OWP voiced items:  $r=.06$ ,  $p=.79$  (Sig. 2-tailed).

### 5.6.3 Different Items in Aural Word Task to Schwa Epenthesis in Oral Word Task

Comparing the accuracy of the different items in the AWP task to the rates of schwa epenthesis of the target codas in the OWP task, again, no significant correlation was found between accuracy rates of the AWP different items and the OWP items that were repaired via schwa epenthesis in the coda:  $r=.09$ ,  $p=.69$  (Sig. 2-tailed).

### 5.6.4 Different Items in Aural Sentence Task to Suppliance in Oral Sentence Task

Then the accuracy of the different items on the Aural Sentence Perception (ASP) task was correlated with the accuracy of suppliance of the target inflections on the Oral Sentence

Production (OSP) task using a Pearson product-moment correlation coefficient to see if there was a positive correlation between the accuracy on these two tasks. There was no significant correlation between the accuracy of these two tasks:  $r=.13$ ,  $p=.43$  (Sig. 2-tailed). Again, overall accuracy on these two tasks was fairly high, and thus, the possibility of finding a distinct positive or negative correlation was low due to the high accuracy rates of perception vs. suppliance of target items.

#### 5.6.5 Different, Voiced Items in Aural Sentence Task to Devoicing in Oral Sentence Task

The accuracy of the voiced, different items in the ASP task was correlated with the rates of devoicing of the voiced items in the OSP task to determine if perceptual accuracy would significantly correlate with production accuracy in terms of devoicing of the target inflections. There was also no significant correlation between accuracy rates of the ASP voiced, different items and the OSP voiced items:  $r=.17$ ,  $p=.43$  (Sig. 2-tailed).

#### 5.6.6 Different Items in Aural Sentence Task to Schwa Epenthesis in Oral Sentence Task

Comparing the accuracy of the different items in the ASP task to the rates of schwa epenthesis of the target codas in the OSP task, a significant correlation was found between accuracy rates of the ASP different items and the OSP items that were repaired via schwa epenthesis in the coda:  $r=.42$ ,  $p<.05$  (Sig. 2-tailed), thus, indicating at least for these subjects that repair of the codas in production via schwa epenthesis was more likely to occur when accuracy of their perception was lower.

The specific relationships between the each inflection type under study regarding the schwa epenthesis repair was examined by comparing like allomorphs in the ASP task to like

allomorphs in the OSP task (e.g., [t] to [t]); a significant correlation was found only between items with the [t] allomorphs:  $r=.55, p<.01$  (Sig. 2-tailed). Thus, when their perceptual accuracy was lower, the subjects in this study were most likely to epenthesize a schwa to repair codas that specifically contained the [t] allomorph. And although there was a significant correlation observed in this relationship between [t] allomorphs in *different* ASP items and [t] in OSP items in terms of schwa epenthesis repair, on the whole, accuracy of L2 inflection and word-final codas was still high, which suggests that these L2 learners are at an advanced level of perception and production. Nevertheless, aural perception and oral production of real items did not appear to correlate overall, except in the isolated case of perception of *different* items and schwa epenthesis within items with [t] allomorphs.

### 5.7 Oral Word Repairs in Real and Nonce Items to Monomorphemic Items

To determine the degree to which word-final repairs pattern after specific allomorphs (stops vs. fricatives), correlations were run to compare the repair patterns of the real and nonce OWP items to the repair patterns in the OWP monomorphemic items that contain word-final target codas. Positive correlations between these items would reveal that the repairs are productive to all target codas that are illicit to L1 syllable phonotactic constraints, rather than due to a prosodic constraint that targets only the real verbal items as claimed by the Prosodic Transfer Hypothesis (PTH).

Devoicing and syllable repairs obviously had a significant effect on the oral production of the target codas in this study. Similar patterns of repair were observed not only in the real and nonce items, but also there were significant degrees of repair that were observed in some of the word-final consonants in the monomorphemic items. To determine if the patterns of repair were

comparable between the real and nonce items and the monomorphemic items, Pearson product-moment correlations were conducted that compared the repair rates of the real and nonce items that contained the target word-final stop and fricative codas to monomorphemic items with the same word-final stops and fricatives. The results showed significant relationships are as follows:

First, there was a significant positive correlation between the degree of devoicing in the real items that contained word-final stops [Vd] and devoicing of the monomorphemic items that contained voiced stops,  $r=.53$ ,  $p=.007$  (Sig. 2-tailed). There was also a highly significant positive correlation between the degree of devoicing in the real items that contained word-final stops [d] and devoicing of the monomorphemic items that contained voiced stops,  $r=.78$ ,  $p<.001$  (Sig. 2-tailed). The fact that devoicing occurs on monomorphemic items as well as real items in a strong positive relationship is, therefore, indicative that the devoicing repair is independent of inflection, and as suggested by other results above, it appears that the devoicing rates are, instead, dependent on the sonority of the coda (i.e., stops vs. fricatives) rather than due to a broader L1 phonotactic constraint against word-final consonants. As such, no significant relationship was observed when comparing the [Vz] and [z] target codas to the monomorphemic voiced fricatives.

When examining the correlations between the nonce items that contained the target [Vd], [t], and [d] stops compared to the monomorphemic word-final stops, again, the same result appeared. That is, there was a significant positive correlation between the degree of devoicing in the nonce items that contained word-final stops [Vd] and devoicing of the monomorphemic items that contained voiced stops,  $r=.68$ ,  $p<.001$  (Sig. 2-tailed). There was also a significant positive correlation between the degree of devoicing in the real items that contained word-final stops [d] and devoicing of the monomorphemic items that contained voiced stops,  $r=.60$ ,  $p=.002$

(Sig. 2-tailed). Again, no significant relationship was observed when comparing the [Vz] and [z] target codas to the monomorphemic voiced fricatives.

Second, there was a significant positive correlation between the degree of schwa epenthesis in the real items that contained word-final stops [Vd] and epenthesis on the monomorphemic items that contained voiced stops,  $r=.66$ ,  $p=.001$  (Sig. 2-tailed). There was also a highly significant positive correlation between the degree of epenthesis in the real items that contained word-final stops [d] and epenthesis of the monomorphemic items that contained voiced stops,  $r=.84$ ,  $p<.001$  (Sig. 2-tailed). In the case of unvoiced [t] and [s], interactions were not significant because many of the monomorphemic items in these categories showed no repair. Nevertheless, the fact that schwa epenthesis occurs on monomorphemic items as well as real items in a strong positive relationship is, again, indicative that the epenthesis repair is independent of inflection, and as suggested by other results above, it appears that the epenthesis rates are, instead, dependent on the sonority of the coda (i.e., stops vs. fricatives) rather than due to a broader L1 phonotactic constraint against word-final consonants. Again, there were no significant effects on the [Vz] and [z] items when compared to the monomorphemic items that ended in voiced fricatives.

Finally, when examining the correlations of schwa epenthesis repair between the nonce items that contained the target [Vd], [t], and [d] stops compared to the monomorphemic word-final stops, again, the same result appeared. That is, there was a significant positive correlation between the degree of epenthesis repair in the nonce items that contained word-final stops [Vd] and epenthesis of the monomorphemic items that contained voiced stops,  $r=.59$ ,  $p=.002$  (Sig. 2-tailed). There was also a significant positive correlation between the degree of epenthesis in the real items that contained word-final stop [d] and epenthesis of the monomorphemic items that

contained voiced stops,  $r=.59$ ,  $p=.003$  (Sig. 2-tailed). Again, there was no significant interaction for [t] or [Vz], [s], or [z]. As will be discussed in Chapter 6, the strong correlations between certain real and nonce target codas and the monomorphemic items, while others showed no significant interaction, supports this researcher's conclusion that inflection is not being affected by an across the board effect of L1 phonotactic constraints, and even more importantly, word-final codas are not being repaired due to a deficit within the prosodic representation of the L2 grammar.

## Chapter 6: General Discussion

### 6.1 Evidence of Morphosyntactic Representation

In their Representational Deficit Hypothesis (RDH), Hawkins and Liszka (2003) have claimed that variability of suppliance of English inflections in oral production is generally indicative of a deficit in representation when the L1 of an adult L2 learner lacks the necessary syntactic features to be transferred from the L1 to the L2. Thus, in the case of adult Mandarin speakers, for example, who lack the instantiation of tense in their L1, Hawkins and Liszka claim that the inability to transfer the tense feature [+/-past] to the L2 results in variable suppliance of tense in the L2 even in the end-state grammar. As should be expected, this claim would entail unsystematic or arbitrary omission across all L2 inflections (i.e., past *-ed* and 3<sup>rd</sup> singular *-s*) in all Mandarin speakers' English inflectional production at all levels of acquisition. Additionally, with regard to perception of L2 inflections, if adult Mandarin speakers of English were unable to acquire the English tense feature due to a critical period for L2 acquisition of features that are not instantiated in the L1 as Hawkins and Liszka also claim, this lack of instantiation within the L2 grammar would undoubtedly lead to an inability to regularly perceive the past tense as well because, as this researcher and others have claimed, the target grammar must contain the grammatical representation for successful parsing of both the input (i.e., perception/comprehension) and the output (i.e., production). In fact, the results of this dissertation did not reveal an across the board reduction of the accuracy rates of target L2 inflection in either production or perception. On the contrary, the subjects revealed target-like knowledge of the regular past and 3<sup>rd</sup> singular inflections in the written grammaticality screening tasks, scoring at a mean of 98% suppliance of target L2 inflection on the GJT1 and mean of 96%

suppliance of target L2 inflection in GJT2 with the groups collapsed. The grammaticality judgment screening task was designed to determine if these subjects' knowledge of L2 inflection was advanced. It was found to be advanced for all of the subjects.

However, what was also observed in the results was the near-ceiling suppliance rates of target L2 inflections in the oral sentence production task for both the lower and higher proficiency groups collapsed (mean = 93%) further supporting this researcher's claim that a morphosyntactic deficit for past tense in the L2 grammar is not attested in these Mandarin speakers. Thus, although not a main research question within this dissertation, contra Hawkins and Liszka's claim that variable production of past tense L2 inflection is a result of lack of L2 representation for past tense, the results of this dissertation provide further evidence for the claim that acquisition of English past tense is possible for adult Mandarin speakers who have learned English as a L2, and based on these near-ceiling suppliance rates, the past tense and 3<sup>rd</sup> singular grammatical representation for these L2 learners is target-like.

## 6.2 Effects of L1 Phonotactics, Salience, and Processing Load on Perception and Production

Knowledge of L2 inflection was clearly attested for these L2 learners, and thus, it was ruled out as a potential cause of variability of the L2 surface inflections. As such, the following three main factors were hypothesized to affect the accuracy of these L2 learners' perception and production of L2 inflection: 1) L1 phonotactic interference, 2) salience (i.e., relative acoustic perceptibility) of L2 inflectional allomorphs, and 3) increased processing load leading to an overload on attentional resources that cause a reduction in accuracy of perception and production. In addition, it was hypothesized that proficiency level would have an overall effect

on the relationship between perception and production such that accuracy rates of perception and production would correlate.

### 6.2.1 L1 Phonotactic Effects

In terms of L1 phonotactic interference, it was hypothesized that L1 phonotactic constraints would negatively affect the perception and production of L2 inflections such that when a syllable coda violated the L1 phonotactic constraints, performance accuracy for that coda would be significantly reduced in both perception and production. Accordingly, it was claimed that the reduction in performance for any codas that contain target inflections would be exhibited mainly in deletion or repair of one or more consonants in the final target cluster, resulting in significantly reduced accuracy rates of the target codas and inflections.

First, L1 phonotactic interference revealed no across the board effects on the accuracy of perception or suppliance of L2 inflection in either the word or sentence tasks. It was expected that because the coda clusters that make up word-final inflections in English are illicit according to the syllable phonotactics of Mandarin, accuracy in perception would be significantly reduced and deletion in production of the L2 inflections would be significantly increased across all of the target codas including real, nonce, and monomorphemic items. The study did not find this across the board result for all inflections. In fact, the rates of deletion of target codas and L2 inflections were surprisingly low. However, as hypothesized, significant repairs in terms of devoicing of final segments, schwa epenthesis within the word-final clusters and after the word-final segments, and a combination of the two repairs (i.e., devoicing and schwa epenthesis) were observed. Although these results are not supportive of some research regarding L1 phonotactic interference on perception (e.g., Dupoux et al. 1999), it is supportive of Brown's (1998),

Davidson's (2005, 2006a, 2006b), and Kabak and Idsardi's (2007) claims that under the appropriate conditions, L2 learners can overcome L1 phonological constraints and acquire novel L2 phonological features and syllable constraints; the results are also supportive of Goldrick's (2004) and Kabak and Idsardi's (2007) claims that instead of a general syllable-level condition effect on performance, a more fine-grained analysis of syllable-contact conditions is warranted because it appears that general syllable conditions and segmental-level effects may act independently. In fact, the systematic asymmetrical application of coda repairs that was observed and that will be discussed are indicative of phonetic and performance effects as proposed.

Additionally, within the word production task, the accuracy rates are even more in conflict with an across the board L1 phonotactic interference effect on deletions of target codas; accuracy rates for the real and nonce items are between 80% and 100% except for one allomorph [Vz], which was presumed to be omitted mainly because of an articulatory issue such that the final coda of the root words were all sibilants causing the subjects to have articulation issues of the final [zVz] syllable. In other words, the possible cause of this significant reduction in accuracy compared to the other target codas may be due to fact that target items that required the syllabic [Vz] already contained a final sibilant in the root coda prior to the inflectional ending (e.g., *raise, dance, fix, touch*). Thus, the subjects' lack of articulation of the final [Vz] may have been due to articulation challenges caused by the two corresponding sibilants in the word final context, resulting in a moderate level of deletion of the final syllable in some cases. However, because this phonemic level analysis was beyond the scope of this study, it was not investigated further. Nevertheless, the accuracy of the [Vz] allomorph was between 70% and 75%, even though the subjects' suppliance of this allomorph in the grammar and oral sentence tasks was

target-like, suggesting that this reduction of accuracy may be indicative of a performance or processing effect, and not a deficit in the grammar.

Within both the word perception and word production tasks, the accuracy rates range from near ceiling to at ceiling, and monomorphemic items in the oral word task are at ceiling in terms of suppliance of target codas. Target-like suppliance of the codas in the oral word task alone cannot provide unequivocal evidence against a strict Prosodic Transfer Hypothesis deficit account, in which reduced accuracy of L2 inflection is a result of the lack of the L2 prosodic structure due to differences between Mandarin and English prosodic structure. However, as will be discussed, the systematic asymmetries that are observed on specific coda types and allomorphs across the word and sentence tasks indicate that perception and production are mediated by specific phonetic effects, depending on the phonetic qualities of the coda consonants, and that performance effects (such as processing load) increase these effects on accuracy rates of target codas.

In fact, it can be concluded that when L1 syllable phonotactics disallow word-final obstruents (i.e., [t], [d], [s], and [z]), L1 phonotactics do not disallow these obstruents across the board for these advanced L2 learners. As such, a more fine-grained analysis of the coda repairs below will provide illuminating conclusions about the specific syllable and segmental-level repairs of the codas and the effects of these repairs on performance. As will also be discussed, proficiency level effects were evident under some specific conditions, but these subjects clearly showed an advanced to target-like ability to perceive and produce the target codas and target inflections.

### 6.2.2 Salience Effects

It was hypothesized that varying degrees of salience of target codas would have a negative effect on performance such that reduced accuracy would be observed in both perception and production for the less salient (i.e., non-syllabic) allomorphs ([t], [d], [s], and [z]) than for the more salient (i.e., syllabic) allomorphic variants ([Vd], [Vz]), following Solt et al., (2004) and Klein et al., (2004). In addition, considering the varying degrees of sonority between stops and fricatives (cf. Ladefoged, 2006), accuracy was also expected to be lower for stops ([t] and [d]) compared to accuracy rates of the fricatives ([s] and [z]), and lower on the non-syllabic, unvoiced inflections ([t] and [s]) compared to their non-syllabic, voiced counterparts ([z] and [d], respectively).

The results revealed a clear effect of salience on the subjects' accuracy across both perception and production, as was also found in Solt et al. (2004) and Klein et al. (2004). First, as noted above, because the accuracy rates of perception and production in the word tasks were near ceiling for nearly all of the target codas in the real items, an across the board effect of salience within the two word tasks was not observed. There was a noticeable, but non-significant reduction in the perception of the real items that contained target coda [t] in the AWP word perception task and a significant reduction of the perception of nonce items that contained the coda [d] in the AWP task. Nonce, *different* items within the AWP task showed the greatest effect of salience on accuracy revealing that the effects of salience were productive even on non-lexicalized items. The lower rates of accuracy on the *different*, nonce items (vs. real items) were likely due to the fact the subjects could not recall these items from the lexicon and, thus, presumably perceived these items at a purely acoustic level because they could not rely on lexical recall. There was no effect of syllabicity in the OWP word production task, which is not

surprising considering that the suppliance of target codas in the OWP task was at or near ceiling. In fact, most of the coda types in the AWP and OWP tasks were perceived and produced at ceiling. As will be discussed, the subjects' advanced aural and oral proficiency allowed for advanced to target-like word-level perception and production of the target items.

That said, there was a significant effect of sonority between the stops (i.e., [t] and [d]) and the sibilants (i.e., [s] and [z]) in the AWP task with sibilants [s] and [z] exhibiting significantly higher accuracy than stops [t] and [d], revealing that accuracy in perception of word-final codas is affected by the degree of sonority between segments that have different phonetic features. However, significant effects were not observed in the OWP word production task with regard to deletion and sonority differences. There was also no significant effects of voicing differences between the unvoiced [t] and [s] and voiced [z] and [d] in either task.

Furthermore, as will be discussed in Section 6.3 below, significant correlations were found between the perception of items that contained the coda [t] and the degree of schwa epenthesis that occurred in these same items in production. This result suggests not only that accuracy of perception and production of specific word-final segments do appear to trend together at least for some codas (following Jia et al., 2006), but also it further supports the approach to analysis of grammatical variability that is proposed by Davidson (2005, 2006a, 2006b), Goldrick (2004), Kabak and Idsardi (2007), and Trenkic (2007)—that is, analyzing variability is not simply a matter of observing reduction of accuracy in suppliance of L2 inflection across the board, but the analysis must closely examine the asymmetries in both perception and production to determine the specific causes of the variability before one can make conclusions with regard to deficits in the grammar.

The asymmetries in perceptual accuracy with regard to salience showed an even stronger effect in the sentence perception task (ASP). The *same* items were separated from the *different* items as noted above because it was presumed that perceiving the difference between an inflected and uninflected form of a verb in sentences would obviously put more processing pressure on the language parser because the uninflected item within the pair requires the parser to access the grammar to determine if an item is grammatically correct, while presumably *same* items could be judged at a phonetic level without needing to access the level of L2 inflection. In other words, within the *same* items, although the subjects may be accessing the grammar to parse the sentences, they could also be correctly matching these items simply on the grounds of surface acoustics without parsing the L2 inflection. As such, the effects of salience on accuracy rates are increased as the level of processing pressure increases. This presumed difference between *same* and *different* items was evident in the accuracy rates of the two separate item types; consequently, only one comparison of syllabicity of the allomorphs within the *same* sentence items reached significance, which was the syllabic [Vz] items and the non-syllabic [z] items. However, this syllabicity effect was not seen when comparing the [Vd] and [d] items within the *same* items. Thus, as claimed, it is likely that performance pressure was lower on the *same* items leading to higher accuracy overall when compared to the results of the *different* items.

Consequently, perceiving *different* items put significantly greater processing pressure on the parser, as indicated by the much lower accuracy rates on the *different* items in the sentence perception task. For example, with regard to syllabicity, non-syllabic past regular inflected items with [t] and [d] codas showed significantly lower accuracy rates collapsed across the two groups ([t]=32% and [d]=50%) when compared to the syllabic past regular inflected items ([Vd]=76%). Further, for the lower level proficiency subjects, the non-syllabic items with [z] codas also

showed a much lower accuracy at 58% than the corresponding syllabic items ([Vz]=82%). As expected, the higher proficiency level subjects scored significantly higher than the lower proficiency level subjects. However, the [s] inflected items showed no significant effect of reduced salience due to syllabicity effects. It was concluded that the high degree of sonority created by the higher acoustic intensity of the [s] codas caused them to be more perceptually salient, and thus, these sibilant codas were significantly more accurately perceived than the [z] items. This result is further supportive of Kabak and Idsardi's (2007) and Goldrick's (2004) claims that general syllable-level effects (i.e., L1 phonotactic constraints) may be independent of more fine-grained segmental conditions. It was concluded that in the case of these Mandarin speakers, they have overcome the general syllable-level effects of their L1 phonotactic constraints as evidenced by a lack of across the board (i.e., on all codas and inflections) reduction of accuracy in perception and production; however, the subjects are still subject to the effects of salience, specifically with regard to syllabicity and sonority effects at the phonetic-acoustic level that resulted in asymmetrical performance of word-final target codas between the stops and sibilants that are even more apparent in the sentence tasks.

Overall, within the sentence perception results (especially in the *different* items), accuracy of some specific inflectional allomorphs showed a significant reduction in accuracy. For example, for the lower group in the sentence perception task (i.e., *different*, real items), the subjects produced the [t] allomorph at 16% accuracy, the [d] allomorph at 37% accuracy, and the [z] allomorph at 58% accuracy. In the higher group, although more accurate than the lower group, their accuracy rates were also noticeably low for the [t] allomorph at 48% and the [d] allomorph at 62%. The nonce items in the sentence perception task also showed lower accuracy rates as well on specific inflectional allomorphs: e.g., the lower group = [t] at 69%; [d] at 24%;

[s] at 71%; and the higher group = [Vd] at 73%; [t] at 75%; [d] at 45%. However, accuracy rates of the real *same* items in sentence perception never dropped below 75% and nonce *same* sentence perception items never dropped below 70% with most allomorphs being perceived at or above 80% accuracy, suggesting that performance effects due to higher processing pressure on the *different* items resulted in lower accuracy compared to the *same* items.

Additionally, suppliance of L2 inflection in the OSP sentence production task was quite high. However, no significant effect of syllabicity on production of L2 inflection in the real or nonce target items was found. This same result was found in the oral elicited production results in Martohardjono et al. (2012). There was, however, a significant effect of inflection type in the nonce target items in the oral sentence production. That said, the reduction of accuracy was more pronounced on the syllabic [Vz] and [Vd] than on the non-syllabic [t], [d], [s], and [z] allomorphs. This result is contradictory to the syllabicity results in the ASP task. It was concluded that the reduction in accuracy, especially of the syllabic [Vz], was due to challenges in articulation just as in the OWP task. Again, the possible cause of this significant reduction in accuracy compared to the other allomorphs may be due to fact that target items that required the syllabic [Vz] already contained a final sibilant in the root coda prior to the inflectional ending (e.g., *raise, dance, fix, touch*). Thus, the subjects' lack of articulation of the final [Vz] may have been due to articulation challenges caused by the two corresponding sibilants in the word final context, resulting in a moderate level of deletion of the final syllable in some cases. Although this phonemic level analysis was beyond the scope of this study, it reveals that the asymmetries in the word level tasks are productive at the sentence level as well and that this particular systematic asymmetry is more indicative of performance effects than deficits in the grammar.

Overall, in terms of salience and its effect on accuracy of performance, the ASP sentence perception task results revealed an obvious effect of syllabicity on L2 inflection, especially in the *different* items. There was also an overall effect of sonority between the stops and sibilants as indicated by the much higher rate of accuracy on the [s] and [z] inflected items compared to the [t] and [d] inflected items in the AWP and ASP. However, the voicing differences between the [t] and [d] allomorphs and the [s] and [z] allomorphs, respectively, did not reveal an effect of improved perception due to voicing vs. the unvoiced allomorphs—a difference in accuracy was only obvious between [t] (32%) and [d] (50%), but this effect was not consistent across the tasks.

Finally, the effects of salience were more pronounced on the nonce items, revealing that higher processing pressure may have increased the effects of salience, especially in perception. In addition, the effects of sonority on accuracy was more pronounced for the lower level proficiency subjects than for the higher level subjects, providing evidence that even at an advanced level of L2 grammatical knowledge, performance effects at the acoustic perceptual level causing variability of L2 inflection can still be susceptible to proficiency level effects.

It is important to note that accuracy of perception of target inflections in sentences was lower when compared to the suppliance of L2 inflections in the sentence production tasks; thus, even though perception of the target codas was significantly in error, suppliance of the same target items was not. As such, this result reveals that when conducting studies that analyze L2 inflectional variability, reliance on either perception or production data only is not sufficient. In fact, researchers must examine both modes of performance to ensure accurate measures of L2 inflectional knowledge. This result contradicts Brown (1998), who claimed that perception is a more reliable test of grammatical knowledge.

### 6.2.3 Processing Load Effects

As hypothesized, accuracy rates of target L2 inflections were higher on items that required lower degrees of processing resources (e.g., shorter, less complex sentences) than on items that required higher degrees of processing resources (e.g., longer, more complex sentences) within the same task. Processing load effects (i.e., due to increased processing pressure) that revealed significant differences in accuracy between shorter and longer sentences were observed in the sentence perception task for both the real and nonce items, and for the nonce items in the sentence production task. However, there was no significant difference in accuracy of the longer and shorter real items in the OSP task. This result is not surprising considering the high rates of suppliance that were observed in the real items within the OSP task for both the lower and higher proficiency levels.

Thus, it can be concluded that sentence length leading to processing load effects on accuracy of target L2 inflection had a significant effect on the perception of the target inflections in real and nonce items within sentences and on the suppliance of the target inflections within nonce items in sentences. The fact that accuracy of the nonce target items in the longer sentences was significantly lower supports the processing load claim in production even though the subjects produced L2 inflection on real items at a high level. Additionally, even though the subjects' oral production accuracy in sentences was high overall (>85% suppliance), their suppliance of target inflections on the sentence production task was lower than it was on the word production task indicating an effect of performance pressure that was likely related to processing load effects due to the higher task demands of sentence tasks over word tasks. Thus, similar to the results in Hopp (2009), Gordon et al., (2002), and Trenkic (2007), a significant

effect of processing load was observed, but this effect was minimized on the real items because these subjects were at an advanced level of L2 proficiency.

The fact that real items did not reveal processing load effects is explainable within the proposed effects of this study. To reiterate, it was proposed that even though variability of L2 inflection would occur for these L2 learners, native-like attainment is possible for L2 inflection; as such, it was hypothesized that as proficiency level increases, L2 learners will produce target inflections at or near-ceiling. In fact, it appears that these subjects are in fact approaching native-like attainment and are, therefore, revealing non-significant differences with regard to suppliance even under increased processing loads for real items in production (following Davidson, 2005, 2006a, 2006b and Hopp, 2009).

### 6.3 Repair Strategies in Oral Tasks

A broad acoustic analysis<sup>28</sup> of the subjects' oral recordings of target item repair strategies revealed some important outcomes that will likely further advance the debate regarding L2 inflectional errors. Following Davidson (2005, 2006a, 2006b), it was proposed that word-final coda repairs such as schwa epenthesis could occur in variation with native-like sequences in the interlanguage of the L2 learner leading to output that is native-like, but is still variably produced. In fact, this researcher observes this exact phenomenon in the results—that is, L2 inflection is produced at native-like levels, yet variable, but systematic repairs (e.g., devoicing and schwa epenthesis) are also observed in production. Importantly, the repairs are systematic because they are significantly more prominent on specific L2 codas and inflections (e.g., especially non-syllabic and arguably least salient [t], but also more prominent on less sonorous stops over

---

<sup>28</sup> This analysis was conducted via a spectrographic analysis in Praat to determine if the final coda clusters containing the target inflections would reveal repairs such as final C devoicing or schwa epenthesis.

sibilants) and are more significantly produced under controlled performance or processing conditions (e.g., more significantly in longer sentences). As such, these systematic asymmetrical patterns in perception and production are not supportive of L2 representational deficits. In addition, as will be discussed below in section 6.4, repairs were also significantly more prominent for the lower proficiency level subjects than for the higher subjects, which again provides evidence that supports Davidson's claims that these repairs are phonetically motivated.

### 6.3.1 Devoicing Word-Finally: Evidence of Markedness Effects

In both the OWP word production and OSP sentence production tasks, there was a high degree of devoicing of word-final consonants (i.e., [d] and [z]) in both the syllabic and non-syllabic voiced inflections (i.e., [Vd], [Vz], [d], [z]), an outcome that appears consistent with the markedness effects investigated by Broselow (2004) and Broselow et al. (1998). This effect was more pronounced in the OWP task with the percentage of devoicing of final consonants ranging from 58% ([d]) to 68% ([Vd]) for the higher proficiency group and 47% ([d]) to 90% ([Vz]) for the lower proficiency group. In the OSP task, the higher proficiency group devoiced final consonants from 14% ([Vd]) to 32% ([d]) and the lower proficiency group devoiced final consonants 32% ([d]) to 61% ([Vz]). This result is important because it appears that even though these Mandarin speakers were not omitting L2 inflections in oral production with any significance, their interlanguage grammars were outputting surface inflections that have been phonetically modified to reflect universal constraints that do not reflect the grammar of the L1 or the L2. For example, *cleans* typically was produced as [kli<sup>v</sup>ns] with the expected word-final consonant allomorph [z] being produced without voicing. Because Mandarin disallows word-final consonants and does not have voice contrasts in its phonology, and English voicing

assimilation rules require this final [z] to be voiced in the surface form, the devoicing results are not consistent with either the L1 or L2. Following Broselow and Broselow et al., who employed an Optimality Theoretic approach, which claims that all constraints of the grammar are available to the L2 learner regardless of the L1 grammar, it would appear that the L2 grammars of the Mandarin speakers in this study have reorganized the constraints within their grammars to attempt to output the L2 surface inflections. However, as evidenced by the high degree of devoicing, an OT account would claim that markedness constraints against voiced final consonants are ranked above other constraints that allow for surface voicing within the coda.<sup>29</sup> As a result, if L2 inflection acquisition has occurred, as evidenced by the high accuracy of suppliance of the target inflections, it may be the case that production of the word-final segments is being variably influenced by universal phonological constraints that are not consistent with the L1 or L2 but that are, consequently, higher ranked within the grammar.

A full OT analysis is beyond the scope of this dissertation, but the devoicing result clearly suggests that restructuring of the grammar is likely occurring, and the degree to which devoicing occurs depends on the interaction between the type of L2 inflection (i.e., [Vd] vs. [d]) and upon the proficiency level of the subject. In sum, the devoicing of L2 inflections is the result of interlanguage grammars that are outputting the L2 inflections at a native-like level as evidenced by the high rates of suppliance (i.e., no representational deficits), but that are still reflective of variability within the grammar that systematically affects the production of the L2 surface inflections depending on the type of coda or inflection. It is also important to note that the devoicing rates for the OSP sentence production task were lower than for the OWP word production task. This reduced devoicing effect in the sentences is a result of the fact that the

---

<sup>29</sup> A full OT analysis is beyond the scope of this study, but it is important to note that within an OT approach, all L1 and L2 constraints are available in the grammar (i.e., ultimate attainment is possible) even when the ranking of constraints outputs non-native surface forms.

subjects were able to carry the voicing from the word-final segment in the target verbs over to the vowel in the following word, regularly resyllabifying the verb-final consonant to the onset of the following word which always began with a vowel (e.g., [kli<sup>v</sup>n.zIt] or *cleans it*). Nevertheless, the pattern of asymmetry of devoicing still occurred at different rates on the syllabic and non-syllabic items just as in the OWP task. In addition, this effect was again more pronounced in the nonce items than in the real items and was productive in both the word and sentence tasks, suggesting that it is not a lexicalized (i.e., memorized) effect, but that it is productive at the segmental level (cf. Brown, 1998; Kabak & Idsardi, 2007).

### 6.3.2 Schwa Epenthesis: Evidence Against a Strict Prosodic Transfer Hypothesis

In addition to the devoicing of final consonants within the target inflections, although occurring much less frequently than the devoicing, schwa epenthesis was also employed in some of the target codas and inflections to break up the word-final consonant clusters. With regard to the tabulations, epenthesis that occurred either between the word-final root consonant and word-final inflectional consonant (e.g., *walked* = [wɔ.kə.tə]) or after the word-final inflectional consonant (e.g., *caused* = [kɔz.də]) was included. Both of these errors were considered a schwa epenthesis error in this dissertation,<sup>30</sup> as they both exhibited an attempt to preserve the consonants in the coda cluster via resyllabification. However, they were not considered errors of deletion because the required target inflection was clearly supplied.

---

<sup>30</sup> Based on personal communication with Elaine Klein, she believes it is important to make clarify what was an error of schwa epenthesis because it may just be in her words, “a mis-pronunciation.” However, first, the determination as to the nature of the schwa insertion at the surface or the phonological level is not a main question of this research—it is only what surfaces in the phonetic form and how this affects surface variability with regard to word-final clusters that is under examination. Further, within the sentence production task (OSP), the subjects had to conjugate the verb from its infinitive form (*to walk*), and thus, it is assumed in this task that any epenthesis was inserted at the phonological level. In either case, word-final clusters were analyzed at the phonetic syllable level within Praat and any non-target vowel insertion within the word-final cluster was, therefore, included in the error rates for schwa epenthesis.

In the OWP word production task, the schwa epenthesis predominantly occurred within the items that contained the [t] codas (lower group = 26%; higher group = 30%), and a significant difference due to inflection type was observed across the inflections,  $F(5, 18)=3.14$ ,  $p<.05$ . However, the pattern of schwa epenthesis reveals that this repair was most productive on the items that contained word-final stops. Within the OSP sentence production task, there was also a higher degree of schwa epenthesis such that there was a significant effect of inflection type on the degree of schwa epenthesis,  $F(5, 18)=10.97$ ,  $p<.001$ , with the rates of repair again occurring most on the [t] allomorphs and much higher for the lower group (lower group = 39%; higher group = 6%). This repair was also high for the [d] allomorph for both groups (lower group = 29%; higher group = 25%) and to a lesser degree on the [s] allomorph (lower group = 16%; higher group = 14%) and [Vd] allomorph (lower group = 14%; higher group = 18%). The other sibilant-coda allomorphs ([z] and [Vz]) in the sentence task exhibited 3% and 4% of schwa epenthesis, respectively, with the groups collapsed, indicating that the non-syllabic stops ([t] and [d]) were the most effected.

In general, the subjects' propensity to re-syllabify the word-final coda clusters across all of the inflections was low. In fact, because the rate of schwa epenthesis systematically affects certain allomorphs (i.e., [t] and [d]) significantly more than others (i.e., [s] and [z]), and because target-like suppliance of L2 inflections has already been established with regard to the oral production rates in the sentence task, the pattern of schwa epenthesis within the oral tasks is clearly not supportive of a representational deficit in the L2 prosodic structure, as is posited by the PTH. To reiterate, the PTH claims that because the L1 prosodic structure cannot accommodate the L2 tense morphology that requires that L2 inflection be attached at the PWD, resyllabification of the L2 inflection must occur to accommodate the L2 inflection into the L1

prosodic structure. However, this claim would only be supportable if the effect was consistently applied across all L2 inflections. In fact, the subjects in this dissertation did not produce this result.

On the contrary, the subjects repaired certain L2 allomorphs significantly more than others such that the items that contained the stops [t] and [d] were repaired significantly more than items that contained the sibilants [s] and [z]. In fact, repair via schwa epenthesis for the sibilants with [Vz] and [z] allomorphs occurred at rates of 5% or below. Not only does this repair pattern after the hypothesized sonority effect as proposed by this researcher, but also the rate of repair increased on the nonce items with the same pattern (i.e., stops were significantly more repaired than sibilants). Hence, the low incidence of syllable repair alone is not enough to contradict the PTH; however, as this researcher concludes, the rates of repair are clearly patterning after both effects of sonority, proficiency level, and in the case of the nonce items, increased processing pressure. Thus, this researcher concludes that the PTH is not predictive of variability of L2 inflections in these data.

Additionally, although the rates of schwa epenthesis do not strongly support an across the board L1 phonotactic interference effect on all illicit codas, the type of asymmetrical schwa epenthesis repair is attested in the literature by several researchers including Davidson (2005, 2006a, 2006b). Following an OT approach, she suggests that variable rankings of phonotactic constraints along with other phonetic constraints within the grammar result in variability of the output, which is supportive of the conclusion that asymmetrical surface variability of inflections can be systematically related to performance and phonetic-acoustic effects rather than to permanent representational deficits such as hypothesized under the PTH.

When examining the items in which both schwa epenthesis plus devoicing of word-final consonants occurred (cf. Tables 22 and 23 vs. 20 and 21), rates of repair again were highest on the non-syllabic items [d] and [z], reaching significance in both the OWP word production and OSP sentence production tasks, suggesting that syllabicity had an across the board effect on these items with items that were less salient being produced with greater repair. These results are again supportive of the fact that repair was not due to morphosyntactic or prosodic deficits in production, but that they are due to non-representational factors that affect parsing at the acoustic level (as suggested by Davidson, 2005, 2006a, 2006b and Solt et al., 2004) or at the segmental-contact level (as suggested by Goldrick (2004) and Kabak & Idsardi (2007)). In fact, the rates of repair pattern systematically after the hypothesized effects of L1 phonotactic interference (in terms of repairs), salience (in terms of syllabicity and sonority), processing pressure, and, to some degree, proficiency level, although the obvious advanced level of these subjects mitigated some of the proficiency level effects in production.

Finally, in an analysis of monomorphemic items in the OWP word production task, it was observed that items with stops (e.g., *mitt*, *mid*, etc.) and fricatives (e.g., *price*, *prize*, etc.) in the coda were devoiced at equally high rates as the real and nonce items that contained the target codas, providing further support that devoicing was productive at the level of the coda in the interlanguage and was blind to the type of target item (i.e., inflected, verbal items vs. uninflected, non-verbal items). Importantly, the rates of schwa epenthesis reflected similar results when compared to the real and nonce OWP items such that stops revealed higher rates of repair than fricatives. In addition, when examining the combination of devoicing plus schwa epenthesis, the rates of repair again increased significantly, revealing that stops were again more significantly repaired than fricatives. When the *schwa epenthesis only* and *devoicing plus schwa epenthesis*

means were combined, there was clearly a pronounced effect of repair on stops over fricatives, as was observed in the real and nonce items. Thus, if the patterns of repairs only targeted inflected items as would be predicted under the PTH, these repairs would not have occurred in the same asymmetrical but systematic patterns in the monomorphemic items as in the real and nonce verbal items. In fact, as will be discussed in Section 6.4 below, the patterns of repair in the stops and fricatives in the real and nonce verbal items showed significant correlations to the patterns of repair in the monomorphemic items. As such, the patterns of repair provide further support against a L2 prosodic deficit for these L2 learners.

#### 6.4 Proficiency Level Effects

There appeared to be a definitive effect of proficiency level within the ASP sentence perception task such that the lower proficiency level group revealed significantly lower accuracy of the target inflections, but only on the *different*, real items. *Same* and nonce items did not reveal any effect of proficiency level on accuracy. Additionally, this effect was clearly prominent with regard to the short, *different* sentences in which the higher group performed more accurately than the lower group. Although there were clear processing load effects, the effect of proficiency level was not definitive between the groups on the longer, *different* sentences, which appeared to affect the accuracy rates of the two groups similarly such that they both performed with similar reduced accuracy. There was no effect of proficiency level on oral productive accuracy in terms of deletions, which was unexpected. This result may be due to the fact that these two groups were fairly close in proficiency level as they were only 2 years apart in college (i.e., freshmen vs. juniors), and the expected level of English proficiency to enter university in China is relatively high. The effect of proficiency level would likely reveal a much greater and significant

difference between L2 learners who had not yet taken the college English entrance exam (e.g., juniors in Chinese high schools) vs. English majors who were already in their junior year of college in China. Thus, although omission of L2 inflections was not significant for either of these subject groups, it was concluded that the advanced suppliance of L2 inflections is supportive of the fact that ultimate attainment of L2 inflections is possible even when the L2 features are not instantiated in the L1 grammar. This result is supportive of the results and conclusions in Hopp (2009) and Trenkic (2007) in which they suggest that performance effects are mitigated by increased proficiency.

Although proficiency level effects did not reach significance with regard to omissions of L2 inflections in the OSP task, there were several significant interaction effects observed between proficiency level and the degree of repair in the OSP task. Again, as concluded, these L2 learners were advanced in their knowledge of L2 inflections, so both groups revealed no proficiency level differences in the suppliance of inflections. Only specific effects of target coda repairs on some of the target inflections were significantly different between the two proficiency levels. With regard to devoicing, the [Vd], [Vz], and [z] real and nonce items were significantly more devoiced for the lower-level group than for the higher-level group. The devoicing rates for the two proficiency levels for [d] allomorphs in real and nonce items were nearly the same and did not show a significant difference between the proficiency levels. With regard to schwa epenthesis, only the [t] allomorphs showed significant effects of proficiency level with the lower-level subjects repairing word-final syllables with schwa epenthesis significantly more in these items. However, in the nonce items, all of the non-syllabic allomorphs ([t], [d], [s], and [z]) were repaired via schwa epenthesis significantly more for the lower-level subjects than for the higher-level subjects. Thus, although, an across the board main effect of proficiency level was not

observed with regard to target coda repair on all of the target items, lower proficiency level subjects clearly repaired target codas at a higher level. Thus, it can be concluded that proficiency level has a greater effect on word-final syllable repair depending on the target inflection. Importantly, because the higher-level subjects showed significantly less repair of L2 inflections than the lower-level subjects, and schwa epenthesis was not consistently applied across the L2 codas (e.g., epenthesis was much lower on sibilants than on stops), it can also be concluded that these repairs are not mediated by a deficit of L2 inflections or prosodic representation, but they are consistent with accounts that maintain that ultimate attainment of L2 inflections for adult L2 learners is possible.

#### 6.5 Correlations between Perception and Production

Finally, it was hypothesized that the subjects' accuracy of perception would correlate with their accuracy of production (cf. Flege, 1995; Jia et al., 2006). The accuracy of perception and accuracy of production were analyzed using the following three comparisons:

- 1) **Different** items (e.g., “walk/walks”) in Perception to **Suppliance** in Production (e.g., “She walks on the sidewalk.”)
- 2) **Different, Voiced** items (e.g., bleeds/bleed) in Perception to **Devoicing** in Production (e.g., “bleeds” [bliyds])
- 3) **Different** items (e.g., walk/walks) in Perception to **Schwa Epenthesis** in Production (e.g., “walkse’ [wɔk.sə])

To reiterate, in the first comparison, as accuracy of perception increases, an increase in accuracy of production should be observed such that as perception increases, suppliance of target items should also reveal an increase in accuracy. A positive correlation in this first comparison

would suggest that perception and production of L2 inflection increase together incrementally, thus, providing support for a positive relationship between these two aspects of performance with regard to L2 inflectional acquisition.

In the second and third comparisons, decreased accuracy in perception should reveal an increase in repairs. These analyses looked at reduced perception accuracy to reduced accuracy in target production of the codas—in other words, with regard to repairs, reduced accuracy in production was indicated by a higher rate of repairs, either devoicing or schwa epenthesis within the target coda clusters. Thus, these analyses were also analyzed as positive correlations within SPSS with decreased accuracy in perception compared to decreased accuracy (i.e., increased repair rates) in production.

Of the above correlations, only accuracy rates in the *different* items in the ASP sentence perception task compared to schwa epenthesis rates in the OSP sentence production task produced a significant result. When the accuracy of the *different* items in the sentence perception task was correlated to the percentage of schwa epenthesis in the sentence production task, a significant correlation was found:  $r=.42, p<.05$  (Sig. 2-tailed). Thus, when accuracy of aural perception of L2 inflections increased, accuracy of their oral production increased. In other words, in the case of schwa epenthesis, the subjects repaired via schwa epenthesis to a higher degree in oral production when their aural perception decreased. However, when comparing like inflections in perception to like inflections in production, a significant correlation was found only between the inflections that ended in the [t] allomorph. Considering the high proficiency level of the subjects in this study, the fact that only the [t] allomorph reached significance is not surprising. Although the accuracy rates of the [d] allomorphs in the sentence perception task compared to the degree of schwa epenthesis in the sentence production task did not reach

significance, there was a clear pattern of reduced perception to increased schwa epenthesis with regard to the allomorph [t] as well.

In fact, the rates of mis-perception and of schwa epenthesis of the allomorph [d] in the sentence tasks patterned similarly to the [t] allomorph, especially for the nonce items and the lower proficiency level subjects, such that as perception accuracy decreased, the schwa epenthesis repair increased. Again, significant correlations were not found across both the allomorphs [t] and [d]; nevertheless, this result is important to consider within this dissertation as it strongly points to a relationship between perception and production for L2 inflections. Thus, following Brown (1998), Hopp (2009), and Jia et al., (2006), this researcher concludes that analysis of perception (i.e., comprehension) of L2 inflections is critical to ground any study of variability of L2 inflections in production. As such, this finding provides further evidence that spontaneous production alone cannot be used as the sole source of data to make claims about L2 grammatical representation. Additionally, although the results of this study did not produce across the board correlations between perception and production, considering the high level of accuracy in the word tasks and oral sentence production, the lack of significant positive correlations was not surprising. Further comparisons of this type with lower level subjects are warranted to determine the extent to which perception and production are related.

## 6.6 Correlations between Inflected and Uninflected Items in Production

Because similar patterns of repair were observed not only in the real and nonce items, but also in similar word-final consonants in the monomorphemic items, correlations were run to determine the degree to which these repairs would pattern together. This analysis was conducted to determine the degree to which the repairs in oral production for the real and nonce items were

productive to uninflected items, which would indicate that the repairs were independent of L2 inflection and modulated by surface conditions such as coda type. First, a significant positive correlation ( $r=.53, p=.007$ , Sig. 2-tailed) was observed with regard to devoicing between the [Vd] codas in the real verbal items and the voiced stops in the monomorphemic items. A significant correlation ( $r=.78, p=.001$ , Sig. 2-tailed) was also observed with regard to devoicing between the [d] codas in the real verbal items and the voiced stops in the monomorphemic items. Thus, the devoicing repair occurs on codas independent of L2 inflections such that rates of devoicing repair patterned together in both the real verbal items and non-verbal monomorphemic items. Significant positive correlations were also observed between the nonce verbal items [Vd] ( $r=.68, p=.001$ , Sig. 2-tailed) and [d] ( $r=.60, p=.002$ , Sig. 2-tailed) when compared to the voiced monomorphemic stops. No significant correlations were found when comparing the sibilants [Vz] and [z] to the voiced fricatives monomorphemic items. Thus, this result is further supportive of a lack of overall effect of repair across all items. Instead, it supports this researcher's previous conclusion that sonority of the word-final coda appears to have a significant effect on the production of these codas in both real/nonce verbal items as well as in monomorphemic items whose codas mirror the target inflections under investigation.

Finally, correlations between rates of schwa epenthesis of real and nonce verbal target items and non-verbal monomorphemic stops revealed significant positive correlations between [Vd] codas in nonce items and monomorphemic items that end in stops ( $r=.59, p=.002$ , Sig. 2-tailed) and between [d] codas in real items and monomorphemic items that end in stops ( $r=.59, p=.003$ , Sig. 2-tailed). There were no significant correlations for comparisons with [t], [Vz], [s], or [z] and monomorphemic items, as the rates of schwa epenthesis were generally pretty low overall on these items in the monomorphemic items. In some cases, there was no repair at all on

the items. As previously noted, this researcher claims that the advanced proficiency level of the subjects caused a low degree of deletion and repair overall in production. Nevertheless, the fact that schwa epenthesis is observed only between the word-final stops in real and nonce items when compared to the word-final stops in the monomorphemic items, and not between the word-final sibilants [Vz], [s], and [z] and monomorphemic items that contain word-final fricatives, again provides evidence for a phonetic-acoustic effect on production of specific target codas that is not reflective of a prosodic deficit as claimed under the PTH.

Overall, with regard to the debate over the causes of variability of L2 inflections, the asymmetrical, yet systematic application of repair across both target L2 inflections and monomorphemic items is in conflict with the PTH claim that Mandarin speakers lack the prosodic representation for L2 inflections.

## Chapter 7: Conclusions and Limitations

### 7.1 Conclusions

This dissertation examined the perception and production of L2 English inflectional endings in learners who exhibited knowledge of the morphosyntactic properties of such items. It was observed that for these learners, the asymmetries in perception and production are consistently applied with regard to specific codas in real and nonce items (e.g., more to [t] than [s]); and the variability of the target codas and L2 inflections can be systematically linked to differing degrees of salience (i.e., syllabicity and sonority) and processing load effects. In addition, because the same repairs are systematically applied to monomorphemic items that are uninflected (e.g., “mist” [mɪs.tə]), but that mirror the target verbal items (e.g., “missed” [mɪs.tə]), this asymmetrical, but systematic application of repairs to non-verbal items provides direct evidence against the application of the Prosodic Transfer Hypothesis (PTH) to these data. In general, there was a significant degree of error on codas that were less acoustically perceptible [t] and [d] than on codas that were more acoustically perceptible [Vd], [Vz], and [s]. This pattern of lower accuracy appeared to be systematic across both perception and production and word and sentence tasks, and as such, the patterns are considered by this research to be systematically applied; however, they are *asymmetrical* in their application (i.e., applied significantly more to codas that contained stops than to codas that contained sibilants).

As such, this dissertation has illustrated that performance effects must be systematically considered and controlled in any analysis of L2 inflectional variability, and that patterns of error must be examined for their application to specific codas and allomorphs to determine what other non-morphosyntactic effects may be observable. Thus, because the effects in the results of this

study revealed these systematic, but asymmetrical patterns, this researcher proposes the Performance Asymmetries Hypothesis (PAH), which can be defined as follows:

L2 learners who are found to be advanced in their L2 morphosyntactic knowledge will continue to perceive and produce L2 inflections with systematic but asymmetrical variability due to the following factors, irrespective of the L1:

- 1) Salience or acoustic perceptibility of the L2 inflections due to syllabicity and sonority effects.
- 2) Processing pressure caused by varying degrees of target item length and complexity.

The PAH predicts that if a L2 learner reaches an advanced level of proficiency in the target language, persistent errors in performance of L2 inflections will reflect the above salience and processing pressure effects and that these effects will pattern similarly across the perception and production modes. Thus, the interlanguage is predicted to reflect not an across the board reduction in accuracy as would be observed if the L2 representation is in deficit, but it is predicted to reflect systematic, but asymmetrical variability that reflects specific effects of syllabicity, sonority, and processing pressure. As is posited by this researcher, the PAH also assumes that access to UG is possible for adult L2 learners, and therefore, native-like attainment of the L2 grammar is possible.

In addition, considering the variable effects that were observed across perception and production, a methodological concern that must be considered when examining L2 competence and L2 inflectional deficits is that we must systematically test both perception and production. Thus, any analysis of L2 inflectional variation in production must also include a similar analysis

of perception of the same L2 inflections to ensure that we are observing the full picture of a learner's competence and performance.

This study has systematically examined the performance variability of L2 inflectional morphology for adult L2 learners. It was found that even when L2 features such as tense and agreement are not instantiated within a L2 learner's L1 syntax, these L2 features can be acquired, as attested by target-like accuracy on grammaticality judgment tasks and even within oral sentence production tasks. However, even when suppliance of L2 inflections is advanced, L2 learners will likely continue to repair word-final consonant clusters by devoicing target codas or applying schwa epenthesis as a repair strategies that affect certain allomorphs more than others.

Furthermore, the results of this study have shown that non-syntactic or non-prosodic factors such as syllabicity and acoustic features such as sonority clearly affect accuracy of perception and production in a systematic way. Thus, representational deficit claims that are based solely on the variability of surface inflections for L2 learners who reveal target-like knowledge of morphosyntax must be reexamined to control for these performance effects. In sum, these factors must be considered in future studies of inflectional variability.

Finally, although this study did not directly examine mapping deficits as a possible effect on inflectional omission, it can be concluded from the high accuracy of suppliance of target inflections within the oral sentence task that mapping deficits is unlikely causing reduced performance for these L2 learners. Additionally, considering the effects of the different factors on inflectional variability within this study, the Missing Surface Inflection Hypothesis (MSIH) of Prévost and White (2000) and the Feature re-Assembly Hypothesis (FAH) of Lardiere (1998a, 1998b, 2003) should be reexamined in light of this new information regarding asymmetries in surface inflectional variation. It should also be reiterated that that predictions of the

Representational Deficit Hypothesis (RDH) of Hawkins and Liszka (2003) and the Prosodic Transfer Hypothesis (PTH) of Goad, White, and Steele (2003) are not borne out in the performance of these subjects and should, therefore, be reexamined in the light of the results of this study.

## 7.2 Limitations

This study could be extended as follows:

1) The lower group of L2 learners in this study appears to be at a higher level than expected. It would be beneficial to examine more distinct proficiency levels in order to determine if and to what extent L1 phonotactic constraints and the other factors under study in this dissertation have an effect on deletion of L2 inflections. At lower levels of proficiency, it is likely that L1 phonotactic constraints have a much greater effect on word-final codas both in terms of deletion of the codas and repair of the word-final clusters. It would also be beneficial to examine other languages that contain other types of interfering and non-interfering L1 syllable constraints for comparison to these data. It would be assumed under the PAH proposed by this researcher that even if the L1 syllable constraints do not disallow the L2 target codas, there should still be measurable effects on L2 inflections of syllabicity, sonority, and processing pressure especially at lower L2 proficiency levels.

2) The items in this study were balanced in terms of the surface complexity of the coda clusters that contained the target inflections (i.e., in terms of number of and type of word-final consonants), and the global effect of stops over fricatives was shown to be significant. However, because it was beyond the scope of this study, the specific (i.e., more local) effects of discrete acoustic interactions between the consonants in the word-final the clusters or the effects of

preceding vowel length before the coda clusters on perception and production accuracy were not examined. This additional analysis of segmental-level acoustic interaction would provide more acoustic-phonetic depth that may reveal even more specific or local patterns of word-final repair.

3) A final limitation and one that must be considered in similar studies is to also test these same subjects using spontaneous tasks such as those used by Hawkins and Liszka and Goad, White, and Steele. The purpose of adding such tasks to the analysis is that the subjects would likely reveal the same asymmetrical pattern of performance and salience effects observed in this dissertation, but to a significantly higher degree. This result would further support the application of the PAH to these data and provide additional evidence against claims of representational deficits as the cause of L2 inflectional variability for advanced L2 learners.

## Appendix A: Data Tables from Pilot Study

**Table A1: Overall performance on WGJ Task - mean percentage of correct inflections by group**

	Mean % correct
NS (n=2)	100%
NNS (n=10)	96.0%

**Table A2: WGJ Task – number and percent correct by verb type and participant**

# of Type	12	12	12	12	12	12	6	6	6	6	
Part	3sg	3pl	[t]	[d]	[Id]	irreg	is	are	was	were	Total % Correct
<b>FN</b>	12	11	12	12	12	12	6	5	6	6	
	100	91.7	100	100	100	100	100	83.3	100	100	97.5%
<b>HW</b>	12	12	12	12	12	11	6	5	6	6	
	100	100	100	100	100	91.7	100	83.3	100	100	97.5%
<b>PY</b>	12	11	11	12	11	12	6	5	6	6	
	100	91.7	91.7	100	91.7	100	100	83.3	100	100	95.8%
<b>RT</b>	10	11	12	12	11	12	6	4	6	6	
	83.3	91.7	100	100	91.7	100	100	66.7	100	100	93.3%
<b>CC</b>	11	11	12	12	11	12	6	4	6	6	
	91.7	91.7	100	100	91.7	100	100	66.7	100	100	94.2%
<b>CL</b>	12	12	12	12	12	12	6	5	6	6	
	100	100	100	100	100	100	100	83.3	100	100	98.3%
<b>JM</b>	12	11	12	12	11	12	6	4	6	6	
	100	91.7	100	100	91.7	100	100	66.7	100	100	95.0%
<b>JY</b>	12	9	12	12	12	11	6	5	6	6	
	100	75	100	100	100	91.7	100	83.3	100	100	95.0%
<b>XQ</b>	12	12	12	12	12	12	6	4	6	6	
	100	100	100	100	100	100	100	66.7	100	100	96.7%
<b>XX</b>	12	12	12	12	12	12	6	5	6	5	
	100	100	100	100	100	100	100	83.3	100	83.3	96.7%

**Table A3: Aural Task – number and percent correct by verb type and participant**

# of Type	6	6	6	6	6	6	6	6	6	6	
Part	3sg	3pl	[t]	[d]	[Id]	irreg	is	are	was	were	Total % Correct
<b>FN</b>	6	5	5	5	5	6	5	6	4	6	
	100	83.3	83.3	83.3	83.3	100	83.3	100	66.7	100	88.3%
<b>HW</b>	5	6	4	6	6	6	5	5	5	5	
	83.3	100	66.7	100	100	100	83.3	83.3	83.3	83.3	88.3%
<b>PY</b>	6	5	2	4	6	6	4	6	4	5	
	100	83.3	33.3	66.7	100	100	66.7	100	66.7	83.3	80.0%
<b>RT</b>	5	3	4	6	6	5	5	5	4	5	
	83.3	50	66.7	100	100	83.3	83.3	83.3	66.7	83.3	80.0%
<b>CC</b>	5	6	2	3	5	5	6	5	4	3	
	83.3	100	33.3	50	83.3	83.3	100	83.3	66.7	50	73.3%
<b>CL</b>	6	5	5	6	6	5	6	6	5	5	
	100	83.3	83.3	100	100	83.3	100	100	83.3	83.3	91.7%
<b>JM</b>	6	5	3	6	6	5	6	5	5	5	
	100	83.3	50	100	100	83.3	100	83.3	83.3	83.3	86.7%
<b>JY</b>	6	4	2	6	6	5	5	5	6	3	
	100	66.7	33.3	100	100	83.3	83.3	83.3	100	50	80.0%
<b>XQ</b>	6	4	3	4	5	5	5	6	5	6	
	100	66.7	50	66.7	83.3	83.3	83.3	100	83.3	100	81.7%
<b>XX</b>	6	2	5	4	6	6	4	5	4	4	
	100	33.3	83.3	66.7	100	100	66.7	83.3	66.7	66.7	76.7%

**Table A4: SPEAK Task – number and percent correct by type and participant**

# of type variable												Total % Correct
Part	3sg	3pl	[t]	[d]	[Id]	irreg	is	are	was	were		
<b>FN</b>	4/4	3/4	N/A	1/2	1/2	2/4	10/10	0/1	0/2	1/1		
	100	75	N/A	50	50	50	100	0	0	100		58.3%
<b>HW</b>	3/3	1/1	N/A	N/A	N/A	0/6	12/14	1/2	0/2	1/1		
	100	100	N/A	N/A	N/A	0	85.7	50	0	100		62.2%
<b>PY</b>	3/4	4/5	0/1	0/1	0/1	3/3	25/26	2/4	0/5	0/1		
	75	80	0	0	0	100	96.2	50	0	0		40.1%
<b>RT</b>	14/18	1/6	N/A	N/A	N/A	N/A	22/22	0/8	N/A	N/A		
	77.8	16.7	N/A	N/A	N/A	N/A	100	0	N/A	N/A		48.6%
<b>CC</b>	3/10	3/3	1/2	0/1	1/1	6/8	23/24	1/3	2/4	N/A		
	30	100	50	0	100	75	95.8	33.3	50	N/A		59.3%
<b>CL</b>	4/9	9/11	3/3	3/3	1/1	1/2	30/30	3/4	2/2	1/1		
	44.4	81.8	100	100	100	50	100	75	100	100		85.1%
<b>JM</b>	1/2	0/1	1/2	0/1	N/A	3/9	18/20	6/6	2/3	N/A		
	50	0	50	0	N/A	33.3	90	100	66.7	N/A		48.8%
<b>JY</b>	4/6	3/5	1/1	1/1	1/1	2/2	12/13	1/1	0/2	0/1		
	66.7	60	100	100	100	100	92.3	100	0	0		71.9%
<b>XQ</b>	4/4	7/8	N/A	N/A	N/A	N/A	10/12	4/5	N/A	0/1		
	100	87.5	N/A	N/A	N/A	N/A	83.3	80	N/A	0		70.2%
<b>XX</b>	2/4	1/3	N/A	2/3	N/A	4/7	23/25	2/4	N/A	N/A		
	50	33.3	N/A	66.7	N/A	57.1	92	50	N/A	N/A		58.2%

## Appendix B: Pilot Test Items

### B1: Written Grammaticality Judgment Task

Name: \_\_\_\_\_ Date: \_\_\_\_\_ Time to complete: \_\_\_\_\_

In order to learn more about your grammatical ability, I'd like you to complete the following short task. Please answer each item to your best ability.

**Part A: Circle the correct verb in the parentheses. There is only one correct answer for each question – no other information is needed to complete the sentence:**

Example A. When the man (**makes** / **made**) cake, he always forgets to add the sugar.

Example B. Last week, she (**has** / **had**) lots of homework to do.

- 1) He (**looked** / **look**) at the sky when the sun was setting.
- 2) The boy always (**asks** / **ask**) for a cookie after dinner.
- 3) The man (**danced** / **dance**) at the party last night.
- 4) The children always (**drink** / **drinks**) all of their milk.
- 5) The police officer (**are** / **is**) extremely angry.
- 6) Before he left for vacation, he (**were** / **was**) very stressed.
- 7) When he exited the building, he (**know** / **knew**) which way to turn.
- 8) The mechanics (**fixes** / **fix**) cars very well.
- 9) He (**missed** / **miss**) an appointment with his advisor yesterday.
- 10) I read my Biology book while my roommate (**ate** / **eat**) dinner.
- 11) In the early 1900's, cowboys (**was** / **were**) everywhere in the western United States.
- 12) They (**plan** / **planned**) the party last week.
- 13) The student (**are** / **is**) in school all day today.
- 14) They always (**thinks** / **think**) about their families when they are at school.
- 15) He usually (**buys** / **buy**) two used cars for the price of one.
- 16) While walking to school yesterday morning, she (**buy** / **bought**) a coffee.
- 17) The women (**is** / **are**) in the store.
- 18) He (**planted** / **plant**) two beautiful trees in his back yard yesterday.
- 19) I saw the bird as it (**soared** / **soar**) across the sky this morning.
- 20) Two of my friends (**was** / **were**) at home when she called last night.
- 21) She (**loved** / **love**) baseball when she was a child.
- 22) John finally (**pass** / **passed**) his exam last week.
- 23) The boys always (**enjoys** / **enjoy**) their yearly trip to California.
- 24) Last night, someone (**was** / **were**) in my house and stole my computer.
- 25) The teachers (**is** / **are**) in the office waiting for the students right now.
- 26) Last night, she (**slept** / **sleep**) on my couch all night.
- 27) After he took his final exams, he (**rest** / **rested**) for a week.
- 28) They (**was** / **were**) happy that tomorrow was his birthday.
- 29) My brother (**writes** / **write**) to me every week from China.
- 30) The student (**finish** / **finished**) his writing assignment before he went to the party.
- 31) He always (**waste** / **wastes**) his money on useless items.
- 32) He (**cook** / **cooks**) breakfast every morning for his wife.
- 33) The painter (**paint** / **painted**) two bedrooms yesterday.

- 34) Yesterday morning, she (**throw / threw**) the ball over the fence.
- 35) The clerk (**change / changed**) the price of the milk this morning.
- 36) The law (**were / was**) passed by the government last week.
- 37) My advisor (**brought / bring**) the article to class last week.
- 38) The little girl (**tie / tied**) her shoes by herself yesterday.
- 39) His mother (**is sew / is sewing**) his pants right now.
- 40) While he was cooking last night, he (**taste / tasted**) all of the different foods.
- 41) The mechanics (**is / are**) always very polite to me.
- 42) When she had the flu, she (**coughed / cough**) all night.
- 43) The students always (**feels / feel**) bored in his class.
- 44) Everyone (**wait / waited**) in line for tickets yesterday morning.
- 45) The girls (**is / are**) happy today.
- 46) She (**start / started**) the car at 9am today.
- 47) They (**study / studied**) for the test last night.
- 48) The man (**built / build**) his house last year.
- 49) The students (**fly / flies**) to Europe every summer.
- 50) She (**sails / sail**) her boat in the ocean every weekend.

**Part B: Rewrite the correct form of the verb in the space provided for each sentence.**

Example A. Yesterday, her friend \_\_\_\_\_ left \_\_\_\_\_ (**to leave**) for vacation.

Example B. I \_\_\_\_\_ went \_\_\_\_\_ (**to go**) back to my home country yesterday.

- 1) Yesterday, they \_\_\_\_\_ (**to accept**) an invitation to the party.
- 2) My wife \_\_\_\_\_ (**to swim**) in our pool yesterday.
- 3) She \_\_\_\_\_ (**to search**) for a new apartment last week.
- 4) They always \_\_\_\_\_ (**to clean**) their house every day.
- 5) The children \_\_\_\_\_ (**to speak**) with their mothers several times per day.
- 6) The woman \_\_\_\_\_ (**to count**) her money in front of the cashier.
- 7) Last week, we \_\_\_\_\_ (**to watch**) a movie at work every day during lunch.
- 8) She \_\_\_\_\_ (**to be**) singing in the shower yesterday morning.
- 9) My friends \_\_\_\_\_ (**to talk**) to their fathers every night after class.
- 10) She is usually very hungry, so she \_\_\_\_\_ (**to spend**) 20 dollars per day on groceries.
- 11) Yesterday, the boy \_\_\_\_\_ (**to be**) at home all day.
- 12) She \_\_\_\_\_ (**to cancel**) her trip to the movies last night.
- 13) His professor \_\_\_\_\_ (**to want**) him to talk about his research yesterday.
- 14) The students always \_\_\_\_\_ (**to complain**) to their friends every day.
- 15) He usually \_\_\_\_\_ (**to smoke**) cigarettes when he is very stressed.
- 16) Last night, we \_\_\_\_\_ (**to listen**) closely to the new president speak about change.
- 17) They \_\_\_\_\_ (**to be**) in the pool all day yesterday.
- 18) That student \_\_\_\_\_ (**to lose**) his backpack and had to buy new books.
- 19) She \_\_\_\_\_ (**to be**) in the classroom at this very moment.
- 20) He \_\_\_\_\_ (**to want**) a job at my company last year.
- 21) After she graduated from college, she \_\_\_\_\_ (**to be**) very happy.
- 22) My friend \_\_\_\_\_ (**to fall**) on the ice and broke her leg.
- 23) It was very sad when they \_\_\_\_\_ (**to be**) very sick.
- 24) When the weather is bad, the woman usually \_\_\_\_\_ (**to drive**) to work.

- 25) She \_\_\_\_\_ (**to be**) with her advisor right now.
- 26) Yesterday, he \_\_\_\_\_ (**to notice**) her singing and loved it.
- 27) I \_\_\_\_\_ (**to complete**) the application last week.
- 28) Our friend \_\_\_\_\_ (**to pay**) for dinner last night.
- 29) The landlord \_\_\_\_\_ (**to fix**) my door but forgot to fix my window.
- 30) The girl \_\_\_\_\_ (**to be**) at home right now.
- 31) The wind \_\_\_\_\_ (**to blow**) strongly last night.
- 32) Now that they have jobs, the students \_\_\_\_\_ (**to be**) living on 25<sup>th</sup> Street.
- 33) It \_\_\_\_\_ (**to rain**) very hard yesterday.
- 34) Whenever they travel, they \_\_\_\_\_ (**to be**) never alone.
- 35) The employee \_\_\_\_\_ (**to pass**) his test and received the promotion last week.
- 36) She \_\_\_\_\_ (**to borrow**) her father's car last week.
- 37) The investigators always \_\_\_\_\_ (**to find**) the clues and solve the case.
- 38) The girl \_\_\_\_\_ (**to sing**) in the shower yesterday.
- 39) The employees \_\_\_\_\_ (**to be**) always on time for work.
- 40) She \_\_\_\_\_ (**to run**) every day after work to stay healthy.
- 41) Before the market crash, many people \_\_\_\_\_ (**to be**) very rich.
- 42) Yesterday, he \_\_\_\_\_ (**to miss**) his bus and had to wait for an hour.
- 43) The dog \_\_\_\_\_ (**to be**) very hungry and tired yesterday.
- 44) When she is busy, the woman typically \_\_\_\_\_ (**to work**) over 60 hours per week.
- 45) The stores \_\_\_\_\_ (**to wait**) until 9am to open this morning.
- 46) The government official \_\_\_\_\_ (**to allow**) his staff to take a vacation last week.
- 47) Usually, every student \_\_\_\_\_ (**to study**) at least 2 hours per day.
- 48) The leaders of the group \_\_\_\_\_ (**to be**) at the party right now.
- 49) She \_\_\_\_\_ (**to kiss**) her boyfriend in the store yesterday.
- 50) The parents always \_\_\_\_\_ (**to read**) to their daughter before they go to bed.

## B2: Aural Sentence Perception Task

- Yesterday, she accepts a job offer. (Incorrect) (II)  
 He paints his room every month.  
 Yesterday, she saw the man on top of the hill.  
 Her teaching is excellent.  
 The graph were clearly designed. (Incorrect) (II)  
 The dog are in the yard. (Incorrect) (IC)  
 His brother should plant a tree.  
 The employee asked for a raise yesterday.  
 His brothers pray in the church every day.  
 The girl were in the park yesterday. (Incorrect)  
 The boys reads the newspaper. (incorrect) (II)  
 Our teacher explain our homework every day. (Incorrect)  
 The children made the cake yesterday.  
 The computer are well designed. (Incorrect) (II)  
 The dog \_\_\_ lazy. (Incorrect) (MC)  
 The girl is on the phone.  
 Their clothes \_\_\_ clean. (Incorrect) (MC)  
 Her brothers develop a plan yesterday. (Incorrect)  
 The man accepted a job offer yesterday.  
 They always pay their rent on time.  
 This morning, they wasted their time.  
 After the test, the student noticed his incorrect answer.

Your suitcases is on the plane. (Incorrect) (II)  
The man buy some new clothes last month. (Incorrect) (MI)  
The children were cold.  
Her sisters fixed my car yesterday.  
His friend complained about his boss this morning.  
Usually, she forget about the weekly meeting. (Incorrect)  
The boys were in the store yesterday.  
It rains all day yesterday. (Incorrect) (II)  
The woman sings loudly.  
Last month, the boy pray in the church. (Incorrect) (MI)  
They write the essay together last night. (Incorrect) (II)  
The two boys are very tall.  
He was a smart man.  
The girls were excited about the new job.  
His brother complete school yesterday. (Incorrect)  
It rained hard yesterday.  
The boys should go home today.  
The children love their mothers.  
Yesterday, she plant a tree in her yard. (incorrect) (MI)  
She talks about her vacation after she arrived. (Incorrect) (II)  
Yesterday, the woman sings beautifully. (Incorrect) (II)  
The food \_\_ delicious. (Incorrect) (MC)  
The boy were happy to see her. (Incorrect)  
She usually packs her clothes in her suitcase.  
After school yesterday, the child walk home. (Incorrect) (MI)  
Last week, my father send a present to my mother. (Incorrect) (MI)  
Many Americans \_\_\_ hard workers. (Incorrect) (MC)  
The boys was extremely happy. (Incorrect) (II)  
Yesterday, her daughter opened her presents.  
They ties the rope to the boat.  
Yesterday, the woman filled all of the boxes.  
The three chairs is green. (Incorrect)  
The student are with the teacher. (Incorrect)  
Their sisters are in the house.  
The dogs finished all the food last night.  
Her brothers are very lazy.  
Yesterday, their friends was very hungry. (Incorrect)  
The child were sad last night. (Incorrect)

### B3: SPEAK Task

Download a sample SPEAK Test at the following website:  
<http://www.lehigh.edu/~inesl/Speak/studyingForSpeak.htm>

## References

- Anderson, J. I. (1982). Syllable simplification in the interlanguage of ESL learners. Paper presented at the 16<sup>th</sup> Annual International Conference of Teachers of English to Speakers of Other Languages (TESOL).
- Anderson, J. I. (1983). The difficulties of English syllable structure for Chinese ESL learners. *Language Learning and Communication*, 2(1), 53-61.
- Baddeley, A. (1986). *Working Memory*. New York: Oxford University Press.
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, (158-173).
- Baker, W., Trofimovich, P., Flege, J. E., Mack, M., & Halter, R. (2008). Child-adult differences in second-language phonological learning: the role of cross-language similarity. *Language and Speech*, 51(4), 317-342.
- Barcroft, J., & VanPatten, B. (1997). Acoustic salience of grammatical forms: the effect of location, stress, and boundedness on Spanish L2 input processing. In W. R. Glass & A. T. Perez-Leroux (Eds.), *Contemporary Perspectives on the Acquisition of Spanish. Vol. 2: Production, Processing, and Comprehension* (pp. 109-121). Somerville, MA: Cascadilla Press.
- Best, C., McRoberts, G. and Sithole, N. (1988). Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 345–60.
- Bonner, T. E., & Martohardjono, G. (2012). Performance factors trump representational deficits: Perception and production of English inflections by L1 Mandarin speakers. In A. K. Biller et al. (Eds.), *BUCLD 36 Proceedings* (pp.74-86). Somerville, MA: Cascadilla Press.
- Bozic, M, Marslen-Wilson, W. D., Stamatakis, E. A., Davis, M. H., & Tyler, L. K. (2007). Differentiating morphology, form, and meaning: neural correlates of morphological complexity. *Journal of Cognitive Neuroscience*, 19(9), 1464–1475.
- Broselow, E. (2004). Unmarked structures and emergent rankings in second language phonology. *International Journal of Bilingualism*, 8(1), 51-65.
- Broselow, E., Chen, S., & Wang, C. (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition*, 20, 261-280.
- Brown, C. (1998). The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research*, 14 (2), 136-193.

- Brown, R. (1973). *A First Language*. Cambridge, MA: Harvard University Press.
- Caplan, D., & Waters, G. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, 22, 77-126.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper-Row.
- Clahsen, H. & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27, 3-42.
- Davidson, L. (2005). Addressing phonological questions with ultrasound. *Clinical Linguistics and Phonetics*, 19(6/7), 619-633.
- Davidson, L. (2006a). Phonotactics and articulatory coordination interact in phonology: evidence from non-native production. *Cognitive Science*, 30(5), 837-862.
- Davidson, L. (2006b). Phonology, phonetics, or frequency: influences on the production of non-native sequences. *Journal of Phonetics*, 34, 104-137.
- DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty & M. Long (Eds.), *Handbook of Second Language Acquisition* (pp.313-348). Oxford: Blackwell.
- Dell, G. S., Reed, K. D., Adams, D. R., & Meyer, A. S. (2000). Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1355-1367.
- de Villiers, J., & de Villiers, P. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research*, 2, 267-278.
- Dulay, H. C. & Burt, M. K. (1974). Natural sequences in child second language acquisition. *Language Learning*, 24, 37-53.
- Demany, L., & Semal, C. (2005). The slow formation of pitch percept beyond the ending time of a short tone burst. *Perception and Psychophysics*, 67(8), 1376-1383.
- Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25(6), 1568-1578.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: a psychometric study. *Studies in Second Language Acquisition*, 27, 141-172.

- Ellis, R. (2008). Investigating grammatical difficulty in second language learning: implications for second language acquisition research and language testing. *International Journal of Applied Linguistics*, 18(1), 4-22.
- Epstein, S. D., Flynn, S., & Martohardjono, G. (1996). Second language acquisition: theoretical and experimental issues in contemporary research. *Behavioral and Brain Sciences*, 19(4), 677-758.
- Eubank, L. (1993). On the transfer of parametric values in L2 development. *Language Acquisition*, 3, 183-208.
- Eubank, L., & Grace, S., (1996). Where's the mature language? Where's the native language? In A. Stringfellow et al. (Eds.), *BUCLD 20 Proceedings* (pp. 189-200). Somerville, MA: Cascadilla Press.
- Felser, C., & Clahsen, H. (2009). Grammatical processing of spoken language in child and adult grammars. *Journal of Psycholinguistic Research*, 38, 305-319.
- Fernandez, E. (2000). *Bilingual sentence processing: Relative clause attachment in English and Spanish*. Doctoral Dissertation, CUNY Graduate Center.
- Flege, J. E. (1991). The interlingual identification of Spanish and English vowels: orthographic evidence. *Quarterly Journal of Experimental Psychology*, 43A, 701-731.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (233-277). Timonium, MD: York Press.
- Flynn, S., & Martohardjono, G. (1994). Mapping from the initial state to the final state: the separation of universal principles and language specific properties. In B. Lust, J. Whitman & J. Kornfilt (Eds.), *Syntactic Theory and First Language Acquisition: Crosslinguistic Perspectives. Vol. 1: Heads, Projections and Learnability* (pp. 319-335). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Gass, S. M., & Mackey, A. (2007). *Data elicitation for second and foreign language research*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Gass, S. M., & Selinker, L. (2001). *Second Language Acquisition* (2<sup>nd</sup> ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition*, 68, 1-76.
- Goad, H., White, L., & Steele, J. (2003). Missing inflection in L2 acquisition: defective syntax or L1-constrained prosodic representations? *Canadian Journal of Linguistics*, 48(3/4), 243-263.

- Goad, H., & White, L. (2005). Representational ‘deficits’ in L2: Syntactic or Phonological? In A. Brugos et al. (Eds.), *BUCLD 29 Proceedings* (pp. 216-227). Somerville, MA: Cascadilla Press.
- Goad, H., & White, L. (2006). Ultimate attainment in interlanguage grammars: a prosodic approach. *Second Language Research*, 22 (3), 243-268.
- Goad, H., & White, L. (2007). Prosodic transfer and determiners in Turkish-English interlanguage. In H. Caunt-Nulton et al. (Eds.), *BUCLD 31 Proceedings* (pp. 230-238). Somerville, MA: Cascadilla Press.
- Goldrick, M. (2004). Phonological features and phonotactic constraints in speech production. *Journal of Memory and Language*, 51, 586-603.
- Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory load interference in syntactic processing. *Psychological Science*, 13(5), 425-430.
- Hahne, A., Muller, J., & Clahsen, H. (2006). Morphological processing in a second language: behavioral and ERP evidence for storage and decomposition. *Journal of Cognitive Neuroscience*, 18, 121-134.
- Hatch, E. (1983). Simplified input and language acquisition. In R. Andersen (Ed.), *Pidginization and Creolization as Language Acquisition* (pp. 64-86). Rowley, MA: Newbury House.
- Hawkins, R., & Liszka, S. (2003). Locating the source of the defective past tense marking in advanced L2 English speakers. In R. van Hout et al. (Eds.), *The Lexicon-Syntax Interface in Second Language Acquisition* (pp. 21-44). Amsterdam: John Benjamins.
- Haznedar, B. (2006). Persistent problems with case morphology in L2 acquisition. *Interfaces in Multilingualism*, 179-206.
- Ionin, T., & Wexler, K. (2002). Why is ‘is’ easier than ‘-s’?: acquisition of tense/agreement morphology by child second language learners of English. *Second Language Research*, 18(2), 95-136.
- Jacquemot, C., Dupoux, E., Decouche, O., & Bachoud-Lévi, A-C. (2006). Misperception in sentences but not in words: speech perception and the phonological buffer. *Cognitive Neuropsychology*, 23(6), 949-971.
- Jacquemot, C., & Scott, S. K. (2006). What is the relationship between phonological short-term memory and speech processing? *TRENDS in Cognitive Sciences*, 10(11), 480-486.
- Jia, G., Strange, W., Wu, Y., Collado, J., & Guan, C. (2006). Perception and production of English vowels by Mandarin speakers: age-related differences vary with of amount of L2 exposure. *Journal of the Acoustical Society of America*, 119(2), 1118-1130.

- Joanisse, M. F., & Seidenberg, M. S. (1998). Specific language impairment: a deficit in grammar or processing? *Trends in Cognitive Sciences*, 2(7), 240-246.
- Juffs, A. & Harrington, M. (1996). Garden path sentences and error data in second language sentence processing. *Language Learning*, 46(2), 283-326.
- Kabak, B., & Idsardi, W. J. (2007). Perceptual distortions in the adaptation of English consonant clusters: syllable structure or consonantal contact constraints? *Language and Speech*, 50(1), 23-52.
- Kang, Y. (2003). Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology*, 20, 219-273.
- Klein, E. (2004). Beyond syntax: performance factors in L2 behavior. In Van Patten, B. Williams, J, Rott, S. & M. Overstreet (Eds.), *Form and Meaning Connections in Second Language Acquisition* (pp. 155-180). New York, NY: Erlbaum Associates.
- Klein, E., & Martohardjono, G. (1999). Investigating second language grammars: some conceptual and methodological issues in generative SLA research. In E. Klein & G. Martohardjono (Eds.), *The Development of Second Language Grammars: A Generative Approach* (pp. 3-34). Philadelphia: John Benjamins.
- Klein, E., Stoynezhka, I., Adams, K., Rose, T., Pugach, Y., & Solt, S. (2004). Past tense affixation in L2 English: The effects of lexical aspect and perceptual salience. Poster presented at the 28th annual Boston University Conference on Language Development.
- Ladefoged, P. (2006). *A Course in Phonetics*, 5<sup>th</sup> ed. Boston: Thomson Wadsworth.
- Lardiere, D. (1998a). Case and tense in the 'fossilized' steady state. *Second Language Research*, 14(1), 1-26.
- Lardiere, D. (1998b). Disassociating syntax from morphology in a divergent L2 end-state grammar. *Second Language Research*, 14(4), 359-375.
- Lardiere, D. (2003). Second language knowledge of [±past] vs. [±finite]. In J. M. Liceras et al. (Eds.), *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference (GASLA 2002)* (pp. 176-189). Somerville, MA: Cascadilla Proceedings Project.
- Lenneberg, E. (1967). *Biological Foundations of Language*. New York, NY: John Wiley & Sons.
- Levelt, W. J. M. (1989). *Speaking. From Intention to Articulation*. Cambridge, MA: MIT Press.

- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1-75.
- Li, C. N., & Thompson, S. A. (1987). Chinese. In B. Comrie (Ed.), *The World's Major Languages* (pp. 811-833). New York, NY: Oxford.
- Lin, J-W. (2005). Time in a language without tense: the case of Chinese. *Journal of Semantics*, 23, 1-53.
- Major, R. C. (2008). Transfer in second language phonology. In J.G.H. Edwards & M.L. Zampini (Eds.), *Phonology and Second Language Acquisition* (pp. 63-94). Amsterdam: John Benjamins.
- Martohardjono, G. (1998). Measuring competence in L2 Acquisition. In S. Flynn, G. Martohardjono, & W. A. O'Neil (Eds.), *The Generative Study of Second Language Acquisition* (pp. 151-158). New York, NY: Erlbaum Associates.
- Martohardjono, G., Otheguy, R., Gabriele, A., de Goeas-Malone, M., Szupica-Pyrzanowski, M., et al. (2005). The role of syntax in reading comprehension: a study of bilingual readers. In J. Cohen et al. (Eds.), *Proceedings of the 4<sup>th</sup> International Symposium on Bilingualism*, eds. (pp. 1522-1544). Somerville, MA: Cascadilla Press.
- Martohardjono, G., Valian, V., & Klein, E. (2012). The tense puzzle in second language acquisition: What part representation? What part performance? Manuscript submitted for publication.
- Mody, M., Studdert-Kennedy, M., & Brady, S. (1997). Speech perception deficits in poor readers: auditory processing or phonological coding? *Journal of Experimental Child Psychology*, 64, 199-231.
- Paradis, M. (2004). *A Neurolinguistic Theory of Bilingualism*. Amsterdam: John Benjamins.
- Peperkamp, S. (2004). A psycholinguistic theory of loanword adaptation. In M. Ettliger et al. (Eds.), *Proceedings of the 30th Annual Meeting of the Berkeley Linguistics Society* (pp. 341-352). Berkeley: Berkeley Linguistic Society.
- Peperkamp, S., & Dupoux, E. (2003). Reinterpreting loanword adaptations: The role of perception. In *Proceedings of the 15th International Congress of Phonetic Sciences*, 367-370.
- Peters, A. M. (1985). Language segmentation: operating principles for the perception and analysis of language. In D. I. Slobin (Ed.), *The Cross-Linguistic Study of Language Acquisition, Vol 2: Theoretical Issues* (pp. 1029-1067). Hillsdale, NJ: Erlbaum.
- Pisoni, D. B., & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception and Psychophysics*, 15, 285-290.

- Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Acquisition*, 16(2), 103-133.
- Sahin, N.T., Pinker, S., Cash, S.S., Schomer, D., & Halgren, E. (2009). Sequential processing of lexical, grammatical, and phonological information within Broca's area. *Science*, 326(5951), 445-449.
- Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: evidence that speech production can precede speech perception. *Minnesota Papers in Linguistics and Philosophy of Language*, 8, 48-70.
- Smolensky, P. (1999). Grammar-based connectionist approaches to language. *Cognitive Science*, 23, 589-613.
- Solt, S., Pugach, Y., Klein, E.C., Adams, K., Stoyneshka, I., et al. (2004). L2 perception and production of the English regular past: evidence of phonological effects. In A. Brugos, et al. (Eds.), *BUCLD 28 Proceedings* (pp. 553-564). Somerville, MA: Cascadilla Press.
- Song, J. Y., Sundara, M., & Demuth, K. (2009). Phonological constraints on children's production of English third person singular -s. *Journal of Speech, Language, and Hearing Research*, 52, 623-642.
- Steriade, D. (2001). The phonology of perceptibility effects: the P-map and its consequences for constraint organization. Ms, UCLA.
- Strange, W. (1995a). Cross-language studies of speech perception: a historical review. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 3-45). Timonium, MD: York Press.
- Strange, W. (1995b). Phonetics of second-language acquisition: past, present, future. In *Proceedings: ICPHS 95, Stockholm*, 4, 76-83.
- Strange, W., & Shafer, V. (2008). Speech perception in second language learners. In J. G. H. Edwards & M.L. Zampini (Eds.), *Phonology and Second Language Acquisition* (pp. 153-191). Amsterdam: John Benjamins.
- Sun, C. (2006). *Chinese: A Linguistic Introduction*. United Kingdom: Cambridge.
- Szupica-Pyrzanowski, M. (2009). *Morphological and phonological factors in the production of verbal inflection in adult L2 learners and patients with agrammatic aphasia*. Doctoral Dissertation, CUNY Graduate Center.

- Thompkins, A. C., & Binder, K. S. (2003). A comparison of factors affecting reading performance of functionally illiterate adults and children matched by reading level. *Reading Research Quarterly*, 38(2), 236-258.
- Trenkic, D. (2007). Variability in second language article production: beyond the representational deficit vs. processing constraints debate. *Second Language Research*, 23(3), 289-327.
- Ullman, M. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, 92, 231-270.
- Valian, V. (2006). Young children's understanding of present and past tense. *Language Learning and Development*, 2, 251-276.
- Van Wijnendaele, I. (1998). Reading in a Second Language. *Psychologica Belgica*, 38(3/4), 149-161.
- Vainikka, A., & Young-Scholten, M. (1994). Direct access to X'-theory: evidence from Turkish and Korean adults learning German. In Hoekstra, T. and Schwartz, B. D. (Eds.), *Language Acquisition Studies in Generative Grammar* (pp. 265-316). Amsterdam: John Benjamins.
- Wade-Woolley, L. (1999). First language influences on second language word reading: all roads lead to Rome. *Language Learning*, 49, 447-471.
- Wade-Woolley, L & Geva, E. (2000). Processing novel phonemic contrasts in the acquisition of L2 word reading. *Scientific Studies of Reading*, 4, 295-311.
- Wang, Y., Jongman, A., & Serano, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America*, 113(2), 1033-1043.
- Walter, C. (2008). Phonology in second language reading: not an optional extra. *TESOL Quarterly*, 42(3), 455-474.
- Werker, J. F., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception and Psychophysics*, 37(1), 35-44.
- Werker, J., & Tees, R. (1984a). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7, 49-63.
- Werker, J., & Tees, R. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America* 75, 1866-78.

- Wode, H. (1995). Speech perception, language acquisition, and linguistics: some mutual implications. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 321-347). Timonium, MD: York Press.
- Wode, H. (1997). Perception and production in learning to talk. In S.J. Hannahs and Martha Young-Scholten (Eds.), *Focus on Phonological Acquisition* (pp. 153-191). Amsterdam: John Benjamins.
- Zheng, Y., & Cheng, L. (2008). Test reviews: college English test (CET) in China. *Language Testing*, 25, 408-417.