FIRST-LANGUAGE TRANSFER AND UNIVERSAL MARKEDNESS IN SECOND-LANGUAGE PRODUCTION AND PERCEPTION OF WORD-FINAL OBSTRUENTS AND OBSTRUENT CLUSTERS

By

MILLA CHAPPELL

A THESIS PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

UNIVERSITY OF FLORIDA

2008

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Arts

FIRST-LANGUAGE TRANSFER AND UNIVERSAL MARKEDNESS IN SECOND-LANGUAGE PRODUCTION AND PERCEPTION OF WORD-FINAL OBSTRUENTS AND OBSTRUENT CLUSTERS

By

Milla Chappell

August 2008

Chair: Caroline Wiltshire
Major: Linguistics

While all adult second language learners are influenced by their first language and by universal principles of markedness, Chinese learners of English have a unique challenge as they move from a relatively unmarked language to one with highly complex syllable structures and marked segments. Our study examined the role of first-language influence and universal markedness in adult Chinese learners of English. More specifically, it describes the production of word-final consonant clusters; also, it examines the perception of word-final obstruents and obstruent clusters by listeners from different language backgrounds, Mandarin, Spanish, and English, in order to explore the effect of first-language influence on perceptual cues.

Overall findings of this study show that Chinese learners of English have difficulty producing syllable structures that are disallowed in their own language; additionally, it seems that listeners are attuned to different perceptual cues, as all three listening groups performed differently on the perception task, with the Chinese group having the greatest level of understanding and the Spanish group having the lowest. Findings also suggest that the place of articulation of the key segment and its position in the phrase also predicts success in production and perception.

CHAPTER 1
BACKGROUND AND METHODS

Introduction

While all adult second language learners are influenced by their first language and by universal principles of markedness, Chinese learners of English have a unique challenge as they move from a relatively unmarked language to one with highly complex syllable structures and marked segments. This study seeks to examine the role of first-language influence and universal markedness in adult Chinese learners of English. More specifically, it describes the production of word-final consonant clusters; also, it examines the perception of word-final obstruents and obstruent clusters by listeners from different language backgrounds, Mandarin, Spanish, and English, in order to explore the effect of first-language influence on perceptual cues.

What follows in this chapter is an overview of the relevant literature as well as a description of the methodology and the expectations of the study. Overall findings of this study show that Chinese learners of English have difficulty producing syllable structures that are disallowed in their own language; additionally, it seems that listeners are attuned to different perceptual cues, as all three listening groups performed differently on the perception task, with the Chinese group having the greatest level of understanding and the Spanish group having the lowest. Findings also suggest that the place of articulation of the key segment and its position in the phrase also predicts success in production and perception.

Background

Second Language Acquisition

When learning a language, it is generally argued that earlier contact is better, as learners who are exposed to a new language at a young age tend to become more proficient in a shorter amount of time than those who begin learning in adulthood. The critical period hypothesis for language learning states that "there is a limited developmental period during which it is possible to acquire a language, be it L1 or L2, to normal, nativelike levels. Once this window of opportunity is passed…the ability to learn language declines” (Birdsong, 1992, p. 1). First-language research has for some time proposed that as humans mature neurologically, the ability to learn the L1 decreases (Lenneberg, 1969), and examples of late learners of L1s show the negative effects of withholding language until later in life (Curtiss 1977, 1989).

Discussions of the critical period hypothesis have also been applied to second-language learning. When learning a second language, it is clear that “children are more successful than adults” (Birdsong, 1992, p. 4), and researchers such as Hurford (1991) and Pinker (1994) have attributed this to physiological changes that accompany puberty, such as loss of neural plasticity and loss of the language learning mechanism. However, research has shown that adult second language learners can be quite successful in language learning; in many documented cases, late-learners of a second language are able to attain native-like proficiency. Birdsong (1992) found native-like competence in American learners of French who had moved to France at a mean age of 28.5 years; Mayberry (1993) shows that late learners of American Sign Language attained native-like proficiency; and White & Genesee (1996) studied acquisition of English by native speakers of French and found that many native-like speakers had first studied English after age 12. These studies are just examples from a body of literature that has shown that adult learners of a second language are often able to attain native-like competence, even if the onset of language learning occurred after puberty.

So, although learning a second language may be more difficult for adults than for children, it is certainly not impossible. And, success in second language learning depends on other factors besides age of acquisition, such as similarity between the L1 and L2 and learner adaptability. For example, the Speech Learning Model (SLM) developed by Flege and his colleagues (1995) states that L1 phonetic categories can be adapted later in life as a second language is learned, and that learners can "perceptually relate positional allophones in the L2 to the closest positionally defined allophone (or 'sound') in the L1" (p. 238).

Other models of speech learning state that a learner's ability to perceive and produce non-native contrasts depends on universal markedness of language segments, and articulatory and acoustic similarity of the L1 and L2. Best's perceptual assimilation model (PAM) hypothesizes that a speaker's ability to perceive and produce non-native contrasts will be largely dependent upon the articulatory properties of the sounds in the speaker's first language(s). Best (1995) says that when learning an L1, children "discover correspondences between higher-level invariants of relations among gestures…and linguistically functioning elements" (p. 178) and create gestural categories into which all sounds are placed. First-language sounds and second-language sounds that resemble those in the first language are more easily categorized; however, second-language learners may have difficulty knowing how to classify sounds that don't fit pre-existing categories.

Another model by Kuhl proposes that we compare heard sounds to stored prototypes. Kuhl et al. (1992) describes prototypes as "ideal representatives of a given phonetic category" (p. 606) as identified by adult speakers of that language. For learners of an L2, the L1 prototype could make L2 sounds more difficult to understand and therefore produce, as the L1 categories and prototypes could "pull" close L2 sounds into the wrong perceptual space, producing what Kuhl calls the magnet effect.

While these models focus on the similarity of the L1 and L2, the Markedness Differential Hypothesis (Eckman, 1977) addresses universal markedness, as it states that certain language structures are more difficult to produce, and are therefore more marked cross-linguistically.

And, the hypothesis makes claims about language learning that are based on this universal principle of markedness, namely, that language learners who have more marked structures in their L1s should have an easier time learning languages which have less marked structures while on the other hand, learners with unmarked structures in the L1 will have a difficult time acquiring the marked structure of many L2s.

In general, these models predict that all individuals acquiring the phonology of a second language after childhood will face such challenges as a decline in neural plasticity, interference from a first language, and the markedness of a second language. Adult Chinese learners of English are certainly no exception, and as the segmental phonology of their first language is less marked than the learned language of English, they are faced with unique difficulties in pronunciation and perception.

Chinese Learners of English

Studies have identified that for adult Chinese learners of English, syllable-final obstruents and obstruent clusters are difficult to produce. According to these theoretical models, this difficulty should be a result of the difference between the L1 (Mandarin) and the L2 (English), for the Mandarin language is more specific than English with respect to the types of consonants that are allowed, and in which syllable position they may be found. The difficulty should also be attributed to universal markedness, for the English language contains structures that are highly marked compared to the structures found in Mandarin.

Mandarin Chinese prefers a simple (CV) syllable structure and does not allow any obstruents or consonant clusters in word-final position; however, the target language of English allows single voiced obstruents /b,d,g/ and voiceless obstruents /p,t,k/ in word-final position as well as complex codas such as CC (eg. learn, soft) and CCC (eg. world, first) and CCCC (eg.

twelfths) (Cheng, 1991). English also differs from Mandarin in its use of marked structures; in general, English allows segments that are highly marked and are therefore difficult to acquire cross-linguistically. For example, research has shown that voiced obstruents are more marked in coda position and are therefore more difficult for second language learners to acquire than voiceless obstruent codas, regardless of whether the learner's first language has this voicing distinction, (Eckman, 1981; Flege & Davidian, 1984; Wang, 1995; Broselow et al., 2004). While English allows all types of voiced obstruents in coda position, Mandarin Chinese does not have such a voicing distinction, and this creates difficulty for Chinese learners of English.

According to Broselow's Syllable Structure Transfer Hypothesis, the differences between syllable structures will cause a problem for Mandarin learners of English. Broselow (1984) states, "when the target language permits syllable structures which are not permitted in the native language, learners will make errors which involve altering these structures to those which would be permitted in the native language" (p. 263). Studies have shown that Broselow's assertion is true for Chinese speakers of English, for in general, when faced with word-final obstruents, many native speakers of Mandarin epenthesize a schwa after the final consonant or simply delete the final consonant. Additionally, participants may devoice the final obstruent in an effort to both retain the original form while making the form less marked and therefore easier to articulate (Anderson, 1983; Yin, 1984; Eckman, 1981; Weinberger, 1988; Wang, 1995). Though Mandarin treatment of word-final obstruent clusters has not been well documented, segment deletion and obstruent devoicing may be common simplification strategies to ease articulation as well.

Studies have shown that these simplification strategies are common cross-linguistically since open (CV) syllables are preferred cross-linguistically to closed syllables (CVC), and when words do contain obstruent codas, voiceless obstruents are easier to articulate than voiced obstruents. Since English contains both voiced and voiceless obstruents in word-final position, language learners from various backgrounds have difficulty with the complex syllable structure, and they use similar strategies to ease production.

An early study by Eckman (1981) used word lists to examine word-final obstruent production by two native speakers of Mandarin. Eckman found that early learners of English used epenthesis to offset the CVC structure, and he attributed this to interlanguage transfer and to the markedness hypothesis. According to Eckman (1977), "those areas of the target language which differ from the native language and are more marked than the native language will be difficult" (p. 321). In this case, the English structure CVC is marked, so participants in his study pronounced /rob/ as [rab∂] and /started/ as [statid∂], therefore creating a final syllable structure of CV which is most common in Mandarin Chinese and is least marked universally. Eckman found no examples of deletion or devoicing in his study; however, the number of participants was small, and these learners were highly proficient, high-intermediate/advanced students in a University intensive English program, which could have contributed to their choice to use a this form.

Studies by Yin (1984) and Wang (1995) replicate the findings of Eckman. Yin examined production of word-final stops, word-final nasals, and word-final consonant clusters by native speakers of Mandarin, Taiwanese, and Hakka. Ten Mandarin speakers participated in the study by producing free speech on a given topic. Yin analyzed the word-final stops, nasals, and consonant clusters for changes, and like Eckman, she found that participants often produced an epenthetic schwa as well as other vowels. The study also documents deletion of final stop consonants, though this change was less frequent than epenthesis. Wang's study provides further support for the markedness hypothesis, for although word-final stops are not allowed in Mandarin in either voiced or voiceless form, when native Mandarin speakers read words with word-final voiced and voiceless stops aloud, they made more errors producing voiced stops, which are universally more marked than the voiceless counterpart.

The practice of devoicing final voiced stops has also been explained by the notion of recoverability. According to this idea, language learners may choose to devoice final obstruents to partially preserve the underlying form, instead of deleting the final stop completely. Weinberger (1987) refers to the impact of context on recoverability, stating that when words are read in isolated contexts such as a word lists or short frame sentences, participants will try to retain the recoverable form and will therefore epenthesize or devoice more often than delete. However, when words are found in spontaneous, natural speech in which the context may serve as a cue for meaning, native speakers of Mandarin will be more inclined to delete the final obstruent. Weinberger's (1988) study found that Mandarin learners often devoiced final voiced obstruents in formal context-free situations, and Eckman (1981) reported similar findings for native speakers of Cantonese.

Though complete deletion of a final stop consonant is found much less often in the literature, there are two studies by Tarone (1980) and Anderson (1983) that document this type of deletion by native speakers of Cantonese and Mandarin, and it seems that the task could have influenced the decision to delete rather than epenthesize or devoice. Tarone asked Cantonese speakers to describe a short cartoon in their own words, and Anderson asked participants to talk about their own cultural festivals and holidays. The fact that speakers were producing target words in an informal, even personal context could have contributed to the speech being produced less carefully and with less phonologically recoverable target words, as the context provided the information necessary for understanding to occur.

While studies have not specifically documented Chinese production of obstruent clusters in word-final position, it could be assumed based on the previous findings that these clusters will also cause difficulty. Mandarin prefers a simple (CV) syllable structure and does not allow consonant clusters in word-final position; however, the target language of English allows complex codas such as CC (eg. learn, soft) and CCC (eg. world, first) and CCCC (eg. twelfths) (Cheng, 1991). Most commonly, English final clusters are made up of two consonants; however, it is possible to have as many as four consonants in a word-final cluster. Any consonant may be found word-finally except h, r, w, and j, and where two-consonant clusters are found, there is either a consonant preceded by a pre-final consonant, or a consonant followed by a post-final consonant (Roach, 2000, p. 73). Pre-final consonants, or those that precede consonants in clusters, are m, n, ŋ, l, r and s as seen in 'bump' [bamp] 'bent' [bɛnt] 'bank'[bæŋk] 'belt'[belt] 'ask' [æsk], and 'burp' [bərp] (Roach, 2000, p. 73). Post-final consonants, or those that follow other consonants in clusters are s, z, t, p, d, and ɵ, as seen in 'bets' [bɛts] 'beds' [bɛdz] 'backed' [bækt] 'bagged' [bægd] 'bump' [bamp], and 'eighth' [eItɵ] (Roach, 2000, p. 73). In addition to two-consonant clusters, there are two types of three-consonant clusters: pre-final + final + post-final, as seen in 'helped' [helpt] and 'banks' [banks], or more than one post-final cluster, as seen in 'fifths' [fifɵs] and 'next' [nekst] (Roach, 2000, p. 73). Most four-consonant clusters are analyzed as a pre-final + final + post-final + post-final, as in 'twelfths' [twelfɵs] and 'prompts' [prompts].

While research on Chinese simplification strategies has described single-consonant codas or complex onsets (Tarone, 1980; Eckman, 1981; Anderson, 1983; Weinburger, 1988; Yin, 1984; Flege et al., 1996; Wang, 1995), cross-linguistic research on first and second-language learning has shown that that similar simplification strategies take place when learners are acquiring consonant clusters in coda position (Roberts et. al., 1990; Ohala, 1999; Jongstra, 2003; Pater & Barlow, 2003; Yoo, 2004; Wiltshire, 2006). Young children and beginning second-language learners often delete one segment of the cluster or epenthesize a vowel between the cluster consonants while older children and more advanced learners often produce clusters with less marked features. When deletion occurs, it is common that language learners will retain the least sonorous segment, and when epenthesis occurs, it is usually an unmarked schwa which is inserted.

Research on first-language acquisition has shown that the younger children are, the more likely they are to simplify consonant clusters by deleting one of the segments (Roberts et. al., 1990; Jongstra, 2003). According to Roberts et. al. (1990), there is a significant correlation between the age of the child and the rate of cluster reduction, and it is most often the most sonorous segment which is deleted (Ohala, 1999; Pater & Barlow, 2003). For all children, two-segment clusters are easier to acquire than three-segment clusters (McLeod, 2001), and by the age of two, children are able to produce consonant clusters; however, the clusters are typically simplified in some way and are therefore less marked than those produced by adults (McLeod, 2001).

Second-language research shows similar trends. A longitudinal study by Yoo (2004) examined Korean learners of English, finding that complex codas were acquired after simple ones and that the most common cluster simplification strategy was consonant deletion.

These cross-linguistic findings seem to suggest that universal markedness governs consonant cluster production. However, there is also evidence that first-language background influences acquisition of second-language clusters. Wiltshire (2006)'s study of Indian speakers of English shows that speakers of Indian English simplify complex codas by deleting one consonant, and she also shows that rate of deletion is contingent upon first-language background, for speakers of languages that did not allow clusters more frequently deleted a segment of the cluster than did speakers of languages which allowed clusters.

Rationale for this Study

While many studies have clearly described the production of English word-final stops by Mandarin learners, a detailed description of Chinese production of English word-final consonant clusters is missing in the literature. And, in addition to describing production patterns of Mandarin speakers, it is also important to examine whether these speakers are understood by the people around them when producing these particular sounds.

Previous perception studies have yielded some information about how segmental and syllable features influence cross-cultural understanding. Magen (1998) examined native speakers of Spanish, showing that altered syllable structure (such as vowel epenthesis) and consonant factors, such as /t□-□/substitution and /s/ deletion (p. 392) contributed to higher accent ratings by native speakers of American English. When examining intelligibility in native speakers of English, Bradlow et. al. (1996) found that phonological changes such as vowel space reduction, consonant deletion, and incorrect attachment of segments to syllables were related to decreased intelligibility. However, these studies have focused on native-speaker ratings, and since many Chinese learners of English live and work in multicultural environments such as national universities or international businesses, it is important to examine not only native speaker judgments of Mandarin-accented speech but to also elicit judgments from other language groups.

Worldwide, Mandarin Chinese is spoken by the largest number of people; 873 million people speak Chinese as a first language. English is the second most popular language with 340 million native speakers, and Spanish is the fifth most common language with 322 million native speakers (Gordon, 2005). In the United States, Chinese is the second most common foreign language and is spoken by .78% of the population according to the U.S. Census Bureau (2000). The most common foreign language is Spanish, as it is spoken by 10.71% of the population according to U.S. Census Bureau (2000). The number of Spanish speakers in the United States outnumbers all other foreign language groups by more than 20 million speakers. These two language groups comprise a significant percent of the population; therefore, examining the interaction between these two language groups in an English-speaking environment is important.

Another reason for studying perceptions of Mandarin and Spanish groups as well as native speakers of English is the difference between the syllables structure and phoneme inventories of the three languages. When specifically examining voiced and voiceless obstruents, Mandarin is more restrictive than either Spanish or English (see Table 1-1). As mentioned previously, Mandarin allows voiceless obstruents in word-medial position, but it doesn't allow voiced obstruents at all; additionally, voiceless obstruents are not allowed in word-final position. Spanish also allows both voiced and voiceless obstruents; however, only the voiced obstruent /d/ is allowed in word-final position (Bedore, 1999). The English language allows voiced and voiceless obstruents in word-final position.

With regard to syllable structure, the Mandarin language is again the most restrictive, for Mandarin prefers a CV syllable structure and does not allow syllable-final consonant clusters. Spanish allows two-consonant clusters syllable-finally, and English allows as many as four-consonants in a syllable-final cluster.

The differing phoneme inventories and syllable structures are important, for if L1 phonetics influences L2 perception and production, then it would be possible to see perception differences between language groups, and it may be possible to describe speech cues that influence understanding. Therefore, the aim of this study is to answer the following questions: do adult Chinese learners of English simplify complex codas in the same way that research has documented simplification of simple obstruent codas? And, does listener group play an important role in understanding? For example, are Chinese listeners able to better understand Mandarin-accented speech than English or Spanish listeners? And if so, what are the cues that listeners are attuned to?

Methods

This experiment consists of two tasks: a production task and a perception task. Eleven Chinese learners of English were recorded performing multiple speaking tasks, and native speakers of Mandarin, English, and Spanish listened to the speech of two speakers in order to judge the accuracy of the speech. The speaking data was analyzed for trends in acoustic production, and the perception data was analyzed using simple t-tests and multiple comparisons.

Production Task

Recorded materials were collected from eleven Chinese learners of English (8f/3m). Participants recorded three types of speech: a two-minute segment of free speech describing life in the United States, a read list of forty-six longer conversational sentences, and a read list of sixty short frame sentences with embedded target words. Speech was recorded in a sound-attenuated room with a Marantz PMD660 solid state recorder and head-mounted Shure SM10A microphone that was situated 1 inch from the participant's mouth.

The present study examines only the fifty-six short frame sentences; however, data was collected using three different methodologies in order to elicit data produced in three different environments with varying levels of naturalness and control. Using lists of words or short frame sentences to elicit speech from participants is somewhat unnatural; however, this method for collecting data allows for key words to be produced in an extremely controlled environment without intonation or personal word choice. Longer sentences or paragraphs allow for more natural intonation and pronunciation while still stipulating the words that should be read. And finally, free speech is the most natural and unhindered; however, the researcher has little control over the words that could be produced, and the speaker has the ability to avoid words and structures that are difficult to pronounce.

Previous literature examining word-final consonants and consonant clusters most commonly uses word lists and free speech to elicit data, and differences between the results suggest that methodology can influence production strategies of Chinese learners of English. Tarone (1980), Anderson (1983), and Yin (1984; in Wang, 1995) use free speech to elicit data, and they find less instances of epenthesis and deletion than Eckman (1981) and Weinberger (1983), who both use a combination of word lists, paragraph reading, and free speech. This could be attributed to recoverability, as the meaning of words produced in free speech can be recovered based on context. Based on these findings, the goal of this larger study is to analyze data collected in all three environments; however, the present paper examines only the frame sentences in order to examine the key words produced in controlled environments in both sentence medial and sentence final position.

Thirty-two of these short frame sentences contained word-final obstruents; twenty-four of the frame sentences contained words ending in consonant clusters. The words containing word-final obstruents were divided into eight sets of minimal pairs that were read in two different frame sentences, and each pair differed only in the voicing of the word-final obstruent. The obstruents examined in this experiment were bilabial /b/-/p/, alveolar /d/-/t/, and velar /g/-/k/. The minimal pairs were /mob/-/mop/ and /cub/-/cup/ for the /b/-/p/ contrast; /mad/-/mat/, /god/-/got/, /bid/-/bit/ for the /d/-/t/ contrast; and /pig/-/pick/, /bag/-/back/, and /tag/-/tack/ for the /g/-/k/ contrast. Twenty-four of these short frame sentences included word-final clusters made up of two or three consonants and containing at least one voiced or voiceless obstruent. The clusters examined in this experiment were /ld/, /lb/, /lpt/, /rp/, /rb/, /rd/, /mp/, /rkt/, /rk/, /rst/, /nk/, and /vd/. The read words, in standard English orthography are gold, bulb, sculpt, burp, curb, bird, lamp, marked, park, burst, bank, and moved.

Each word was read twice, once in the sentence-final context of "Repeat __." and once in the sentence-medial context of "Repeat __ again." The two types of frame sentences were ordered randomly in the list, and the target words were spaced so that minimal pairs did not occur in back-to-back sentences. In order to elicit the most natural speech as possible, participants were given time to practice pronouncing the sentences while wearing the microphone before recording began; additionally, many participants asked for pronunciation advice from the researcher. Since read speech is not as natural as that produced in real-world situations, these measures were taken to ensure maximal comfort and naturalness from the participants.

All speaking participants were intermediate or high-intermediate learners of English enrolled in graduate school at the University of Florida. Proficiency level was measured by TOEFL-IBT or SPEAK test scores; TOEFL-IBT scores of 24 or SPEAK test scores of 45 were accepted as evidence of intermediate standing while TOEFL-IBT scores of 26 or SPEAK test scores of 50 were accepted of evidence of high-intermediate standing. In addition, a self-reported proficiency questionnaire was completed by each participant. This form asked for age of beginning study in English, time spent in an English-speaking county, time spent speaking English per day, and the language spoken at home, as well as a 5-point likert scale rating of proficiency in reading, writing, speaking, and listening in English. All participants started studying English after the age of nine and had been living in the United States for at least one month and at most four years. Additionally, all participants gave their speaking ability a likert scale rating of at least three, which was also accepted as evidence of intermediate standing. All participants were paid for their time.

After all speaking data was collected, the data was transferred to a computer using a USB 2.0 cable and was analyzed using Praat version 4.5.19 by Boersma & Weenink (1992-2007). Additionally, the key words were specifically analyzed for vowel length and aspiration of the obstruent in order to determine the production patterns among speakers. Aspiration intensity was measured by highlighting the main burst in Praat and choosing "get intensity" from the "intensity" pull-down menu. Vowel length was measured from the onset of vowel voicing to the offset. The perception study examines all fifty-six frame sentences while the production study examines only the cluster data.

Perception Task

After the recorded material was collected, a perception task was completed. In this experiment, listeners from three different language backgrounds listened to the speech of two speakers and completed a judgment task and a rating task. Listeners were native speakers of English (8), native speakers of Spanish (6), and native speakers of Mandarin Chinese (12). All English listeners were graduate students enrolled at the University of Florida. All Mandarin and

Spanish listeners were intermediate to high-intermediate learners of English who were enrolled in graduate school at the University of Florida. The proficiency standing of listeners were decided by the same method described for speaking participants.

In the first part of the perception task, each listener was seated in a sound-attenuated booth while 108 sentences, 54 from each each of two speakers, were presented through headphones with five-second intervals between each sentence. After each sentence, listeners were asked to first "write the key word in English, as best you can. If you don't understand the word, just write down what part you could understand" and then to "rate how easy each word was to understand on a seven-point scale, with 1 being "difficult to understand" and 7 being "easy to understand." Each participant was given a form with blanks for the key word and a blank likert scale.

In the second part of the perception task, each listener was seated in a sound-attenuated booth and was presented with the same 108 sentences in a different order, this time with only a two-second interval between each sentence. After each sentence, participants were asked only to "write the key word in English, as best you can. If you don't understand the word, just write down what part you could understand." This section did not contain a rating task.

Since it could be assumed that less proficient speakers would be more likely to produce simplified forms, the two speakers used for the perception task were the two participants with the lowest test scores and the least time spent in the United States. After the perception task was completed, multiple comparisons were completed for both perception data and rating data.

Expectations

When developing this study, I expected that the differences between syllable structures and phoneme inventories would provide difficulties for intelligibility between language groups.

Mandarin only allows the least-marked syllable structures of CV or CVC(sonorant); however, this study examines the most marked syllable structure, CVC(stop) and CVCC(C)(C). Therefore, I expected that Mandarin learners would have difficulty producing final consonants and consonant clusters and might try to correct this problem with various simplification strategies. Since the proficiency of speakers in this study is high-intermediate, it might be expected (based on the notion of recoverability) that participants will epenthesize or devoice more than delete, in order to retain the original form to the greatest extent possible.

It was also expected that Chinese speaking and listening participants may use aspiration as a cue to voicing. According to Wang (1995), in the Mandarin language, aspiration in obstruents is phonemic; so, aspirated stops are differentiated from unasperated stops. Therefore, "it is very common among Mandarin speakers to differentiate between English voiced and voiceless stops by means of aspiration. Consequently…/b/ is often perceived and produced as unaspirated /p/, and /p/ as /p^h/ regardless of its context" (39). Therefore, I expect to find that participants may produce /p/ with burst of aspiration and /b/ with none. And, I would also expect that Mandarin listeners would be aware of this voicing cue and could therefore have a higher rate of understanding than Spanish or English listeners. Another expectation is that Spanish listeners would have a more difficult time determining word-final [p] than would native speakers of English, as the voicing distinction does not exist word-finally in Spanish. Whether or not these cues are interpreted correctly by native speakers of English or by native speakers of Spanish is not documented in the literature.

Finally, I expect to find that Mandarin listeners rate the speakers as more "easy to understand" than Spanish or English listeners do. It is possible that Spanish listeners will rate the speech as more "difficult to understand" relative to English listeners based on the fact that

English allows more complex structures and a word-final voicing distinction. Additionally, English speakers will be more familiar with the key words and could therefore have an easier time understanding the speech.

Table 1-1. Phoneme inventory: Mandarin, English, and Spanish.

	Bilabial		Alveolar		Velar	
Mandarin	*p*	p^h	*t*	t^h	*t*	t^h
Spanish	*p*	*b*	*t*	**d**	*t*	d
English	**p**	**b**	**t**	**d**	**t**	**d**

CHAPTER 2
RESULTS

Introduction

Overall production results show that intermediate Chinese speakers of English have learned to produce some word-final clusters but still have difficulty producing clusters containing voiced obstruents in particular. Additionally, these speakers are understood to varying degrees by people from different language backgrounds, for Mandarin listeners were able to understand key words better than English or Spanish listeners. This could be due to differing voicing cues between Mandarin Chinese and English. Additionally, findings show that other factors such as the key word's position in a sentence and the place of articulation of the word-final obstruent play a significant role in production and in understanding.

Production Results

Cluster production data show that Chinese speakers have difficulty producing word-final consonant clusters containing voiced obstruents and that these clusters are usually simplified by either obstruent devoicing or consonant deletion. Word-final consonant clusters containing voiceless obstruents undergo less change, however (Table 2-1).

When simplification occurs, there is a preference for deleting one consonant of three-consonant cluster and for devoicing or deleting one consonant of a two-consonant cluster, and when deletion takes place, learners prefer to delete the most sonorous member of a cluster. When devoicing occurs, it is the obstruent that is devoiced, often along with the preceding consonant. What follows are some examples of simplification strategies found in the data.

As seen in (1), the three-consonant clusters are usually simplified by deletion, and the deletion usually takes the most sonorous member of the 3-consonant cluster. When two consonants are deleted, the final consonant is retained.

(1) /markt/ → [makt]

/skəlpt/→[skəpt]
[skət]

In example (2), the CC clusters often undergo devoicing and deletion. For example, the cluster /vd/ is sometimes devoiced and produced [ft] while sometimes the more sonorant [v] is deleted while the obstruent [d] is retained. In other CC clusters, the obstruent is devoiced.

(2) /muvd/ → [muft]
[mud]

/burd/ →[burt^h]

/curb/ → [crp^h]

Voiceless obstruent clusters are produced with less difficulty; however, some simplification strategies do take place. Example (3) shows that /bank/ undergoes deletion and weakening while /p□rk/ just undergoes deletion.

(3) /benk/ →[bæk]
[bærs]

/park/ →[pak]

Sentence position is an important factor in production of word-final obstruent clusters, for clusters in phrase-medial position, before the word "again", are simplified less often than those in phrase-final position. In particular, when clusters with voiced obstruents are positioned at the end of a phrase, far more mistakes occur than when clusters are phrase-medial. The effect of sentence position is negligible when the clusters contain voiceless obstruents. Numerical results are shown in Table 2-2 while overall percentages are given in Figure 2-1.

The difference in production patterns between clusters containing voiced and voiceless obstruents is striking. Production results show that as Chinese speakers move from an unmarked, native syllable structure to the non-native structure of English, they produce many complex clusters correctly while still simplifying highly marked structures such as voiced obstruents.

Perception Results

The perception study examined listener understanding of key words with two types of codas: complex obstruent codas and simple obstruent codas. Since previous research shows that simple obstruent codas are frequently devoiced and deleted by Chinese learners of English, and since the results of the first experiment show that devoicing and deletion are also common simplification strategies for complex coda structures, it was expected that listeners would have difficulty identifying key words with these coda types. It was also expected that there might be differences between listener groups with regards to understanding.

In general, perception results shows that all three listener groups of Mandarin, English, and Spanish understand Chinese production of word-final voiceless stops and clusters more than word-final voiced stops and clusters; additionally, there is a significant difference in perception between the groups, as Mandarin listeners best understand the Mandarin-accented speech, followed by English listeners and then Spanish listeners. Both voiced and voiceless stops in sentence-medial position are identified as voiced more often than those in sentence-final position. Additionally, place of articulation and vowel duration seem to affect perceptions of word-final stops and word-final clusters. There was no significant effect of speaker ($p=.3$); therefore, data was analyzed using both speakers throughout.

Results show that regardless of speaker, the most significant predictor of listener understanding is the intended voicing of the key word. For all listener groups, key words ending in final voiceless obstruents were perceived correctly more often than those ending in final voiced obstruents ($p<.01$); additionally, there is an effect of place of articulation, for the alveolar obstruent [d] was understood more often than the bilabial obstruent [b] (p =.001) and the velar obstruent [g] is understood more often than alveolar (p = .03) . When misunderstandings occurred, the segment was most often heard as a voiceless stop in the same place of articulation; however, in some cases, the listeners perceived a liquid or glide in the same place of articulation. There was no significant effect of place of articulation for voiceless obstruents; however, descriptive data shows that bilabial [p] was understood correctly more often than alveolar [t] or velar [k]. Results are given in Table 2-3 and Figure 2.

There is also a simple main effect of sentence position (p<01). For all groups, words in phrase-final position are understood more often than those followed by a vowel-initial word (Table 2-4 and Figure 2-3). This effect is most pronounced for voiceless obstruents, as phrase-final position allows for greater aspiration, which could be a cue to voicing for Chinese learners of English. This will be discussed later in the paper.

Group Effects

Another significant predictor of understanding is group. Results show that Mandarin listeners are able to understand all key words better than English listeners (p=.005) and better than Spanish listeners (p<.01). Additionally, English listeners perceive key words better than Spanish listeners (p=.0002). There is also a significant effect of syllable type, for simple C codas were understood more than complex clusters overall (p<.01). Surprisingly, although the Mandarin language does not allow word-final obstruents or complex clusters while the English language allows both, Mandarin listeners understood all cluster types more accurately than English and Spanish listeners (Table 2-5 and Figure 2-4).

Even though English and Spanish listeners don't differentiate well between voiced and voiceless obstruents produced by Chinese speakers, it is apparent that the Chinese listeners are more aware of the distinction. As Wang (1995) predicted, Chinese speakers seem to use aspiration to cue listeners to the quality of the final stop consonant, and Chinese listeners could have more accurate perception scores because they are aware of this cue. The word-final [p] segments that were released with a strong aspirated burst (average intensity of 47.9 db) were always judged as [p] by native English listeners. However, those produced with a less intense burst (avg. 44.4 db) are judged incorrectly by native listeners. Though Chinese listeners do not always judge these words correctly, the perception scores suggest that they may be aware of this cue to voicing.

While aspiration appears to be an important cue for Chinese speakers and listeners, an important voicing cue for native speakers of English is vowel length, for when the vowel length is greater than .25s, the following stop consonant is often interpreted as voiced, even when the intention of the speaker was to produce a voiceless consonant. When the vowel length differed by .081s, and both words were interpreted correctly by all listener groups. However, when the vowel length is similar, native speakers interpreted /mob/ as [mop].

This finding is corroborated by the findings of Wardrip-Fluin (1982) and Flege et al. (1992) who report that vowel length is an important cue for native speakers of English when determining the voicing quality of a word-final obstruent. Flege et. al.'s (1992) study reports that native speakers produce vowels before a word-final voiced obstruent with a longer vowel duration than those produced before a word-final voiceless obstruent. Additionally, when judging the /t/-/d/ contrast in speakers of foreign languages, native speakers used vowel length as a significant cue. According to Flege et al., "listeners may tailor their perceptual processing of stops to the range of cues that are readily available" (p. 141). So, in the case of this study, where there was limited contextual information and the target words were only one syllable in duration, vowel length was a readily available cue.

While we have evidence for voicing cues for Mandarin listeners and English listeners, it is unclear what Spanish listeners use to determine voicing in foreign-accented speech. No clear patterns exist for aspiration or vowel length, as with Mandarin and English, so since Spanish listeners performed more poorly than even English listeners, it may be that simple familiarity with English words allowed English speakers to identify key words more accurately.

Important findings include the production patterns of Mandarin learners of English, as there are clear patterns in simplification strategies depending upon the size and voicing of the obstruent clusters. Additionally, perception results show that understanding of different language groups can be influenced by the first language; this may be a result of first-language influence on perceptual cues to voicing.

Table 2-1. Overall production results

Target cluster	English Word	Percent Correct	Actual Production
[rp]	burp	95 (21/22)	[b] (1/22)
[rst]	burst	86 (19/22)	[rd] (1/22)
			[rθt] (1/22)
			[rθ] (1/22)
[nk]	bank	81 (18/22)	[k] (3/22)
			[rs] (1/22)
[mp]	lamp	81 (19/22)	[m] (2/22)
			[p] (1/22)
[lpt]	sculpt	45 (10/22)	[pt] (5/22)
			[t] (2/22)
			[rpt] (3/22)
			[ptə] (2/22)
[rkt]	marked	27 (6/22)	[makt] (13/22)
			[də] (1/22)
			[kIt] (2/22)
[rk]	park	40 (9/22)	[k] (12/22)
[rb]	curb	45 (10/22)	[rp] (12/22)
[rd]	bird	45 (10/22)	[rt] (10/22)
			[r] (1/22)
			[rdə] (1/22)
[vd]	moved	18 (4/22)	[vt]/[ft] (9/22)
			[d] (4/22)
			[vəd] (3/22)
			[vdə] (3/22)
[ld]	gold	4 (2/22)	[d] (8/22)
			[lt] (4/22)
			[ldə] (4/22)
			[t] (4/22)
[lb]	bulb	0 (0/22)	[b] (7/22)
			[p] (7/22)
			[mp] (5/22)
			[lp] (1/22)
			[rb] (1/22)
			[rp] (1/22)

Table 2-2. Phrase position results

Target	Phrase Medial	Percent Correct	Phrase Final	Phrase Final
[rp]	burp again	100 (11/11)	burp	90 (10/11)
[rst]	burst again	100 (11/11)	burst	90 (10/11)
[rkt]	marked again	27 (3/11)	marked	27 (3/11)
[nk]	bank again	81 (9/11)	bank	81 (9/11)
[mp]	lamp again	72 (8/11)	lamp	90 (10/11)
[lpt]	sculpt again	45 (5/11)	sculpt	45 (5/11)
[rk]	park again	36 (4/11)	park	45 (5/11)
[rd]	bird again	72 (8/11)	bird	18 (2/11)
[rb]	curb again	63 (7/11)	curb	11 (1/11)
[vd]	moved again	36 (4/11)	moved	0 (0/11)
[ld]	gold again	18 (2/11)	gold	0 (0/11)
[lb]	bulb again	0 (0/11)	bulb	0 (0/11)

Table 2-3. Effect of place of articulation on perception scores: percent correct

	Bilabial	Alveolar	Velar
Voiced	14%	28%	36%
Voiceless	85%	74%	82%

Table 2-4. Effect of sentence position on perception scores: percent correct

	Voiceless	Voiced
Overall	74%	28%
Sentence Medial	67%	27%
Sentence Final	79%	29%

Table 2-5. Perception of codas by group and coda type: percent correct

	All Key Words	Clusters	Simple C
Mandarin	54%	52%	55%
English	48%	45%	50%
Spanish	39%	33%	44%

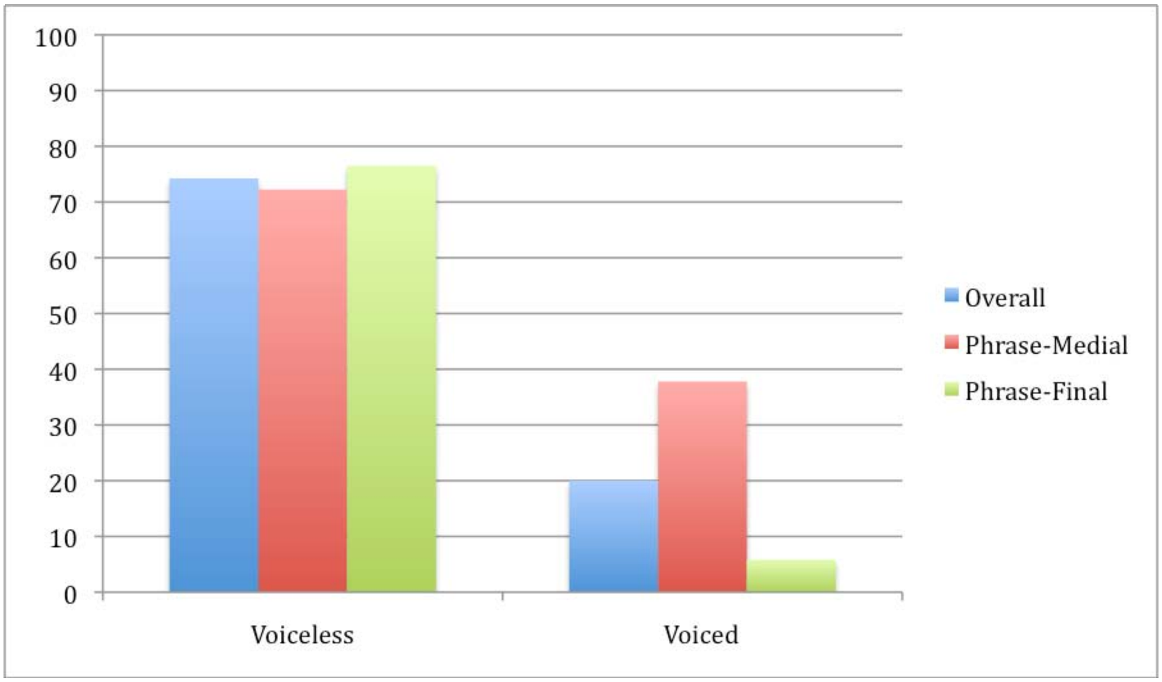

Figure 2-1. Production results – percent correct

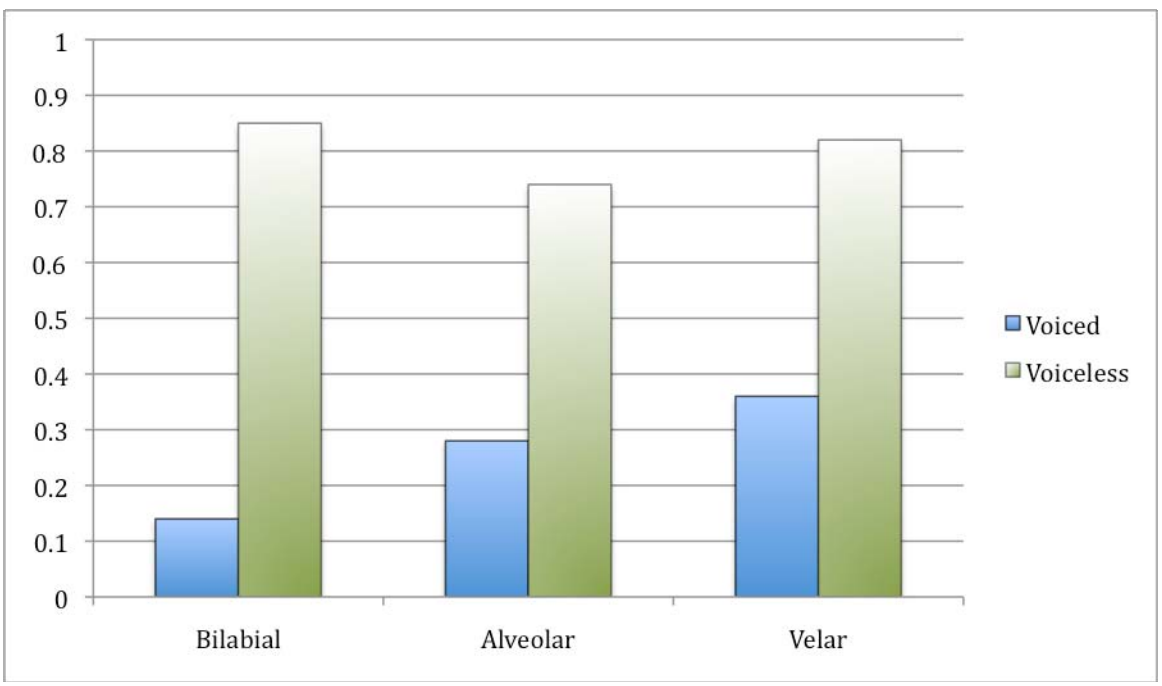

Figure 2-2. Effect of place of articulation on perception scores: percent correct

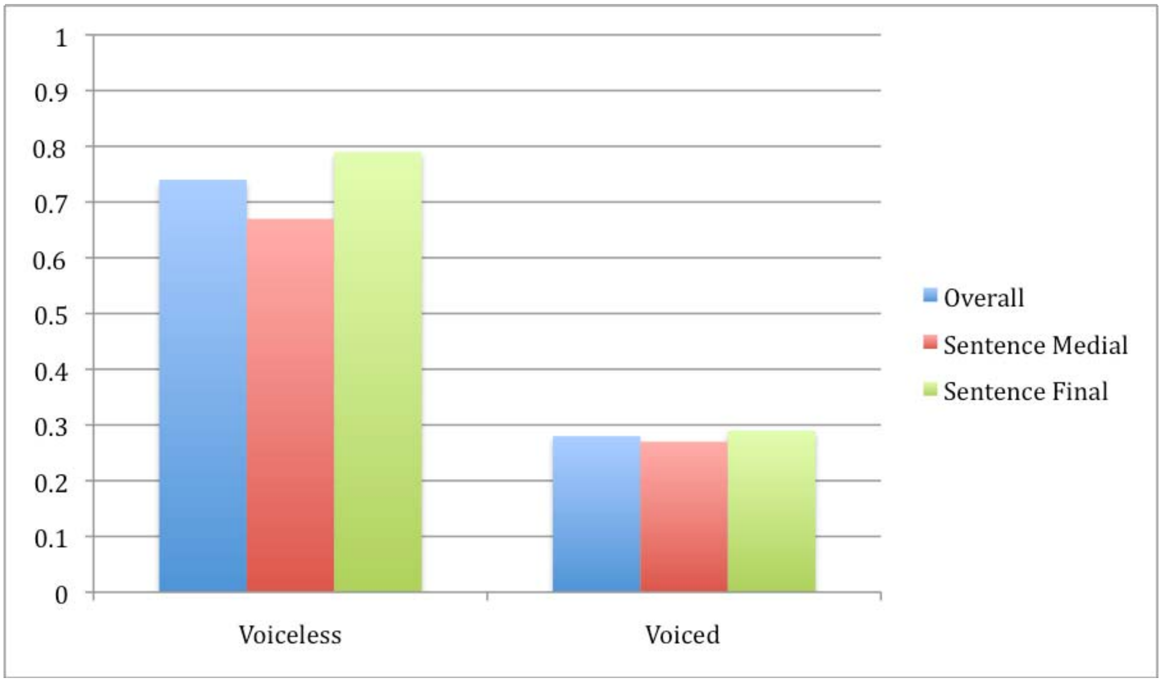

Figure 2-3. Effect of sentence position on perception scores: percent correct

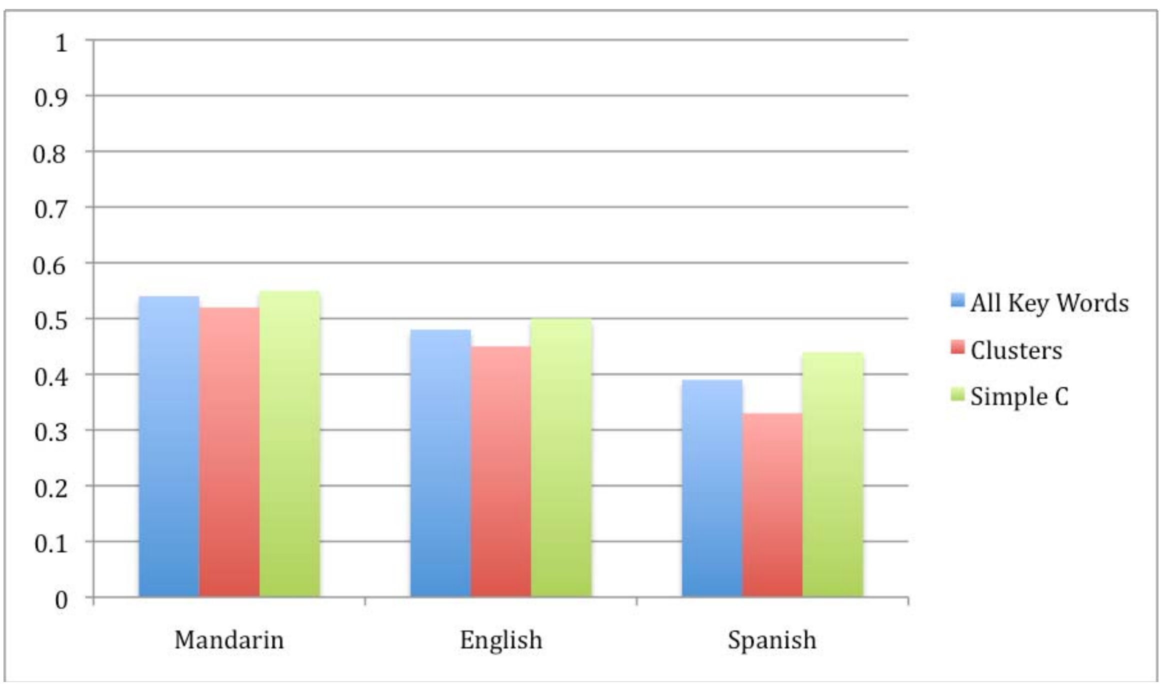

Figure 2-4. Perception of codas by group and coda type: percent correct

CHAPTER 3
ANALYSIS AND CONCLUSIONS

Introduction

In this chapter, the production data is analyzed within the framework of Optimality Theory (OT) in order to explain production as a dynamic interaction between the first language of Mandarin and the second language of English; additionally, a model of language learning by Chinese learners of English is provided within the gradual learning algorithm. Finally, major conclusions from production and perception data are discussed, and limitations of the study as well as ideas for future research are given.

Analysis

Optimality Theory

Production data shows that Chinese learners of English simplify complex codas in various ways, and Optimality Theory (OT) provides a framework within which we may analyze the data by discussing how the interaction between language transfer and universal markedness predict production strategies. OT can be used to analyze second-language data as multiple sets of ranked constraints, showing how speakers move from total faithfulness to the native language ranking to accuracy in the second language. Therefore, the OT constraints needed to account for the Chinese production data will address both faithfulness to the native language of Mandarin and universal markedness (Table 3-1).

According to Broselow et al. (2004), Mandarin Chinese has a constraint ranking which favors markedness constraints and therefore produces unmarked structures. In the following ranking, markedness constraints will keep codas from surfacing at all:

NoCoda, NoVoicedObsCoda, NoObsCoda, *ComplexCoda >> Faithfulness

The ranking of NoCoda above faithfulness is all that is needed for the Mandarin language to avoid codas of any kind, including obstruent codas, complex codas, or voiced obstruent codas, as the other constraints describe. English, on the other hand, ranks faithfulness above NoCoda, therefore producing complex syllable structures and marked segments:

Faithfulness >> NoCoda, NoVoicedObsCoda, NoObsCoda, *ComplexCoda

As Chinese learners of English become more proficient, they must learn to demote markedness constraints below faithfulness constraints. According to Broselow et al. (1998), the reranking process can be attributed to the interaction between language transfer and universal markedness, for learners begin with the ranking of the native language and then,

> "under pressure from interlanguage data, begin to construct an interlanguage grammar in which the rankings of constraints may differ from the native-language ranking. In this case, markedness effects that are not visible in either the native or the target language may become visible in the interlanguage data" (279).

Therefore, it was expected in this experiment and also found that intermediate Chinese learners of English have learned to demote certain markedness constraints, but not all of them, and that intermediate learners therefore produce some complex codas while also relying upon simplification strategies to avoid highly marked structures.

In general, intermediate learners of English simplify complex CC clusters in one of three ways: consonant deletion, voiced obstruent devoicing, or both, and they simplify complex CCC clusters by consonant deletion. Optimality Theory should be able to account for all of these changes. In this section, we will examine each of these simplification strategies separately, giving examples and appropriate constraint rankings.

One common strategy used for correcting voiced obstruent clusters is obstruent devoicing. For example, /burd/ → [burt] 80% of the time and /kurb/ → [kurp] 100% of the time that changes occurred. In order to account for this process within OT, we need the following ranking:

NoVoicedObsCoda >> MaxI-O >> IdentI-O[ObsVce]

Additionally, since complex codas remain, *ComplexCoda and NoCoda must be ranked low (see figure 3-1):

NoVoicedObsCoda, MaxI-O >> IdentI-O[ObsVce] >> ***ComplexCoda, NoCoda**

Another common strategy is consonant deletion. For example, /park/ → [pak] 84% of the time and /bulb/ → [bub] or [bup] 77% of the time that changes occur. The following ranking is needed to allow for consonant deletion:

*ComplexCoda >> MaxIO

When deletion occurs, it is usually the non-obstruent that is deleted. In order to account for this phenomenon, the OT constraint must describe which consonant is deleted.

*Complex Coda, **$\mathbf{MaxC_{[-son]}}$** >> MaxI-O

Additionally, epenthesis is a very uncommon strategy and therefore DepI-O should be ranked higher than the preferred strategy of MaxI-O violations (figure 3.2).

*ComplexCoda, $MaxC_{[-son]}$, **DepI-O** >> MaxI-O

This ranking isn't complete, however, as it should also be able to account for three-consonant coda simplification such as in /rkt/ → [kt]. In this example, we need to introduce the constraint *CCC, and it should be ranked above *ComplexCoda to keep three-consonant clusters from emerging on the surface. Additionally, *ComplexCoda must be demoted, as a complex coda is left on the surface (figure 3-3).

In some cases, both obstruent devoicing and consonant deletion occur. For example, /bulb/ → [bap] and /gold/ → [got]. In this case, the constraints must be ranked to allow both deletion and devoicing to occur (figure 3-4).

NoVoicedObsCoda, $MaxC_{[-son]}$, DepI-O >> *ComplexCoda >> MaxI-O, IdentI-O[ObsVce]

With this simple ranking, we can capture the general trends of the data; however, there is great variability in the data that is not accounted for here. Ideally, the same ranking of constraints should account for any language input, so all simplification strategies should be possible outputs of the theory. However, because of variability in speaker strategies, it is necessary to talk about general patterns in production and in learning. That is the purpose of the Gradual Learning Algorithm.

Modeling Learning in the Gradual Learning Algorithm

Boersma's Gradual Learning Algorithm (GLA) in stochastic OT (Boersma, 1999; Boersma & Hayes, 2001) is a program in which learning can be represented in Optimality Theory. The difference between traditional OT and stochastic OT is the possibility of overlapping constraints. While OT implies strictly ranked constraints that produce the same output on every try, stochastic OT ranks constraints along a continuum (see figure 3-5), therefore allowing for various outputs on different trials (see figure 3-6).

The GLA within Stochastic OT is a valuable model within which to discuss learning, for given an initial state grammar and target-language data, the program can predict possible stages that learners may go through during the language acquisition process. The program does not assume that learners move from one stage to the next abruptly; rather, as they move between stages, variation occurs. When programming an OT grammar, the user gives every constraint a ranking value on a scale of 0 – 100, and noise, or possible variation, is added to each trial. As the algorithm represents error-driven learning, small changes in the ranking value occur as the given grammar compares its constraint ranking with input-output pairs from the target grammar.

The Gradual Learning Algorithm was first used within Stochastic OT to model first-language learning; however, others have applied the program to second-language learning (Xu, 2003 in Broselow et. al., 2004; Wiltshire, 2006) and it is also the purpose of this section to apply the algorithm to Chinese learners of English. Since all learners in this experiment began studying English after the age of nine, it is assumed that the first language had been acquired before English learning began. Therefore, the initial state grammar in this analysis is Mandarin Chinese. As previously discussed, the Mandarin language prefers unmarked structures; therefore, the Mandarin OT grammar is set to favor markedness constraints over faithfulness constraints. All constraints are considered to be on a ranking scale, and following the example of Xu (2003) (in Broselow et. al., 2004), the markedness constraints were given a ranking value of 100 while the faithfulness constraints received a ranking value of 88. The twelve-point difference in the ranking values insures that the constraints to do not overlap in the grammar. The constraints and rankings are given in Table 6. The standard deviation was set at the standard 2.0 (Boersma & Hayes 1999). Given the initial state grammar, Praat's output distribution predicts that the program will give open syllables 100% of the time, until constraints shift in OT ranking. All other outputs should never occur.

According to Broselow's description of the GLA, "the rate at which a markedness constraint is demoted is a function of the frequency with which the constraint is violated by input structures…the more general constraint will be demoted more quickly than the more specific (and therefore less frequently violated) constraint" (137). Therefore, I considered the frequency of complex codas and planned the input-output pairs accordingly. The desired output of the GLA for the input /bulb/ is the faithful form [bulb]. Therefore, I gave the desired output of [lb] the highest score in the input-output pairs. Other outputs are ranked according to frequency, for according to the BYU Corpus of American English, English words ending in liquids [r] and[l] are more common than those ending in voiceless obstruents [p] [t] and [k]. Therefore, since we might expect to see users simplifying [lb] clusters to [l] before [p], [l] is ranked as the second most common in the GLA, and [p] as the third most common. [b] was considered to be the least common output given the marked nature of the segment in word-final position (see figure 3-7).

When the GLA is set to learn at a rate of 300 trials, the grammar accurately describes the learners' production data collected for this experiment. In order to move from the totally unmarked Mandarin grammar to the totally marked English data, the algorithm goes through three stages, thus capturing three important observations about the learning process. The grammar tells us that sonorant codas are acquired before obstruent codas, that voiceless obstruent codas are acquired before voiced ones, and that complex codas with voiced obstruents are acquired last. I will consider each of these in turn.

The first observation is that Sonorant Codas are acquired before obstruent codas. In the first change, the markedness constraint NoObsCoda and the faithfulness constraint Max-IO are ranked above markedness constraints NoComplexCoda, NoCoda, and NoVoicedObsCoda, and the corresponding output is [bul]. This prediction corresponds with research documenting that learners tend to acquire non-complex, sonorous syllable codas before complex, obstruent codas (Yoo, 2004; Wiltshire, 2006); however, it does not directly correspond to the data collected, as learners were much more likely to delete the sonorant consonant in the consonant cluster (Figure 3-8).

I suspect the reason the algorithm produced this output was due to the frequency ratings liquids and obstruents. However, the algorithm doesn't take sonority sequencing of consonant clusters into consideration; therefore, the predicted output was not found in the data.

The second observation is that voiceless obstruent codas are acquired before voiced obstruent codas. After another 300 tries, NoVcdObsCoda and Max-IO were moved in front of Ident-IO[ObsVce] therefore producing optimal outputs with voiceless obstruent clusters (see

Table 8). Again, this prediction by the GLA corresponds to numerous cross-linguistic studies citing that voiceless obstruents are less marked than voiced ones, and more specifically, this prediction corresponds to studies showing that Chinese learners of English tend to acquire voiceless obstruent codas first (Tarone, 1980; Eckman, 1981; Anderson, 1983; Wang, 1995; Broselow et. al., 2004). This prediction by the GLA also corresponds with the production data collected in this experiment, for two speakers were recorded producing [bulb] as [bulp] and [bump].

The third observation is that complex codas with voiced obstruents were acquired last. After a final 300 tries, Max-IO became a highly-ranked constraint while IdentObsVce moved lower, therefore allowing [bulb] to surface as the correct output (Figures 3-9 and 3-10). This makes sense given the constraint rankings and the target data; however, what is not certain is whether language learners actually ever reach native-like competency or native-like constraint rankings.

The results do correspond with the output distributions given by Praat for the initial state grammar that has undergone learning. For the input /bulb/, the output [bul] was predicted to occur 26% of the time; [bulb] was predicted to occur 25% of the time; [bub] was expected 22% of the time; [bup] was predicted 11% of the time; [bulp] was predicted 9% of the time; and [bu] was predicted to occur 0% of the time.

To test the GLA's prediction for complex CCC clusters, the algorithm was once again programmed for the input /markt/. Once again, the initial state grammar is that of Mandarin, with all markedness constraints being ranked above all faithfulness constraints. The OT tableau is given in Table 10; the optimal output is the unmarked syllable form, [ma] (Figure 3-11).

The desired output of the GLA for the input /markt/ is the faithful form [markt]. Therefore, the pair distribution ranked the desired output [markt] the highest. Simplified codas were ranked according to the BYU Corpus of Spoken American English, with [rk] being the most common coda, [r] and [kt] the second most common, and [k] the least common coda in the spoken English database.

When the GLA is set to learn at a rate of 300 trials, the grammar accurately describes the learners' production data collected for this experiment. In moving from the initial state of Mandarin to the target state of English, the algorithm goes through only two stages, one simplified form and the actual target. After 300 trials, the GLA still predicts [ma] as the output; however, after 600 and 900 trials, the GLA predicts [makt] as the output. After 1200 trials, *CCC is demoted, and the GLA predicts the correct output of [markt] (Figures 3-12 and 3-13).

The prediction of the GLA is accurate for language data collected, for /markt/ was only produced as such 27% of the time, and100% of all simplified forms were produced as [makt]. Results also correspond to output distributions given by the Praat program: that [park] will surface 47% of the time after learning; that [pak] will surface 26% of the time; that [par] will surface 24% of the time; and that [pa] will surface 0% of the time.

Conclusions

The production data expands the observation that has been made about word-final obstruent production by Chinese learners of English by showing that similar strategies also apply to production of word-final clusters. These strategies are subject to first-language constraints and constraints on universal markedness, and they are predictable based on the first language (Mandarin) and the second language (English).

The Gradual Learning Algorithm accurately predicts the results of the production task, namely that consonant clusters containing voiced obstruents are the most difficult for Chinese learners of English to acquire and that simplification strategies such as consonant deletion and devoicing frequently occur. Additionally, the general observations about language learning are important, as they reflect the role of universal markedness in language acquisition. What is unrealistic about the GLA, however, it its somewhat static representation of the learning. In actual language acquisition, people don't learn at the same rate, nor do they always progress through the same stages. This could be because of different frequencies of input. For example, language learners immersed in a second-language environment have high frequencies of input and may learn more quickly than those who have almost no input. Additionally, most people never achieve the native-like production; rather, if they are able to communicate effectively, many language learners remain in an intermediate stage.

The perception data shows us that speakers from different first-language backgrounds interpret the same input in different ways. In this study, Chinese listeners performed significantly better at understanding Mandarin-accented speech, and English listeners performed better at understanding the Mandarin-accented speech than did Spanish listeners. Since the input was exactly the same, it seems that these three groups of listeners were paying attention to different cues in the data. Based on the predictions of Wang (1995) related to aspiration and the findings of Flege (1992) about native-speaker attunement to vowel length, the results of the experiment seem to suggest that Mandarin listeners are more aware of aspiration as a cue to voicing while native speakers of English are more attuned to vowel length as the most important cue; however, it is still unclear if Spanish listeners were aware of any specific cues that influenced their performance on the perception task.

This is an important finding, as the use differing cues will certainly lead to misunderstandings between language groups. Additionally, the finding speaks to the relationship between perception and production. Previous literature has proven that perception and production are closely related (Flege & Effting, 1987; Flege, 1993; Bradlow et. al., 1996), while it is still debated whether perception precedes production or vice versa. In this case, it is not clear whether Mandarin learners of English may be perceiving aspiration as a cue to voicing because that is what they are trained to produce. Or, whether they may be using aspiration in their production because they perceive this distinction in the speech of other language learners.

Studies have come to various conclusions about the relationship between production and perception. One study by Flege & Effting (1987) suggests that production precedes perception. This study examined Dutch learners of English production of /t/-/d/ contrast in English, finding that although Dutch learners were able to produce the sounds with relative accuracy, they were not able to perceive the difference between the sounds in a perception study. However, multiple other studies suggest that perception of non-native contrasts both precedes and influences production. For example, a study by Bradlow et. al. (1996) proved that Japanese learners of English who were trained to perceive the /r/-/l/ contrast also improved in their production of the sounds, and another study by Flege (1993) showed that perception of English words ending in /t/ or /d/ by Chinese learners of English was correlated with the amount of foreign accent they were judged to have by native speakers of English. Therefore, Flege concludes that perception of the /t/-/d/ contrast is related to native-like production of these sounds.

The relationship between perception and production is important for our discussion of Chinese learners of English, for if these learners are indeed using non-native cues to voicing in their production and perception, misunderstandings could occur when speaking to other language groups. A better understanding of how cues to voicing are used in perception and production could impact the way we think about second-language learning and teaching. For example, perceptual training methods could be devised that train listeners to use more native-like cues, and if these training methods are able to improve the perception of word-final obstruents and obstruent clusters, it may also positively affect the production of these second-language forms.

The limitations of this study include a small sample size and a limited number of cluster types. Additionally, studying production within such a limited speech environment such as a frame sentence does not result in completely natural production patterns; therefore, findings in this study cannot be generalized to conversational situations. Finally, all transcriptions were completed by one researcher, with no outside judgments, and the proficiency level of speakers and listeners could be more rigidly controlled in order to examine the role of proficiency on perception and production.

Future research could focus on training methods and perceptual acquisition. In particular, I would like to focus future research on psycholinguistic training methods, for previous research has shown that perceptual training can cause improvements in behavioral accuracy and production as well as preattentive reaction times. For example, Tremblay et. al. (1997) examined the effects of training on segment perception. Native English participants were trained to hear and identify a prevoiced labial stop, and a prevoiced alveolar sound was used as a transfer stimuli. After only one training session, the preattentive response increased significantly in duration and area. Bradlow et. al.'s (1997) study also shows that perceptual training can also increase production accuracy. After training Japanese learners of English to perceive the difference between /r/ and /l/, it was found that their ability to produce the segments accurately also increased.

This study only begins to touch upon important issues for Chinese learners of English, namely how the first language and universal principles influence the ability to learn a second language. While these theoretical findings are important, I believe they have implications for further research on training methodology and perceptual cues, which I leave to future research.

Table 2-1. Markedness and Faithfulness Constraints

Markedness	
NoCoda	Syllables are open (Kager, 1999)
*ComplexCoda	Codas are simple (Kager, 1999)
NoObsCoda	No obstruents in the coda (Broselow et al., 2004)
NoVoicedObsCoda	No voiced obstruents in the coda (Broselow et al., 2004)
*CCC	No CCC clusters (Kager, 1999)
Faithfulness	
MaxI-O	Input segments must have output correspondents (Kager, 1999)
DepI-O	Output segments must have input correspondents (Kager, 1999)
IdentI-O	No featural changes (Kager, 1999)
IdentI-O[ObsVce] (Kager, 1999)	Correspondent obstruents are identical in their specification for voice
$MaxC_{[-son]}$	Inputs segments with the feature [-son] must have output correspondents

/rd/	NoVoicedObsCoda	MaxI-O	IdentI-O[ObsVce]	*ComplexCoda	NoCoda
☞ rt			*	*	*
d	*!	*			*
rd	*!			*	*
r		*!			

Figure 3-1. Obstruent devoicing as the preferred strategy

/rk/	*ComplexCoda	$MaxC_{[-son]}$	DepI-O	MaxI-O
rk	*!			
☞ k				*
r		*!		*
rik			*!	

Figure 3-2. Consonant deletion as the preferred strategy

/rkt/	*CCC	MaxC[-son]	DepI-O	*ComplexCoda	MaxI-O
rk		*!		*	*
☞ kt				*	*
r		*!			**
kit			*!		*
rkt	*!				

Figure 3-3. Three-consonant coda simplification

/lb/	NoVcdObsCda	MaxC[-son]	DepI-O	*ComplexCoda	MaxI-O	IdentI-O[ObsVce]
☞ p					*	*
b	*!				*	
lb				*!		
lp				*!		*
lepe			*!			*
l		*!			*	

Figure 3-4. Deletion and devoicing

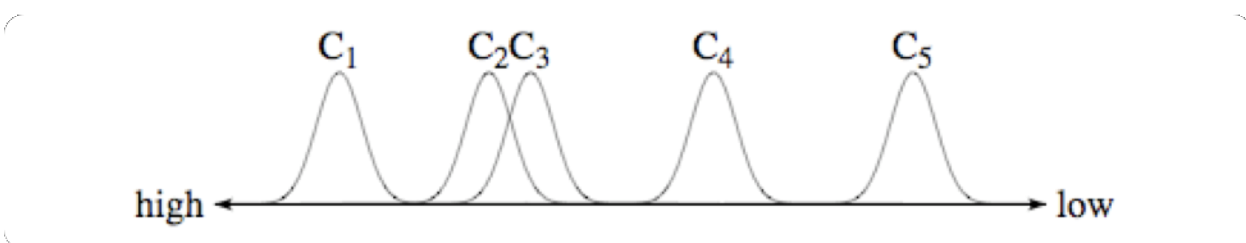

Figure 3-5. Example of ranked constraints in stochastic OT (Boersma, 2003)

strict ranking determines winning candidate form

90% of the time: $C_2 \gg C_3$

10% of the time: $C_3 \gg C_2$

Figure 3-6. Possible output of ranking in stochastic OT

	ranking value	*disharmony*
NoCCC	100.000	100.000
NoCoda	100.000	100.000
NoComplexCoda	100.000	100.000
NoObsCoda	100.000	100.000
NoVcdObsCoda	100.000	100.000
DepIO	88.000	88.000
Ident	88.000	88.000
IdentObsVce	88.000	88.000
MaxIO	88.000	88.000
MaxObs	88.000	88.000

bulb	NoCCC	NoCoda	NoComplexCoda	NoObsCoda	NoVcdObsCoda	DepIO	Ident	IdentObsVce	MaxIO	MaxObs
bup		*		*			*	*	*	
bulp		*	*	*			*	*		
bulb		*	*	*	*					
bub		*		*	*				*	
bump		*	*	*			*	*		
☞ bu									**	*
bul		*							*	*

Figure 3-7. Ranking values and tableau for the initial state – Mandarin Chinese

	ranking value	*disharmony*
NoCCC	100.000	103.729
NoObsCoda	93.721	97.842
MaxIO	96.556	96.775
NoComplexCoda	96.000	94.398
IdentObsVce	91.534	93.291
Ident	91.534	92.336
NoVcdObsCoda	90.187	91.709
MaxObs	94.279	90.612
NoCoda	89.382	88.177
DepIO	88.000	85.558

bulb	NoCCC	NoObsCoda	MaxIO	NoComplexCoda	IdentObsVce	Ident	NoVcdObsCoda	MaxObs	NoCoda	DepIO
bup		*!	*		*	*			*	
bulp		*!		*	*	*			*	
bulb		*!		*			*		*	
bub		*!	*				*		*	
bump		*!		*	*	*			*	
bu			**!					*		
☞ bul			*					*	*	

Figure 3-8. Ranking values and tableau at 300 trials

	ranking value	*disharmony*
NoCCC	100.000	99.333
DepIO	88.000	90.220
NoVcdObsCoda	28.850	31.035
MaxObs	29.783	29.189
Ident	29.164	29.154
MaxIO	27.936	28.140
IdentObsVce	29.462	27.900
NoComplexCoda	27.543	27.464
NoCoda	22.428	25.343
NoObsCoda	25.879	23.092

bulb	NoCCC	DepIO	NoVcdObsCoda	MaxObs	Ident	MaxIO	IdentObsVce	NoComplexCoda	NoCoda	NoObsCoda
bup					*	*!	*		*	*
☞ bulp					*		*	*	*	*
bulb			*!					*	*	*
bub			*!			*			*	*
☞ bump					*		*	*	*	*
bu				*!		**				
bul				*!		*			*	

Figure 3-9. Ranking value and tableau at 600 trials

	ranking value	*disharmony*
NoCCC	100.000	100.503
MaxIO	97.662	98.463
NoComplexCoda	97.038	95.621
MaxObs	93.810	95.549
NoObsCoda	94.190	94.969
IdentObsVce	91.514	92.455
Ident	91.514	91.589
NoCoda	87.790	91.300
NoVcdObsCoda	90.676	89.690
DepIO	88.000	86.598

bulb	NoCCC	MaxIO	NoComplexCoda	MaxObs	NoObsCoda	IdentObsVce	Ident	NoCoda	NoVcdObsCoda	DepIO
bup		*!			*	*	*	*		
bulp			*		*	*!	*	*		
☞ bulb			*		*			*	*	
bub		*!			*			*	*	
bump			*		*	*!	*	*		
bu		*!*		*						
bul		*!		*				*		

Figure 3-10. Ranking values and tableau for output [bulb].

markt	NoCCC	NoCoda	NoComplexCoda	NoObsCoda	NoVcdObsCoda	DepIO	Ident	IdentObsVce	MaxIO	MaxObs
markt	*	*	*	*						
mark		*	*	*					*	*
makt		*	*	*					*	
market		*		*		*				
mak		*		*					**	*
☞ ma									***	**

Figure 3-11. Initial state grammar: input /markt/

markt	NoVcdObsCoda	DepIO	Ident	NoCCC	IdentObsVce	MaxIO	NoObsCoda	NoComplexCoda	NoCoda	MaxObs
markt				*!			*	*	*	
mark						*	*	*	*	*!
☞ makt						*	*	*	*	
market		*!					*		*	
mak						**!	*		*	*
ma						**!*				**
mar						**!*			*	**

Figure 3-12. GLA after 600 trials

markt	NoVcdObsCoda	DepIO	IdentObsVce	Ident	MaxIO	NoCCC	NoObsCoda	NoComplexCoda	NoCoda	MaxObs
☞ markt						*	*	*	*	
mark					*!		*	*	*	*
makt					*!		*	*	*	
market		*!					*		*	
mak					*!*		*		*	*
ma					*!**					**
mar					*!**				*	**

Figure 3-13. GLA predication after 1200 trials

LIST OF REFERENCES

Anderson, J. (1983). The difficulties of English syllable structure for Chinese ESL learners. *Language Learning and Communication*, *2*, 53-61.

Bedore, L. (1999). The acquisition of Spanish. In O. Taylor & L. Leonard (Ed.), *Language acquisition across North America: Cross-cultural and cross-linguistic perspectives* (pp. 157-207). San Diego, CA: Singular.

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

Birdsong, D. (1999). Whys and Why Nots of the Critical Period Hypothesis for Second Language Acquisition. In D. Birdsong (Ed.), *Second Language Acquisition and the Critical Period Hypothesi* (pp. 1-22). Mahway, NJ: Lawrence Erlbaum Associates.

Boersma, P. (1999). Optimality-Theoretic Learning in the PRAAT program. *Proceedings of the Institute of Phonetic Sciences*, *23*, 17-35.

Boersma, P. & Hayes, B. (2001). Empirical tests of the Graduate Learning Algorithm. *Linguistic Inquiry, 32,* 45-86.

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

Broselow, E. (1984). An investigation of transfer in second language acquisition. *International Review of Applied Linguistics, 22*, 253-269.

Broselow, E., Chen, S., & Wang, C. (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition, 20*, 261-280.

Broselow, E., & Xu, Z. (2004). Differential difficulty in the acquisition of second language phonology. *International Journal of English Studies: Advances in Optimality Theory, 4(2),* 135-163.

Cheng, L. (1991). *Assessing Asian language performance: Guidelines for evaluating limited-English proficient students, second edition*. Oceanside, CA: Academic Communication Associates.

Eckman, F.R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning*, *27*, 315-330.

Eckman, F.R. (1981). On predicting phonological difficulty in second language acquisition. *Studies in Second Language Acquisition*, *4*, 18-30.

Gordon, R. (2005). *Ethnologue: Languages of the World, fifteenth edition*. Dallas, TX: SIL International.

Flege, J. (1989). Chinese subjects' perception of the word-final English /t/-/d/contrast: Performance before and after training. *Journal of the Acoustical Society of America*, *86*, 1684-1697.

Flege, J. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 233-277). Baltimore, MD: York Press.

Flege, J. & Davidian, R. (1984). Transfer and developmental processes in adult foreign language speech production. *Applied Psycholinguistics*, *5*, 323-347.

Flege, J., Munro, M., & Skelton, L. (1992). Production of the word-final English /t/-/d/ contrast by native speakers of English, Mandarin and Spanish. *Journal of the Acoustical Society of America*, *92*, 128-143.

Hancin-Bhatt, B. (2000). Optimality in second language phonology: codas in Thai ESL. *Second Language Research, 16*, 201-232.

Hurford, J.R. (1991). The evolution of the critical period for language acquisition. *Cognition*, *40*, 159-201.

Jongstra, W. (2003). Variable and Stable Clusters: Variation in the Realisation of Consonant Clusters. *Canadian Journal of Linguistics*, *48*, 265–288.

Kager, R. (1999). *Optimality Theory*. Cambridge, MA: Cambridge University Press.

Kuhl, P.K. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics, 50*, 93-107.

Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*, 606-608.

Lenneberg, E.H. (1969). *Biological Foundations of Language*. New York, NY: Wiley.

Mayberry, R. (1993). First-language acquisition after childhood differs from second-language acquisition: the case of American Sign Language. *Journal of Speech and Hearing Research*, *36*, 1258-1270.

McLeod, S., Van Doorn, J., & Reed, V. (2001). Consonant cluster development in two-year-olds: general trends and individual difference. *Journal of Speech, Language, and Hearing Reseach*. *44*, 1144-1171.

Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*, *28*, 111-131.

Ohala, Diana (1999). The influence of sonority on children’s cluster reductions. *Journal of Communication Disorders, 32*, 397–422.

Pater, J. and Barlow, J.A. (2003). Constraint conflict in cluster reduction. *Journal of Child Language. 30,* 487-526.

Pilus, Z. (2003). Second language speech: Production and perception of voicing contrasts in word-final obstruents by Malay speakers of English. *Dissertation Abstracts International: The Humanities and Social Sciences*, *63*, 11.

Pinker, S. (1994). *The language instinct: How the mind creates language*. New York, NY: Morrow.

Prince A. & Smolensky, P. (1993). Optimality Theory: Constraint Interaction in Generative Grammar. *Rutgers University Center for Cognitive Science Technical Report 2.*

Roach, P. (2000). *English Phonetics and Phonology: A Practical Course*. Cambridge, MA: Cambridge University Press.

Tarone, E. (1980). Some influences on the syllable structure of interlanguage phonology. *IRAL*, *18*, 139-152.

Tremblay, K., Kraus, N., Carrell, T., & McGee, T. (l997). Central auditory system plasticity: Generalization to novel stimuli following listening training. *Journal of the Acoustical Society of America*, *6*, 3762-3773.

Vanderweide, T. (2005). A Perceptual Analysis of Consonant Cluster Reduction. *Proceedings of the 2005 annual conference of the Canadian Linguistic Association*.

Wang, C. (1995). The acquisition of English word-final obstruents by Chinese speakers. *Dissertation Abstracts International, the Humanities and Social Sciences*, *56*, 10.

Wardrip-Fruin, C. (1982). On the status of temporal cues to phonetic categories: preceding vowel duration as a cue to voicing in final stop categories. *Journal of the Acoustic Society of America*, *71*, 187-195.

White, L., & Genesee, F. (1996). How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research*, *12*, 238-265.

Wiltshire, C. (2006). Word-final Consonant and Cluster Acquisition in Indian English(es). *A Supplement to the Proceedings of the 30th Boston University Conference on Language Development.*

Yoo, H. (2004). A Longitudinal study of consonant cluster acquisition. *Studies in Phonetics, Phonology and Morphology, 10.3*. 481-503.

BIOGRAPHICAL SKETCH

Milla Chappell is a graduate student in Linguistics at the University of Florida who is interested in second language acquisition, particularly that of phonology and phonetics. Milla graduated in 2006 with an M.A. in English from Auburn University and began her studies in linguistics at the University of Florida in August 2006.