MARKEDNESS IN THE PERCEPTION OF L2 ENGLISH CONSONANT CLUSTERS

By

Mahmoud S. AlMahmoud

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Linguistics

2011

ABSTRACT

MARKEDNESS IN THE PERCEPTION OF L2 ENGLISH CONSONANT CLUSTERS

By

Mahmoud S. AlMahmoud

The central goal of this dissertation is to explore the relative perceptibility of vowel epenthesis in English onset clusters by second language learners whose native language is averse to onset clusters. The dissertation examines how audible vowel epenthesis in different onset clusters is, whether this perceptibility varies from one cluster to another, whether an auditory bias toward anaptyxis vs. prothesis, and vice versa, in the perception of vowel epenthesis location exists, and the extent to which each of these inquiries depend on cluster type as a factor.

The dissertation reports on four experiments. Experiment 1 explores Saudi Arabian (SA) listeners' perception of vowel epenthesis in tautosyllabic English onset clusters. The findings suggest that SA listeners are sensitive to the type of cluster in question when perceiving vowel epenthesis, and a hierarchy of perceptual difficulty among different onset clusters is motivated empirically as well as phonetically. Experiment 2 investigates the relationship between vowel length and non-native listeners' ability to perceive vowel epenthesis in different onset clusters. It evaluates listeners' aural sensitivity to the epenthetic vowel along a 5-step duration continuum. Results indicate that duration of the epenthetic vowel as well as type of cluster have a significant effect on listeners' ability to discriminate stimuli correctly. The interaction between duration and cluster type is hardly significant, however. Experiment 3 tests the hypothesis that the choice between perceptual anaptyxis and prothesis is cluster-determined by having

listeners discriminate clusters from their anaptyctic and prothetic forms. Results, however, provide partial evidence for this claim. Experiment 4 tests the perceptual hierarchy of difficulty of Experiment 1 in production. The findings suggest that, like perception, L2 learners produce onset clusters variably as a function of cluster type, suggesting a strong link between L2 perception and production of English consonant clusters.

The formal analysis of the data draws on principles of the Optimality Theory (Prince and Smolensky 1993) and P-map (Steriade 2009) in accounting for SA learners' vowel epenthesis perceptual patterns in onset clusters. Perceptual distinctiveness scales that reflect relative perceptibility of vowel epenthesis and its location asymmetry in onset clusters are projected into context-sensitive faithfulness constraints. It is argued that the ranking of the markedness constraint *COMPLEX^{Ons} relative to DEP-V/A+B, and ONSET relative to DEP-V/A_B capture zero-epenthesis and anaptyxis-prothesis asymmetries, respectively.

ACKNOWLEDGEMENTS

First and foremost, I thank the Almighty for everything. Second, I owe my deepest gratitude to a great many people whose assistance and advice have culminated in the production of this work. My advisor, Yen-Hwei Lin, merits my thanks for her continued help, insightful suggestions and guidance not only in the writing of this dissertation but also in helping me navigate through the linguistics program at Michigan State University. I was indeed fortunate to have her advisorship.

I am also enormously appreciative of the assistance I received from Grover Hudson. His profound knowledge of linguistics, long experience in the field and willingness to generously and selflessly allocate time to review my research made this dissertation possible. Thanks are also due to the rest of my dissertation committee, Deogratias Ngonyani and Anne Violin-Wigent, for their ongoing encouragement and constructive feedback, which has been instrumental in completing the dissertation. I consider myself amazingly fortunate to have had the guidance and wisdom of such a committee.

I would like to acknowledge the financial support I received during my graduate studies. I am greatly indebted to a scholarship from the College of Languages and Translation at Imam Muhammad Ibn Saud Islamic University to pursue my higher education. Many thanks to the Georgetown University graduate school for awarding me a teaching assistantship during the course of my M.S. degree there. I also wish to acknowledge the financial aid I received from Michigan State University, the College of Arts and Letters 2009 Summer Fellowship.

Finally, I would like to thank my parents, siblings, and my in-laws for their patience and encouragement. Last but not least, I am for ever grateful to my wife, Arwa, and our two kids, Sulieman and Reema, for their moral support, solace and endurance over the years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1. Overview of the Dissertation

This dissertation is part of the general psychoacoustic program of research on the perception of non-native phonotactic and syllable structure constraints. The dissertation explores how second language (L2) listeners perceive onset consonant clusters that are illicit in their native language. More particularly, it focuses on the aural perception of tautosyllabic onset clusters by Saudi Arabian (SA) learners of English whose first language (L1) is averse to onset clusters. Previous research has shown that knowledge of the native language phonotactics and syllable structure can affect how L2 listeners perceive and process non-native sequences; L2 listeners tend to hear an illusory epenthetic vowel within consonant clusters that violate native phonotactic constraints (e.g. Dupoux et al. 1999, Kabak and Idsardi 2007, Polivanov 1931). Regarding production, research has also shown that L2 learners produce consonant clusters with a varying degree of difficulty that can be attributed to cluster type (Broselow and Finer 1991, Davidson 2006, Kwon 2005, Major 1996). In this dissertation, I argue that, like production, L2 perception of onset clusters varies in difficulty from one cluster to another according to cluster type.

Empirically, the dissertation attempts to arrive at a hierarchal ranking of perceptual difficulty that yields a principled account of SA listeners' ability to identify vowel epenthesis[1] in different onset clusters. The dissertation also assesses the role of second language proficiency in the acquisition of onset clusters by examining the L2

[1] 'Vowel epenthesis' is used here as a cover term for anaptyxis and prothesis, vowel insertion splitting and prefixing the cluster, respectively.

performance of SA learners with different proficiency levels. To this end, a series of experiments examine how beginning and advanced SA learners of English perceive and process vowel epenthesis as it applies to different English onset clusters.

In Experiment 1, the general hypothesis that perceptibility of vowel epenthesis varies as a function of cluster type as well as linguistic proficiency is tested. Although no significant role of proficiency was found, the findings yielded the following 4-place scale of perceptual difficulty:

(1) Perceptual hierarchy of difficulty among onset clusters

less difficult ←——————→ *more difficult*

[sibilant+stop] >> [sibilant+sonorant] >> [obstruent+liquid] >> [obstruent+glide]

The scale in (1) reflects SA listeners' different perceptions of vowel epenthesis in different onset clusters; vowel epenthesis was least perceptible in [obstruent+glide], more perceptible in [obstruent+liquid], even more perceptible in [sibilant+sonorant], and most perceptible in [sibilant+stop] clusters (see §4.2.4.2). Experiment 1 has also examined whether a cluster-induced bias toward hearing anaptyxis vs. prothesis exists. Results show that SA listeners tended to associate anaptyxis with [obstruent+glide] clusters, and prothesis with [sibilant+stop] clusters. Evidence from phonetics and empirical research on consonant clusters is used to justify the proposed hierarchy as well as this cluster-based difference of anaptyxis and prothesis which SA listeners displayed in their perception of vowel epenthesis (see §4.3 and §4.4).

The perceptual difficulty scale motivated by the results of Experiment 1 is further tested by three follow-up experiments. The first of these probes listeners' sensitivity to the presence of vowel epenthesis along a durational continuum. In this experiment, it is hypothesized, after Dupoux et al. (1999), that the longer the duration of the epenthetic vowel, the more likely presence of the vowel will be recognized.

Further, it is assumed that discrimination on the vowel continuum will vary by cluster type. Results from Experiment 2 show that SA listeners were able to systematically better discriminate stimuli the longer the duration of the epenthetic vowel. The results also suggest a relationship between cluster type and listeners' sensitivity to subtle variations in vowel epenthesis length (see §5.2).

Experiment 3 tests the difference in vowel epenthesis location reported in Experiment 1 by exploring SA listeners' judgments of perceived similarity between onset clusters and their epenthetic forms. It assumes that the choice between anaptyxis and prothesis in the perception of vowel epenthesis is cluster-driven (see §5.3). Although auditory bias toward hearing anaptyxis vs. prothesis and vice versa was expressed by beginner as well as advanced listeners, this finding did not generalize to all types of onset clusters (see §5.3.2.5).

Experiment 4 assesses the extent to which the perceptual hierarchy of Experiment 1 is reflected in the speech of SA learners of English. It tests the main hypothesis that SA learners of English would produce onset clusters to a large degree according to their perceptual biases. Results show that in Experiments 1 and 4 certain onset clusters consistently presented more difficulty for SA learners. For example, in both experiments learners' ability to identify vowel epenthesis as well as produce onset clusters correctly was worse in [obstruent+glide] and [obstruent+liquid] relative to [sibilant+sonorant] and [sibilant+stop] clusters (see §5.5).

Formally, the dissertation attempts to account for L2 perceptual patterns of vowel epenthesis within the constraint-based framework of Optimality Theory (OT) (Prince and Smolensky 1993). It is argued that the ranking of the markedness constraint *COMPLEX$^{\text{Ons}}$ relative to the faithfulness constraint DEP-V/A+B accounts for SA learners' zero-epenthesis patterns, and the ranking of ONSET relative to DEP-

V/A_B accounts for the anaptyxis-prothesis asymmetry in the perception of onset clusters (see §6.3).

1.2. Research Questions

The dissertation is situated in L2 research exploring the possible causes of variable success with which non-native speakers of English perceive and produce onset consonant clusters. In particular, the dissertation endeavors to answer the following research questions:

1. How will SA learners of English, whose first language is averse to CC-sequences, perceive English onset consonant clusters? Will their perception be native-like, non-native-like, or variably both?

Research shows that second language learners tend to perceive non-natively prosodic structures which are not in conformity with their native language. Polivanov (1931) was among the earliest to propose that L2 listeners perceive as well as produce non-native utterances through their L1 phonological system. Polivanov argues that for Japanese speakers of English, the native perceptual grammar introduces the percept of the epenthetic vowels [u] and [o] to avoid initial consonant clusters such as [dr-] in *drama*. More recently, Dupoux et al. (1999), who compare French and Japanese listeners' perception of medial consonant clusters (VCCV), argue that Japanese, but not French, listeners have a strong tendency to perceive vowel epenthesis amidst nonsense words such as *ebzo* → *eb<u>u</u>zo* (vowel epenthesis underlined). This phenomenon of 'perceptual vowel illusion' which interferes with listeners' ability to correctly perceive non-native consonant clusters has been well-established (see Berent et al. 2007, Chang et al. 2007, Kabak and Idsardi 2007, Moreton 2002). These studies demonstrate that native language constraints on syllable structure can exert pressure on the listener's perceptual system, giving rise to the illusory percept of vowel epenthesis in illicit consonant clusters. Thus, for the first question addressed by this dissertation, research seems to suggest that perception of English onset clusters by SA listeners would be non-native with vowel epenthesis being commonly perceived as part of the cluster. In other words, SA listeners are predicted to show variability in their perception of onset clusters, alternating between presence and absence of epenthesis.

2. Will SA listeners' perception of vowel epenthesis vary from one cluster to another? If so, can this relative perceptibility of vowel epenthesis be systematically attributed to cluster type?

The second research question is more central to the dissertation. It concerns relative perceptual difficulty among onset clusters, assessed here by listeners' failure to identify vowel epenthesis in some onset clusters, but not others. L2 research indicates that learners' ability to realize consonant clusters natively varies, depending on the type of cluster acquired (Broselow and Finer 1991, Davidson 2006). Major (1996), for example, demonstrates experimentally how Brazilian Portuguese speakers of English produced the least error rate in onset [sibilant+stop] clusters (e.g. [st-]) relative to other types of clusters. Kwon (2005) shows that [stop+liquid] clusters (e.g. [br-]) are more likely to result in vowel epenthesis than [sibilant+stop] clusters in the speech of Korean learners of English. Berent et al. (2007) demonstrate that English listeners are perceptually sensitive to differences in sonority among onset clusters in that they report hearing vowel epenthesis more in onset clusters with sonority falls (e.g. [lb-]), than in plateaus (e.g. [bd-]), and even less in clusters involving sonority rises (e.g. [bn-]).

Several factors have been put forth as explanation for this variable performance in the acquisition of consonant clusters. These include typological markedness (Eckman 1987, Morelli 2003), sonority (Broselow and Finer 1991), structural differences (Broselow 1993, Giegerich 1992), articulatory ease and/or perceptual contrast maximization (Flemming 1995, Steriade 2009, Zuraw 2007), and frequency (Frisch et al. 2000). Thus, the main question seems to be why some clusters appear to be less subject to epenthesis than others and how, depending on the cluster type, different repair methods are called for? While many L2 studies have examined the issue of relative difficulty in the production of consonant clusters from different perspectives, it remains largely unaddressed in L2 perception. However, evidence from production seems to suggest that perception of L2 onset clusters by SA listeners should vary according to cluster type.

3. How sensitive are SA listeners to the presence/absence of the inserted vowel in onset clusters? Does the degree of sensitivity vary according to vowel length, cluster type or both?

Dupoux et al. (1999) report that their Japanese as well as French listeners systematically heard more instances of vowel epenthesis as the duration of the inserted vowel increased. Presented with a six-step vowel continuum that ranged from no vowel VCCV (e.g. *ebzo*) to full vowel VC<u>V</u>CV (e.g. *eb<u>u</u>zo*), both Japanese and French listeners were better able to identify vowel epenthesis in longer durations, although the effect of vowel length was much clearer in French listeners. This suggests that SA listeners would perceive vowel epenthesis better, the longer the duration of the inserted vowel. While in the Dupoux et al. study sensitivity to vowel epenthesis differed along the dimensions of vowel length and listener group (i.e. Japanese vs. French), the interaction between vowel length and cluster type remains untested. The research question here attempts to address this issue.

4. Will SA listeners demonstrate a bias toward anaptyxis vs. prothesis, and vice versa, in their perception of vowel epenthesis location? If so, will this asymmetry in their performance be cluster-dependent?

The fourth question is related to the location of vowel epenthesis in a cluster. Suggestive evidence for vowel epenthesis place asymmetries comes mainly from studies that examined cluster production. No study, to my knowledge, has investigated vowel epenthesis location asymmetries in L2 perception. In loanword phonology, Fleischhacker (2001, 2005) justifies anaptyxis-prothesis asymmetries in onset clusters by appealing to perceptual similarity between the cluster and its epenthetic form as a determining factor in vowel epenthesis site. She argues that the cluster-dependent preponderance of anaptyxis in [obstruent+sonorant] and prothesis in [sibilant+stop] clusters derives from maximizing auditory similarity between the cluster and its anaptyctic or prothetic form. Both similarity as well as preference judgments elicited from her native English speakers supported this claim.

Kwon (2005) refers to the same notion of perceptual similarity in explaining patterns of vowel epenthesis in the speech of her Korean learners of English. Although Kwon did not examine the asymmetry in vowel epenthesis location per se, her findings suggest that vowel epenthesis can be predicted on perceptual grounds. She argues that the likelihood of vowel insertion in producing a cluster follows from the context in which vowel epenthesis takes place: vowel epenthesis is more likely to occur in contexts where it is less obtrusive perceptually. She explains how vowel epenthesis was common in clusters which were judged more similar perceptually to their anaptyctic forms by Korean speakers, and vice versa. These studies suggest that SA learners of English in this dissertation would associate anaptyxis with some clusters, but prothesis with others depending on the type of cluster in question.

5. Will perception of onset clusters be better in higher proficiency learners of English, compared to lower proficiency learners?

This question has to do with L2 perceptual development, and is based on research that has established an effect of target language proficiency on L2 acquisition (Abrahamsson 2003, Riney and Flege 1998). Bradlow et al. (1997) and Flege (1995) show that higher proficiency can lead to improved L2 pronunciation, suggesting that in this dissertation advanced SA learners will outstrip beginning learners. Broadly speaking, this suggestion challenges claims of the classical Critical Period Hypothesis that place age limitations on the mastery of L2.

6. How closely do L2 learners' perception and production of consonant clusters correspond?

The relationship between perception and production has long been an issue in the study of L2 speech with research emphasizing a strong link between learners' ability to perceive and produce a second language (Barry 1989, Borden et al. 1983, Grassegger 1991, Rochet 1995, Yeon 2003). I delay discussion of this until Chapter 3 of the dissertation. But if indeed this is the case, then it is expected that SA learners would demonstrate some regular correspondence in their perception and production of onset clusters. In Chapter 5, I examine this claim by comparing the perceptual and productive performance of SA learners in producing the same set of onset clusters.

7. Can the variable degrees of difficulty in the perception of English onset clusters by SA listeners be accounted for in formal phonological theory?

Previous research has shown that the interaction between certain markedness and context-sensitive faithfulness constraints in OT (Prince and Smolensky 1993) can provide a principled account of the different vowel epenthesis patterns that exist as part of loanword phonology (e.g. Fleischhacker 2005, Gouskova 2001) and in the speech of L2 learners (Davidson 2001, 2003, Kwon 2006). I devote Chapter 6 to discussion of this question.

1.3. Dissertation Outline

The rest of the dissertation is organized as follows. In Chapter 2, I review the literature on the second language acquisition of consonant clusters. I limit the review to some of the main approaches to the acquisition of L2 consonant clusters such as L1 transfer, linguistic universals (e.g. typological markedness and sonority), structural and frequency accounts, as well as perceptual and articulatory factors.

In Chapter 3, I introduce some of the key concepts and issues in the study of native and non-native speech perception. After briefly reviewing some of the properties involved in speech analysis, I consider how the native language delimits listeners' ability to perceive non-native segments and phonotactic strings. I also discuss in this chapter the relationship between perception and production of L2 sounds.

Chapter 4 motivates a hierarchy of perceptual difficulty among different onset clusters. In Chapter 4, Experiment 1 provides empirical basis for establishing this scale of difficulty by exploring SA listeners' sensitivity toward vowel epenthesis and its location in different onset clusters. I also argue for this hierarchy by presenting phonetic evidence and further empirical findings from other studies on consonant clusters.

Chapter 5 reports on three follow-up experiments which test the validity of the perceptual hierarchy developed in Chapter 4. Experiment 2 evaluates the hierarchy by assessing listeners' ability to perceive vowel epenthesis along a vowel duration continuum. Experiment 3 explores perceptual anaptyxis-prothesis differences by asking SA participants to discriminate clusters from their epenthetic forms. In

Experiment 4, a production experiment tests the predictions of the perceptual difficulty hierarchy in the speech of SA learners.

Chapter 6 presents a formal OT analysis (Prince and Smolensky 1993) of the vowel epenthesis patterns exhibited by SA learners in their perception of English onset clusters. The chapter shows how the interaction between syllable well-formedness and context-sensitive faithfulness constraints (Steriade 2009) yields an array of possible grammatical patterns which include those of SA learners. I also discuss in this chapter how L2 learning takes place under the framework of OT, and point to some of the typological implications of the analysis.

In Chapter 7, I present a summary of the main issues in the dissertation. I highlight the empirical contributions and theoretical implications this dissertation makes. The chapter concludes with a discussion of limitations of the study, suggesting further topics for future research.

Chapter 2

Factors in the Acquisition of L2 Consonant Clusters

2.1. Introduction

While early research attributes errors in the acquisition (i.e. perception and production) of second language consonant clusters to negative transfer from the learner's native language, the role of linguistic universals, in terms of typological markedness and sonority, has been influential. The study of syllable structure has also shed some light on how L2 learners process and produce non-native sequences. However, recent investigation of syllable structure has sparked interest in a rather different phonetically-based approach.

In this chapter, I survey the literature on the acquisition of L2 consonant clusters. The review is limited to discussing studies relevant to the role of L1, linguistic universals, syllable structure, frequency, and phonetics, all of which have been shown to affect L2 learners' ability to acquire non-native clusters. It should be noted that the majority of these studies are concerned more or less with production and only very few deal with perception.

2.2. L1 Transfer

Early on, researchers noted the role of the native language in the acquisition of L2 speech sounds and structures. Polivanov (1931) and Trubetzkoy (1939) were among the earliest to view the phonological system of L1 as being a filter through which all of the L2 sounds and structures are perceived, categorized and ultimately produced (Broselow 1984, Hodne 1985, Tarone 1987). One formalization of this view is Lado's (1957) Contrastive Analysis Hypothesis (CAH), which maintains that difficulty in the acquisition of a second language can well be predicted on the basis of L1 transfer:

only L2 sounds or structures different from those of the L1 pose difficulty for the learner.

Evidence in testing the predictions of the CAH exists although primarily limited in its nature to L2 segmental production (phonemes and allophones) rather than syllable structure (Anderson 1983, Oller 1972). Nonetheless, there have been some transfer-based explanations of errors in the production of L2 syllables. Broselow (1984) attributes prosodic errors in the speech of American learners of Arabic and Egyptian learners of English to negative transfer from the syllabification system of the learner's native language. Broselow (1987) compares epenthesis errors made by speakers of two Arabic dialects, Iraqi and Egyptian, neither of which allows onset clusters. Both Iraqi and Egyptian speakers exhibited patterns of vowel epenthesis when producing English clusters that are characteristic of epenthesis rules in their native dialects. Consequently, Broselow argues that L2 speakers transfer their L1 prosodic structure conditioning constraints over to the new language they are learning (also Broselow 1993).

However, in its strong predictive form, the validity of the CAH has been called into question as it fails to explain why some L1-L2 differences do not lead to negative transfer and why, on the other hand, some L1-L2 similarities still pose a great deal of difficulty for L2 learners (Eckman 1987, Towell and Hawkins 1994, Wardhaugh 1970). This same argument that led to the demise of the strong form of the CAH has encouraged, however, another interpretation of L1 transfer: similarities between the L1 and L2 give rise to L2 difficulty while differences fail to promote difficulty. This view of L1 transfer reverses the roles of positive and negative transfer previously held by the CAH, and is perhaps best captured by Flege's (1987, 1995) acclaimed Speech Learning Model (SLM). According to the SLM, the ability to perceive and produce L2 sounds in a native-like manner is contingent upon the establishment of separate abstract phonetic categories for these sounds.[2] Category formation in the end is regulated by the relationship of similarity between L1 and L2 sounds: ‘the greater the perceived distance of an L2 sound from its closest L1 sound, the more likely it is that a separate category will be established for the L2 sound’ (Flege 1987: 264). See Major and Kim (1996) and Best (1995) for other related similarity-based claims.

The past decade has seen a resurgence of the notion of transfer. For example, Brown (1997), who develops a generative model of speech perception based on the phonological features shared by the learners’ native and target language, claims that the learner transfers the L1 distinctive feature(s) when encountering a new language and whatever difficulty faced in the production (or perception) of a certain L2 contrast is the result of the L1 lacking the needed feature (or features) to discriminate that contrast. Larson-Hall (2004) explains that Brown’s ‘Featural Model’ views the process of L2 phonological acquisition as involving a level of ‘feature organization’.

By no means downplaying the role of L1 transfer, the following section sheds some light on the role of universal linguistic factors in explaining claims of commonalities and variation as far as the acquisition of L2 consonant clusters is concerned.

2.3. Role of Linguistic Universals

Here, I consider two aspects of universals as they relate to the acquisition of consonant clusters: typological markedness and effects of sonority on segmental sequencing.

---

[2] Flege (1995: 239) defines ‘phonetic categories’ as ‘language specific aspects of speech sounds [which] are specified in long term memory representations’.

2.3.1. Typological Markedness

The concept of typological markedness has been thoroughly developed in Greenberg 1976, 1978, and defined by Gundel et al. (1986: 108) as:

> A structure X is typologically marked relative to another structure, Y, (and Y is typologically unmarked relative to X) if every language that has X also has Y, but every language that has Y does not necessarily have X.

Embedded in Gundel et al.'s definition is the implicational relationship between X and Y: the presence of X implies the presence of Y, but not necessarily the opposite. Markedness then obtains both within language and cross-linguistically as a result of the distribution of the two members in question: a marked structure or linguistic feature is one that is less commonly found across languages.[3]

Eckman's (1987) Markedness Differential Hypothesis (MDH) incorporates this notion of typological markedness[4] to address some of the issues Lado's (1957) CAH came short of explaining, such as the directionality of difficulty in the English-German final obstruent voicing contrast. The fact that German allows only voiceless obstruents word-finally is expected, under the CAH, to present a learning problem for English as well as German speakers since it is an area of difference between the two languages. Eckman (1987) argues, however, that only German learners of English suffer from this incompatibility between the two languages while English learners of German show little or no difficulty.

Whereas any structural discrepancy between two languages is deemed problematic in CAH, for the MDH such cross-linguistic structural differences will

[3] Underlying this definition, of course, is the assumption that frequency of occurrence is directly related to unmarkedness or naturalness.

[4] Eckman (1987: 60) defines markedness as follows: 'a phenomenon A in some language is more marked than B if the presence of A in a language implies the presence of B; but the presence of B does not imply the presence of A.'

only lead to learning problems if they are more marked typologically in the target language relative to the learner's native language. For the acquisition of L2 consonant clusters, difficulty would arise then if the L2 has a different *and* typologically more marked structure than the L1. L2 learners normally experience difficulty with complex onsets when only simplex onsets are sanctioned by their L1. Underlying this observation is a typological universal which states that simplex onsets are less marked than complex onsets (Greenberg 1978).

The MDH, as stated in Eckman (1987: 86), can also explain the degree of difficulty among two (or more) L2 structures.[5] A good example of this is the relationship found between cluster length and markedness in L2 where triliteral clusters are considered marked relative to biliteral clusters, and biliteral clusters are marked relative to singletons (Cairns and Feinstein 1982). These markedness claims are based on the implicational typological universal: the presence of triliteral clusters implicates the presence of biliteral clusters, and the presence of biliteral clusters entails the presence of singletons[6] (Greenberg 1978, Kaye and Lowenstamm 1981). We expect, therefore, to see a linear relationship between number of consonants and difficulty (i.e. increased error rate or delayed production) in the acquisition of consonant clusters, given that markedness corresponds to difficulty (Eckman 1987: 60, Eckman and Iverson 1993: 241).

---

[5] Note that the markedness relationship in this case obtains in the target language independent of the L1. Carlisle (1997, 1998) claims this is the most commonly tested form of the MDH, or the Interlanguage Structural Conformity Hypothesis (ISCH). Therefore, it seems two scenarios are possible: two structures in the L2, one of which is marked and the other unmarked, and either the L1 contains the unmarked structure only (e.g. Osburne 1996, Sato 1984), or the L1 contains neither structure (e.g. Anderson 1987, Carlisle 1988). See Carlisle (1998: 257) for further discussion.

[6] Note, however, that the presence of singleton onsets (CV-) does not imply the presence of zero-onset syllables (V-) (Vennemann 1988).

Available evidence seems to be supportive of this prediction. Sato (1984) and Osburne (1996) show longitudinally a clear preference for singleton over biliteral syllable margins by analyzing production errors in the English speech of their Vietnamese subjects. Anderson (1987) examines the production of English consonant clusters by Chinese and Egyptian speakers and finds that both have more errors concerning triliteral than biliteral clusters. Carlisle (1997) examines 11 Spanish speakers on the production of biliteral and triliteral onset clusters to conclude that biliteral clusters are significantly less frequently modified than triliteral clusters. Other examples of this kind of support for the MDH are reported by Benson (1988), Carlisle (1998), Eckman (1991), Eckman and Iverson (1993), and Weinberger (1987).

While a typology-based account alone might explain the difficulty associated with onset length, its success in predicting markedness relations holding within biliteral or triliteral consonant clusters is less transparent. For one thing, a typology account by itself is limited in that it can make claims about the individual markedness of one or more segments in a cluster, but not the entire consonantal string. To illustrate, typologically voiced stops are believed to be marked relative to voiceless stops, and fricatives to stops (Eckman and Iverson 1993). Hence, [p] in the onset cluster [pr-] is considered unmarked relative to [b] in [br-] on the basis of voicing, and to [f] in [fr-] on the basis of manner. However, this claim says nothing about the (un)markedness of the cluster as a whole. Eckman and Iverson attempt to overcome this problem by appealing to Clements' (1990: 313) Sequential Markedness Principle (SMP):

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

According to the SMP, the preference of [pr-] over [br-] and [fr-] is established by the relative markedness of [b] and [f] over [p]. Thus, the SMP provides a way of linking the markedness of sequentially individual segments to a larger constituent, the cluster.

In her dissertation, Morelli (2003) takes a slightly different approach to consonant cluster typological markedness in which the distribution of the cluster as a whole, rather than its constituent members, is considered. Focusing on obstruent onset clusters in 30 unrelated languages, Morelli demonstrates that only 6 patterns are attested in these languages in which onset [fricative+stop], [stop+fricative], [stop+stop], [fricative+fricative] can (co)occur:

	[fricative+stop]	[stop+fricative]	[stop+stop]	[fricative+fricative]
Type 1	X			
Type 2	X			X
Type 3	X	X		
Type 4	X	X		X
Type 5	X	X	X	
Type 6	X	X	X	X

Table 2.1 Typology of obstruent onset clusters (Morelli 2003: 42)

In Table 2.1, Type 1 languages are most restrictive (e.g. English) allowing only one class of obstruent clusters, namely [sibilant+stop], while languages of Type 6 are least restrictive with all four possible classes of obstruent clusters occurring in the language (e.g. Georgian). The remaining types represent languages that vary in their permissibility of obstruent clusters.

Crucial are the observations that [fricative+stop] clusters are (1) the only clusters present in all 6 patterns attested in the languages sampled, and (2) the presence of any other obstruent cluster always entails their presence, as is clear from Table 2.1. Morelli (2003) thus argues on implicational typological grounds that a sequence of [fricative+stop] universally forms the least marked cluster within the domain of onset obstruent clusters. Morelli claims this is due to [fricative+stop] clusters being less marked with regard to the manner feature [continuant] and the place feature [coronal], since the first member of the cluster can only be the continuant coronal /s/. The following scale of markedness summarizes the findings of Morelli with regard to obstruent onset clusters (clusters to the right of >> are more marked):

(1) [fricative+stop] >> [stop+fricative] >> [stop+stop], [fricative+fricative][7]

However, typological universals, implicational or not, built on generalizable observations of certain linguistic phenomena among the world's languages are but synchronic descriptions of the languages' phonologies. That is, they provide formal statements of commonly shared linguistic properties and expressions, but offer little or no explanation of the (un)markedness of the phenomenon in question (although see Morelli (2003) who invokes phonetic factors besides typology).

Much of the criticism of a typological markedness approach comes from the proponents of Universal Grammar (UG). Archibald (1998: 150) notes that markedness claims based on typological generalizations are descriptive of facts which themselves are in need of explanation:

> … this sort of typological universals … provides an interesting description of the phenomenon to be explained. I'm less sure of their [sic] status as an explanation of the observed facts.

Similarly, Gass and Selinker (2001: 154) critique the typological markedness account (see also White 1987):

> It is not sufficient to state that second languages obey natural language

---

[7] Based on the typology of obstruent onset clusters surveyed, no markedness relation between [fricative+fricative] and [stop+fricative] or [stop+stop] clusters could be established.

constraints because that is the way languages are. This only pushes the problem of explanation back one step.

The sine qua non of the argument against a typological account of markedness is that in pointing out that some linguistic features or patterns are unmarked *only* because they occur more frequently among the world's languages, little explanation is given as to why unmarked forms are more pervasive than marked forms, and further, why frequency of occurrence should be associated with (un)markedness to begin with (however, see Eckman's 2004 rebuttal of these claims). As will be discussed in §2.4 and §2.6, other functional (e.g. structural or etiological) factors may be invoked when interpreting the notion of phonological markedness in consonant clusters.

2.3.2. Role of Sonority

Sonority has long been proposed as a principle governing syllable well-formedness and, more recently, a determinant of the errors non-native speakers make when producing L2 consonant clusters. The cross-linguistic generalization that most sonorous segments form the nucleus of the syllable and that sonority should gradually decrease in segments as they move away from the nucleus was observed as early as Sievers (1881). Ever since, the idea has demanded increasing interest and has been called the Sonority Sequencing Principle (SSP) (Clements 1990, Selkirk 1982, Steriade 1982).

The SSP requires sonority in segments to rise in the onset toward the syllable peak (nucleus) and fall in the coda toward the trough. One main assumption of the SSP is that segments are inherently distinct in their sonority. A scale showing relative sonority levels among segments is therefore essential for the SSP to function; one scale which ranks segments from most sonorous (left) to least sonorous (right) is suggested by Clements (1990):

(2) Vowels > Glides > Liquids > Nasals > Obstruents

More distinctively, obstruents typically break down to stops and fricatives, with the latter being more sonorous (Broselow and Finer 1991).[8]

Singh (1985) in discussing Hindi English argues the position of the epenthetic vowel can well be determined by the sonority profile of the consonant cluster in question: anaptyxis splits rising sonority clusters such as /gret/ → [gəret] 'great', while prothesis prefixes non-rising clusters as in /skul/ → [əskul] 'school'. A constraint-based approach such as OT (Prince and Smolensky 1993) captures Singh's generalization on vowel epenthesis location with the sonority sequencing constraint SYLLABLE CONTACT (Murray and Vennemann 1983).[9] For example, Gouskova (2001) argues that the ranking of SYLLABLE CONTACT and faithfulness CONTIGUITY over the markedness constraint *COMPLEX is what drives anaptyxis-prothesis asymmetries in Kirgiz loanwords from Russian (see also Morelli 1999):

(3) SYLLABLE CONTACT: a syllable onset should be less sonorous than the preceding coda.

CONTIGUITY: input segments should be contiguous at output.

ONSET: no onsetless syllables.

*COMPLEX: onsets and codas are simple.

While anaptyxis is enforced through the ranking of SYLLABLE CONTACT over CONTIGUITY, the latter imposes prothesis when SYLLABLE CONTACT is respected as in tableau (4):

---

[8] Other more refined sonority scales have been proposed in the literature (e.g. Selkirk 1984).

[9] In Chapter 6 I discuss in more detail the framework of OT and show how it can be used to formally account for the data.

(4) Anaptyxis in /brek/ 'brake' and prothesis in /stek/ 'steak'

	/brek/			/stek/		
	☞bərek	əbrek	brek	sətek	☞əstek	stek
*COMPLEX			*!			*!
SYLL-CONTACT		*!				
CONTIGUITY	*			*!		
ONSET		*			*	

In (4), the faithful outputs with onset clusters are eliminated by high ranking *COMPLEX. The candidate with prothesis [əb.rek] is penalized by SYLLABLE CONTACT because prothesis in [əb.rek] has resulted in onset /r/ which is more sonorous than the preceding coda /b/; this leaves the anaptyctic output [bə.rek] as the optimal one. On the contrary, SYLLABLE CONTACT is not violated in [əs.tek] since /t/ is less sonorous than /s/; as a result prothesis emerges as the winning candidate due to CONTIGUITY.

Broselow and Finer (1991) propose on sonority grounds a multi-valued Minimal Sonority Distance (MSD) Parameter in L2 syllable acquisition. The model, they argue, determines the set of consonant clusters allowable in an individual language based on sonority differences between adjacent segments. The MSD has several parameter settings that are specified by the sonority index given in (5):

(5) Sonority Index (Broselow and Finer 1991: 38)

Class	**Value**
Stops	1
Fricatives	2
Nasals	3
Liquids	4
Glides	5

The number of clusters allowable in a given language is inversely related to the degree of sonority difference between adjacent segments: languages requiring greater sonority differences between onset segments have very few types of onset clusters while languages requiring lesser sonority differences allow more types of onset clusters. With stops being the least sonorous and glides being the most, an MSD setting of 5 yields a language that has no consonant clusters since the largest possible difference between the values in the sonority index is 4; an MSD setting of 4 permits [stop+glide] clusters in addition to singleton onsets while a language that has a minimal sonority difference of 3 can possibly yield [stop+glide], [stop+liquid] and [fricative+glide] clusters. And a language with a setting of 1 should theoretically allow any cluster whose members differ minimally by 1 as well as all clusters permitted by settings 5, 4, 3 and 2.

Further, Broselow and Finer (1991) claim that the higher the MSD setting of a consonant string, the easier it is to acquire, and vice versa. Thus, the model makes a prediction regarding relative markedness or difficulty across different consonant cluster types: a consonant sequence with a low MSD setting will be more difficult to acquire than a consonant sequence that has a higher MSD setting. For instance, a [stop+liquid] onset cluster has a minimal sonority difference of 3 and is expected, therefore, to be more easily acquired than, say, a [fricative+liquid] cluster, which has an MSD setting of 2.

Broselow and Finer (1991) maintain that cross-linguistically different languages may demonstrate various MSD parameter settings. Given this assumption, the acquisition of different L2 onset clusters becomes regulated by L1 sequencing constraints on segments. In other words, when acquiring L2 consonant clusters the native language's MSD parameter values are transferred over to the L2. Only when the MSD values of the native language are higher than those of the target language will L2 clusters be problematic for the learner. Consider, for example, the case of Arabic and English. Arabic is averse to onset clusters and is believed to have an MSD setting 5, whereas English has a lower MSD parameter setting 1 that permits a wide choice of onset clusters. According to the MSD, difficulty for speakers of Arabic in learning English onset clusters originates from the higher MSD setting Arabic has over English. Broselow and Finer note that given sufficient positive evidence, the initial parameter settings, transferred from the L1, can eventually be reset to match or at least approximate the settings of the target language.

Empirically, the validity of the MSD as a model for L2 acquisition of consonant clusters has been tested by Broselow and Finer (1991) who examine the production of English onset clusters by 32 Japanese and Korean speakers.[10] The predictions of the MSD were generally upheld in that lower-MSD clusters turned out to be more challenging for subjects than clusters with a relatively higher MSD; fewer errors were incurred in [stop+glide] (e.g. [pj-]) vs. [stop+liquid] (e.g. [pr-]) onsets (6 vs. 12), but more errors in [fricative+liquid] (e.g. [fr-]) vs. [stop+liquid] (e.g. [br-]) clusters (18 vs. 12) (however, see Kwon 2006: 33 for a reinterpretation of the Korean results).

Moreover, Hancin-Bhatt and Bhatt (1997) account for L1 Japanese and Spanish speakers' production of English consonant clusters in light of the MSD. They hypothesize that Spanish speakers, due to their low L1 MSD setting 2, would commit fewer mistakes on [fricative+liquid] and [stop+liquid] clusters than would Japanese speakers whose L1 has a setting of 5. Their findings generally support the MSD model in that Spanish speakers outperformed Japanese speakers on the production of [fricative+liquid] and [stop+liquid] clusters.

---

[10] I am unaware of any study that tests the MSD predictions on L2 consonant cluster perception.

To sum up, the MSD parameter is a model for the acquisition of consonant clusters. In addition to the notion of language transfer, it employs a pre-determined index of different sonority settings in predicting consonant cluster difficulty.

Although widely accepted and although it has captured some of the universal generalizations on the organization of segments in syllables, a sonority-based account is not without its problems. The question of what sonority is has been highly controversial; it has been defined as perceptual salience or loudness of a particular sound (Ladefoged 2001), the amount or volume of airflow in the resonance chamber (Goldsmith 1995), and the inherent degree of stricture (Lindblom 1983) (see Wright 2004 for a more recent proposal incorporating perceptual elements in the definition of sonority). Even more controversial is the issue of sonority scales and how they should be constructed and interpreted. Clements (1990) and Selkirk (1984), despite proposing slightly different scales of sonority, claim they are universally shared by all languages. Steriade (1982) and Davis (1990), on the other hand, believe sonority scales should reflect language-specific tendencies in assigning sonority values to their segments.

Eckman and Iverson (1993) criticize the MSD on the grounds that it takes only sonority sequencing as the sole determiner of consonant cluster difficulty. They argue that Clements' (1990) Sonority Dispersion Principle (SDP) along with typological markedness are sufficient to predict which clusters are likely to undergo more restructuring than others without the need for positing L1 transfer of MSD settings. Based on evidence from Korean and Japanese, Eckman and Iverson (239) observe how the MSD over-generates typologically unattested sequences; the onset sequence [nasal+glide] in Korean [myən] 'cotton' and Japanese [myo] 'strange' indicates that both languages have the MSD setting 2. Predictably, other consonant clusters such as [fricative+liquid] and [stop+liquid] would be sanctioned by this low setting. Nonetheless, Eckman and Iverson note that no such consonant clusters exist in Korean or Japanese.

The MSD's predictions with regard to the relative difficulty of some clusters over others have also been called into question empirically. Because onset [stop+liquid] clusters are believed to have a minimal sonority difference of 3 while [fricative+liquid] clusters have an MSD setting of 2, it is expected under the MSD that [fricative+liquid] clusters will be harder to learn; however, this prediction does not obtain in Kim (2000) nor in Kwon (2005), both of which report more instances of vowel insertion in [stop+liquid] onsets and not as many in [fricative+liquid], especially [sibilant+liquid] onsets (see also Barlow 1997, and Pater and Barlow 2003, who report earlier acquisition of /s/-clusters by children). Davidson (2001) undermines the validity of the MSD as an adequate measure of difficulty in cluster production since her native English speakers produced equally poorly Polish onset clusters e.g. [kp-], [vz-] with MSD settings of 1, e.g. [tf-], [zm-], [zr-], all with MSD settings of 0, although the MSD predicts the latter set to be easier. Furthermore, their production of English and Polish clusters that have the same sonority setting (MSD 1) varied substantially.

Sonority explanations can at best be characterized as based upon a universal tendency, not an absolute universal; other studies have hinted at the possibility for an explanation outside the realm of sonority. Following, I review some of these proposals.

2.4. Structural Account

The generative structure-based syllable approach has been motivated by the unique behavior of [sibilant+stop] clusters. For one thing, the [sibilant+stop] onsets [sp-], [st-] and [sk-] violate the SSP by forming a sonority reversal. However, despite the sonority violation, [sibilant+stop] clusters, compared to other clusters, are quite prevalent among the world's languages and have been found fairly easy for L2 speakers of English to acquire (see §4.4.3 and §4.4.4). They usually resist anaptyxis and, when modified by vowel epenthesis, undergo prothesis instead (see §4.4.5). Broselow (1988, 1993), and Selkirk (1982) claim that C_1 and C_2 in [sibilant+stop] clusters share the same root node and are, therefore, analyzed by native as well as non-native speakers of English as a single two-segment constituent onset (Figure 2.1b) rather than a branching onset as in Figure 2.1a (Kenstowicz 1994, but see Archibald 1998, 2003 for a different structural account):[11]

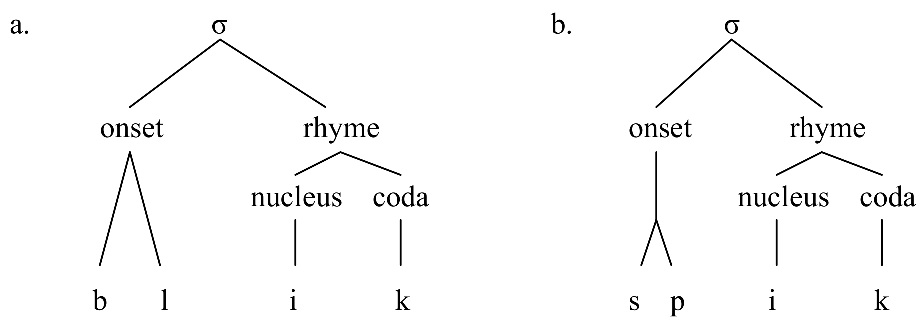


Figure 2.1 Prosodic representation of branching onset (a) vs. complex onset (b)

Unlike in Figure 2.1a, the [sibilant+stop] onset in Figure 2.1b is structurally similar to how affricate onsets would be projected, i.e. parsed as a complex onset with two inseparable segments. Under Broselow's (1993) structural analysis, whenever vowel epenthesis applies to resolve consonant clusters anaptyxis is generally assumed to be the default mechanism, and prothesis in [sibilant+stop] clusters is said to be an exception.

[11] Archibald proposes a phonological L2 parser according to which L2 prosodic structures are acquired. His proposal rests on principles of Government Phonology (Kaye et al. 1990).

Alternatively, Giegerich (1992) and Barlow (2001) analyze [sibilant+stop] clusters as singleton stop onsets (i.e. consisting of one segment) with sibilant /s/ being linked directly to the main syllable node as illustrated in Figure 2.2:

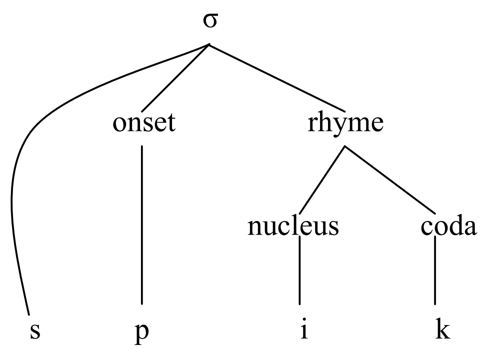


Figure 2.2 Prosodic representation of /spik/ 'speak' with adjunct /s/

Similar to the adjunct analysis in Figure 2.2 is the analysis that treats the sibilant as extrasyllabic (e.g. Goad and Rose 2004, Levin 1985):

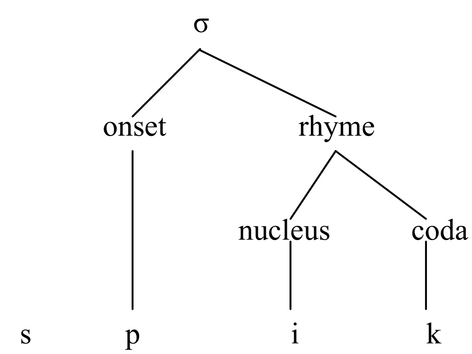


Figure 2.3 Prosodic representation of /spik/ 'speak' with extrasyllabic /s/

One corollary of Broselow's (1993) prosodic analysis is that [sibilant+stop] clusters are structurally different from other clusters, including [sibilant+sonorant] (e.g. [sn-], [sl-]). If the structural representations underlying [sibilant+stop] and [sibilant+sonorant] clusters are distinct, one would expect these to pattern differently with regard to cluster resolution. However, this is not the case; Fleischhacker (2001, 2005) states that in some languages such as Kazakh [sibilant+stop] and [sibilant+sonorant] clusters behave similarly: both equally undergo prothesis (see Gouskova 2001 for further criticism of the structure-based account). Therefore, empirical support for a structural dichotomy between [sibilant+stop] and other types of clusters appears to be limited. Moreover, on a functional phonetic level sibilant /s/ is believed to be highly perceptible due to its internal frication cues (Wright 2004); this renders Broselow's unsplittability argument of [sibilant+stop] clusters, which relies on their exceptionality to the SSP, irrelevant since the main purpose of the SSP is to maximize the perception of individual segments (Fleischhacker 2001: 13).

As mentioned earlier, Giegerich (1992) treats initial [sibilant+stop] sequences [sp-], [sk-], and [st-] as consisting of single segment onsets (i.e. the occlusive) with the sibilant left-adjoined to the syllable head. Consequently, [sp-], [sk-], and [st-] are predicted to function like [p-], [k-], and [t-]. Several predictions follow from Giegerich's analysis. First, [sp-], [sk-], and [st-] should behave like singleton CV onsets where modification owing to syllable structure (e.g. vowel epenthesis) is unlikely to occur. Nevertheless, these clusters, although less commonly than other types of clusters, regularly undergo syllabic restructuring via vowel epenthesis in L2 acquisition (Broselow 1988, Kwon 2006) as well as in loanword phonology (Fleischhacker 2001, 2005). Further, analyzing sibilant /s/ in initial [sibilant+stop] clusters as an appendix segment attached to the main syllable node would suggest, consistent with Halle and Vergnaud (1980) and in fact with Giegerich's (1992: 147-152) analysis of final consonant clusters, that /s/ final [stop+sibilant] sequences (e.g. [-ps], [-ks], [-ts]) be prosodified as rightmost adjuncts dominated immediately by the syllable node as in Figure 2.4 (cf. Figure 2.2):

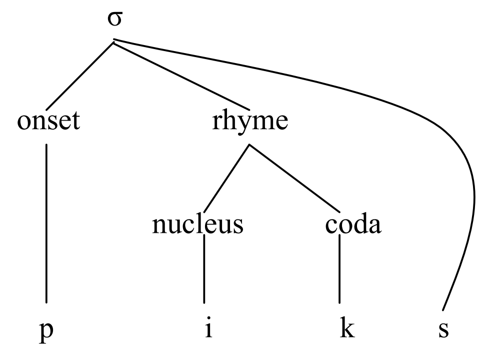


Figure 2.4 Syllabic structure of /piks/ 'peaks' with adjunct /s/

In other words, the onset and coda in these sequences would consist only of the stop, with the sibilant appended to the initial or final word margin. If this is in fact the case as argued by Giegerich (1992: 148), then sequences of initial [sibilant+stop] and final [stop+sibilant] should in principle be equally (un)problematic for L2 speakers, given they are structurally identical. Exactly this prediction has been empirically tested and refuted by Kirk and Demuth (2005), who report better performance on [stop+sibilant] codas than [sibilant+stop] onsets by English-speaking children. In addition, a well-observed difference between onset and coda clusters in general exists and is often reported in L2 studies (e.g. Anderson 1987, Hancin-Bhatt and Bhatt 1997, Kwon 2006, Sato 1984). In Anderson 1987, Egyptian Arabic, Amoy, and Mandarin Chinese speakers of English modified coda clusters more frequently than onset clusters. Similarly, Hancin-Bhatt and Bhatt (1997) asserts that more vowel epenthesis errors are made in onset clusters while Yoo (2004) concludes coda clusters are more prone to consonant deletion.

2.5. Frequency Effects

In general, frequency has been shown to have an effect on the perception and production of speech sounds. Vitevich and Luce (2005), for instance, measure reaction times using a speeded same-different task to find that response times are faster for words occurring with phonemes of high probability in the lexicon; Frisch et al. (2000) demonstrate how lexical frequency affects English listeners' recognition of possible English non-words. Their word recognition task shows English listeners being able to better recall non-words with higher frequency sequences than non-words with lower frequency.

Levelt et al. (2000) show how the acquisition of Dutch syllable structure improves as frequency of different syllable types increases. High frequency CV syllables were early-acquired compared to low frequency CCV syllables, which were acquired later. Kirk and Demuth (2005) find a positive correlation between the frequency with which onset and coda clusters occur and accuracy of production; [stop+sibilant] codas which comprised 22% of the CC clusters database were acquired more easily than [sibilant+stop] onsets, which added up to only 5%. And finally, an interesting study that examines children's sensitivity to frequency versus markedness in the production of English codas is Stites et al. 2004. Its findings show that one of the children first acquired the more frequent, albeit the more marked, stop codas; the other, however, acquired the less frequent and less marked fricative and nasal codas first. Stites et al. suggest that such asymmetry in the acquisition process is learner-based: some learners are frequency-oriented, others are more sensitive to markedness.

Yet others (e.g. Davidson 2006, Pitt 1998) discount any essential role of frequency in phonological acquisition. Pitt (1998), for example, evaluates the role of frequency in the perception of consonant clusters that are illegal for native speakers of American English by measuring listeners' labeling biases on an 8-step /r/ to /l/ continuum. It was predicted that based on frequency alone, and not knowledge of the phonotactics, perceptual bias would exist in liquid labeling regardless of whether the cluster is legal or not: [gr-]-[gl-] (both legal) vs. [tr-]-[tl-] ([tl-] illegal). Moreover, the degree of this bias should reflect frequency differences in various clusters, with effects as follows: [tr-]-[tl-] (142-0) > [dr-]-[dl-] (74-0) > [sr-]-[sl-] (0-49). Results yielded little support for frequency. First, listeners demonstrated aural partiality in illegal clusters only, but not in [gr-]-[gl-]. Second, no differences in magnitude of bias between [tr-]-[tl-] and [sr-]-[sl-] were borne out although the smallest effect among illegal clusters predictably belonged to [dr-]-[dl-].

A similar kind of counter-evidence to frequency in consonant clusters is reported by Davidson (2006). In a repetition task, 20 English native speakers were asked to produce various pseudo-Czech words with initial consonant clusters which are illegitimate in English (e.g. [fs-], [zv-], and [vz-]). The results revealed no positive correlation between frequency of occurrence in the Czech lexicon and accuracy of L2 production: the cluster's higher frequency did not lead to more native-like production (see also Kabak and Idsardi 2007, Moreton 2002).

Thus there seems to be little evidence showing a robust effect of frequency on the acquisition of L2 clusters. In the following section, I turn to a more recent approach involving the implementation of phonetic factors, perceptual and/or articulatory, in the explanation of consonant cluster acquisition.

2.6. Perceptual and Articulatory Factors

Perceptual salience is an active operative force in determining diachronic change (e.g. assimilation, omission, merger, etc.) and ultimately language inventory (Lindblom and Maddieson 1988, Ohala 1990). In the general field of L2 acquisition, the phonetically-based approach is manifested in claims about perceptual and articulatory similarity between the native language and the L2. These claims underpin a number of L2 speech acquisition models (e.g. Flege's 1987, 1995 Speech Learning Model, Best’s 1995 Perceptual Assimilation Model, Kuhl’s 1991 Perceptual Magnet Effect, Major and Kim’s 1996 Similarity Differential Rate Hypothesis).

In the analysis of phonotactic patterns and more specifically consonant clusters, perceptual as well as articulatory considerations have been invoked. Steriade (1999a) argues in favor of a syllable-independent perceptibility-based account in explaining phonotactic generalizations among a number of languages. Steriade (2009) notes the ‘too many solutions’ conundrum that arises when a constraint-based theory such as OT (Prince and Smolensky 1993) yields multiple possible solutions for a single phonotactic violation. As a revision, she proposes the inclusion of the P(erceptibility) Map in the grammar, a set of statements about the degree of perceived distinctiveness among different contrasts in various contexts (see §6.2.2).

Similarly, Coté (2000) argues against the generative structural prosodic approach to vowel insertion and segmental deletion. She regards vowel insertion as a perceptually enhancing mechanism associated with consonant clusters whose members have relatively weak acoustic cues. In constructing the argument for the perceptually-based approach, Coté maintains that a more principled account of phonotactic patterns and processes in many languages (e.g. schwa epenthesis and deletion in French, cluster resolution in Québec French, and vowel epenthesis in Cairene Arabic, Iraqi Arabic, Parisian French, and Basque) can be achieved within a perceptual framework of analysis. The degree of perceptual saliency, defined as the sufficient auditory knowledge by which a segment is recognized, and the mechanism of cue enhancement in syntagmatic contexts underlie Coté’s perceptual account. Perceptual saliency of sequential segments depends on several acoustic factors such as vowel transitions in CV vs. VC contexts, acoustic cues associated with stops (e.g.

auditory burst release), as well as contrast and modulation in the acoustic signal (see §3.3).

Fleischhacker (2005) attributes onset simplification patterns of vowel insertion and consonant deletion to the preferential retention of forms that are perceptually similar to their sources (see Kenstowicz 2003 for a similar proposal). Looking at reduplication and loanword adaptation in a number of languages, Fleischhacker claims [obstruent+sonorant] clusters behave differently from other types of clusters. First, they are more prone to 'skipping' (i.e. deletion or failure to copy the second element of the cluster in reduplication), to use Fleischhacker's term, than are [sibilant+stop] clusters. Second, whenever vowel insertion applies, [obstruent+sonorant] clusters tend to undergo anaptyxis while [sibilant+stop] undergo prothesis. In either case, Fleischhacker argues that these differences follow from a general principle in the grammar which favors perceptual proximity between the input and output. Skipping and anaptyxis lead to minimal departure from the original cluster (i.e. less dissimilarity) in [obstruent+sonorant] but not in [sibilant+stop] clusters, which are rendered more similar to their prothetic counterparts.

Fleischhacker (2001, 2005) provides direct evidence for these similarity effects on consonant deletion and vowel insertion in loanword phonology. In one of her auditory similarity judgment experiments, Fleischhacker presents native English speaking listeners with two epenthetic variants of initial clusters: anaptyctic and prothetic, and instructs them to rate on a 1-7 scale (1 = least similar, 7 = most similar) the similarity between each cluster and its two epenthetic variants. Clusters with anaptyxis were rated more similar in sound to unepenthetic [obstruent+sonorant] clusters, and those with prothesis more similar in sound to unepenthetic [sibilant+stop] clusters (see §4.4.5).

Along the same lines but considering articulation as well is Kwon (2006) who develops a Phonetically Based Consonant Cluster Acquisition Model (PCCAM). Kwon proposes that markedness relations in the production of different non-native onset and coda clusters can be constructed on the basis of the articulatory and perceptual properties holding within each word-initial and final consonant string. Drawing on Articulatory Phonology (Browman and Goldstein 1989), the PCCAM comprises several articulatory parameters such as degree of gestural overlap, articulatory release, stridency and voicing effects. Perceptually, it finds basis in Steriade's P-map (2009) and Coté's (2000) principle of perceptual saliency, in loanword and interlanguage phonology (Fleischhacker 2001, 2005). Kwon (2006) conducts two experiments to assess the validity of the PCCAM in the perception and production of English clusters by native Korean speakers. Both auditory judgments of perceived similarity between clusters and their epenthetic realizations as well as vowel epenthesis errors in production were largely consistent with the predictions of the PCCAM (see §4.4.5).

To sum up, what the discussion so far of cluster simplification in L1, loanword and, L2 phonology shows is that restructuring of consonant clusters can better be understood under a phonetically-based framework. Later in the dissertation (Chapter 4), I show how a perceptual approach can be used to explicate markedness relations among different types of English onset clusters.

2.7. Other Factors

Many other factors besides the ones pointed out so far have been examined in the field of second language acquisition. Although not all pertain specifically to the acquisition of consonant clusters, the effects of biological factors such as the Critical Period Hypothesis (DeKeyser 2000, Lenneberg 1967), age of acquisition (Flege and

Fletcher 1992), gender (Flege et al. 1995, Thompson 1991), individual learner characteristics such as L2 proficiency (Abrahamsson 2003, Riney and Flege 1998), length of residence in the L2 speech community (Flege and Fletcher 1992, Flege and Liu 2001), and degree of motivation in L2 learning (Moyer 1999) have all been well-established. Other task-related variables such as linguistic context (e.g. language-specific morphosyntactic constraints, e.g. Hansen 2004) have been shown to affect the outcome of L2 speech research. In Chapters 4 and 5, I revisit some of these factors when discussing the experimental design of the dissertation.

2.8. Summary

This chapter has considered some of the main approaches to the acquisition of second language consonant clusters. The effects of L1 transfer, language universals such as typological markedness and sonority, structurally-based, and frequency accounts on the perception and production of L2 clusters have been demonstrated. Finally, it was shown that an alternative phonetic approach can better explain the way consonant clusters are realized in interlanguage phonology. The chapter concludes with a brief statement of other possible factors in L2 acquisition.

CHAPTER 3

ISSUES IN SPEECH PERCEPTION

3.1. Introduction

Speech perception is a human trait that allows listeners to segment, decipher and process an incoming stream of information carried by the speech signal. Although some of the elements involved in the process of speech perception are known to us, how exactly speech sounds are perceived is far from clear. Theories of speech perception have yet to answer questions regarding what the smallest unit of analysis is in speech perception (e.g. gesture, phoneme, syllable, etc.), how perceptual normalization is achieved despite variation in the signal, and to what extent the native linguistic experience shapes listeners' ability to perceive native and non-native sound contrasts.

This chapter introduces some of the key concepts and issues recurrent in the study of native and non-native speech perception. While a detailed scientific discussion of the auditory mechanisms and acoustic signal processes that underlie human perception of speech sounds is beyond this dissertation, a careful look at some of the issues in L2 speech perception any theory has to address is necessary. This being said, this chapter concerns itself more with perception and less with audition (see §3.2).

The chapter is organized as follows. In sections 3.2-3.5, a brief overview of some of the properties underpinning speech analysis is given. Section 3.6 looks at how the acquisition of a native language delimits listeners' ability to perceive non-native segments and phonotactic strings, and the section concludes with a discussion of two prominent models of L2 speech perception. Section 3.7 explores the relationship between perception and production of L2 sounds. The chapter concludes in §3.8 with a summary of the main points.

3.2. Audition vs. Perception

Hearing is different from perceiving; we are often not able to *interpret* linguistically the sounds of a new language although we have no trouble *hearing* them. That being said, audition is the essential process by which pressure waves from sound are received, transduced into mechanical vibrations, and eventually transmitted to the brain through electrochemical excitement of the auditory nerves. Perception, on the other hand, is concerned with interpreting and mapping the acoustic information in the signal into linguistically interpretable units.

The difference between speech audition and perception can be better discerned if we observe the effects of deafness and aphasia on language acquisition. Deaf people cannot learn (spoken) language because they are not able to hear speech sounds, not because of brain damage. On the contrary, some aphasic patients, while having intact hearing mechanism, are not able to learn language due to brain trauma interfering with their ability to process speech sounds normally (Borden et al. 2003: 152).

3.3. Acoustic Cues[12]

In the identification of different speech segments, speech perception relies heavily on auditory cues embedded in the acoustic signal. The strength vs. weakness and essentially presence vs. absence of different acoustic cues influence whether and how speech sounds are perceived by listeners. We are not inclined to misperceive a segment with strong acoustic cues, but may do so when the segment is poorly cued.

---

[12] It should not be understood from the discussion here of acoustic cues that visual cues are peripheral to the perception of speech sounds. In fact, a growing body of research has shown clear evidence for bimodality in speech perception (Hardison 2003, McGurk and MacDonald 1976).

This propensity also depends on the larger context in which a segment appears. The process is even further exasperated by noise, as is often the case in the typical speech environment, rendering the role of auditory cues in the perception of speech sounds even more pivotal.

The scope of this section will be limited to a brief survey of the acoustic cues crucial to the perception of vowels and consonants in general, the relevance of which will be clear later in the dissertation (§4.3).

3.3.1. Redundancy of Acoustic Cues

When more than one piece of information in the speech signal provides a cue for the discrimination of one sound from all others, the signal is said to have redundant and robust acoustic cues. The non-linearity and gesturally overlapping nature of speech sounds in the acoustic signal result in redundancy of acoustic cues. A stop, for example, can be identified by the brief amount of silence at the outset due to the occlusion of the vocal tract, the burst release immediately following the occlusion, and by the formant transitions of the adjacent vowels (Borden et al. 2003, Delattre et al. 1955 in Wright 2004).

Speech is made up of an overlapping stream of vowels and consonants, which often results in cues of one segment being spread across those of another, especially preceding or following segments. Cue redundancy is favorable and can enhance our ability to sufficiently recover place, manner and voicing information in the speech signal, especially in poor communication environments (e.g. noise, extremely rapid speech, consonant strings, etc.) as is typical of spoken language.

3.3.2. CV vs. VC

The onset sequence CV is perceptually privileged over the coda VC. First, the prevocalic position entails a consonant occlusion release which 'yields a phonetic burst, a perturbed postconsonantal airstream that clarifies voicing and place of articulation contrasts' (Hudson 1995: 655). Frication noise in this position appears to be more intense too, which provides unambiguous acoustic cues for fricatives (Borden et al. 2003). Second, initial CV formant transitions are more robust than transitions out of a vowel (i.e. final VC transitions) and provide better consonant cues (Wright 2004). On the other hand, preconsonantal plosives tend to be unreleased and the only place cues for stops in this position are found in the offset VC transitions. As for fricatives in this position, they have shorter and much less intense frication noise than in CV. In general, it is much more difficult to recover the acoustic cues of consonants in preconsonantal than prevocalic positions (Coté 2000).

3.3.3. Voicing

A key indication to voicing in consonants is the presence/absence of periodic waves in the signal. Periodicity occurs when the subglottal air pressure forces adducted vocal folds to vibrate systematically resulting in glottal pulses or voicing which can be seen on a spectrogram as a vertical dark line known as the voice bar (O'Shaughnessy 2000).

There are other correlates of voicing in consonants. For stops, the time between the burst release and the onset of phonation or voicing, known as the Voice Onset Time (VOT) (Lisker and Abramson 1964), as well as the duration of the closure provide good cues for voicing. Shorter VOT and closure durations are characteristic of voiced stops. Voiceless aspirated stops have longer VOT and closure durations, while voiceless unaspirated stops are somewhere in between (Borden et al. 2003, Ladefoged 2005). In fricatives, voicing is often accompanied by shorter frication noise.

Listeners can also rely on differences in the duration of a preceding vowel to cue the contrast between voiced and voiceless consonants: short vowel durations cue voicelessness while voicing is cued by longer vowel durations (Borden et al. 2003, O'Shaughnessy 2000).

3.3.4. Acoustic Cues of Speech Sounds

3.3.4.1. Vowels and Diphthongs

Vowels are perceptually very salient since their articulation involves little or no constriction, resulting in clear vocal tract resonances, known as formants, which can be maintained for a relatively long period of time. Vowels may also be long, voiced and high in intensity (i.e. loud) relative to consonants. The steady state spectral frequencies of the first (F1) and second (F2), and sometimes third (F3), formants provide excellent internal acoustic cues to the quality of vowels (Ladefoged 2005, O'Shaughnessy 2000). In connected speech, listeners rely less on steady state formant frequencies and more on formant transitions from adjacent consonants in identifying vowels (Borden et al. 2003, Wright 2004).

For diphthongs, articulation moves from one point to another since they are made up of two sounds. This movement translates as sudden changes in the spectral formant frequencies. These rapid transitions from one formant structure to another are argued to be more reliable cues to diphthongs than are formant frequency values alone (Gay 1970 in Borden et al. 2003).

3.3.4.2. Semivowels: Glides and Liquids

Like vowels, semivowels or approximants are produced with narrowing, but no complete constriction, of the vocal tract. They are usually identified by prominent formant frequency changes, i.e. transitions. /w/ is characterized by a rising F2 and /j/ by a falling one; /r/ is often marked by a sharp onset dip in the third formant and to some degree in the second formant as well. That is to say, F2 and F3 both rise at the outset of /r/, setting it apart from the lateral /l/. In general, formant transitions in approximants change more rapidly than in diphthongs, making them more consonant-like (Borden et al. 2003: 162).

3.3.4.3. Stops

Stop consonants provide a good example, as mentioned earlier, of cue redundancy, since more than one cue can contribute to their identity as stops. For one thing, the brief cessation of the airstream due to the occlusion of the oral cavity followed by the abrupt burst release of the trapped air provides important transitional cues in the identification of stops.

Acoustic cues for stops can also be found in neighboring segments. Vowels and vowel-like consonants such as glides and liquids with consistent clear formant structures can carry information relating to the quality of stops. Formant transitions, the changes in the formant structure that result from the imposition of a constriction on the vocal tract into and out of stops (Wright 2004), contain valuable information about a stop's place of articulation.

3.3.4.4. Nasal Stops

Among the common cues for nasals is the presence of anti-resonances or anti-formants. Anti-resonance is the attenuation of higher formants spectrally visible as white stripes. Another cue for nasal stops is nasal murmur, a low resonance usually below 500 Hz. The combination of weakened formants, or anti-formants, and nasal murmur has been shown to cue nasals, particularly their place of articulation (Borden et al. 2003, Wright 2004). However, according to Malécot (1958) the most reliable cue for the nasal place of articulation is the formant transitions, in particular those of the second formant (F2).

3.3.4.5. Fricatives

Unlike stops, fricatives allow air to pass through the vocal tract with some varying degree of constriction. The result is long turbulence in the airstream known as frication noise strong enough to reliably cue the fricative place of articulation and distinguish it from other consonants (Wright 2004).

Frication is an internal cue and is most helpful in sibilant fricatives and less so in non-sibilants. In sibilants, energy from frication is intense and rather concentrated at higher frequencies, whereas in non-sibilants energy appears to be evenly dispersed and less intense. It is this intensity of frication noise and its diffusion that set apart sibilants from non-sibilants.

3.3.4.6. Affricates

Because a sequence of a stop followed by a fricative is what makes an affricate, its acoustic cues are naturally those of stops and fricatives combined. The airflow stoppage (silence), release (burst) and formant transitions of adjacent vowels cue the stop portion of the affricate acoustic signal, while the continuant portion is identified by its unmistakable frication noise. That is not to say, however, that a stop-fricative sequence is identical acoustically to an affricate. In the case of affricates the transition from a stop to a fricative is much more rapid than that of non-affricate stop-fricative sequences. In addition, in affricates both the stop and fricative tend to be shorter in duration than in non-affricates (O'Shaughnessy 2000: 51).

3.4. Linearity and Segmentation

Speech sounds are physically non-linear. There are no boundaries between one segment and another and the acoustic signal alone cannot be parsed into discrete segments that correspond to phonological units. Instead, the human ear receives simultaneously a mesh of information concerning more than one segment in the acoustic signal due to the coarticulation and overlap of speech sounds; acoustic cues for one segment may span other neighboring segments. This relationship between the acoustic signal and the phonetic message was characterized much earlier by Liberman et al. (1967) as being highly coalescent, complex and non-linear. Nonetheless, we are able to efficiently segregate and recover strings of individual speech sounds from the signal rather effortlessly.

3.5. Variation and Perceptual Constancy

The problem of non-linearity is further exasperated by variation in the acoustic signal. In addition to gestural overlap, acoustic cues in speech sounds can vary substantially from one context to another. They can be better or worse depending on the relative position of the segment(s) in the speech signal (e.g. prevocalic vs. postvocalic consonants (CV vs. VC), two consonant strings vs. three consonant strings (CCV-CCCV), initial vs. final clusters (CCV-VCC), etc.).

A number of other factors can contribute to lack of invariance in the acoustic signal. One factor is the rate at which speech is articulated and transmitted; rapid speech decreases the temporal space (i.e. time) between one segment and another which increases degree of coarticulation. Others include the signal loudness, masking noise, inter-speaker and intra-speaker variation (e.g. in pitch, clarity), identity of the talker (e.g. male, female, child), and even the quality of the talker's voice (e.g. breathy, creaky, harsh, etc.).

Despite this apparent lack of invariance in the acoustic signal, we are efficiently able to constantly and accurately perceive and segment speech sounds into single interpretable units. This leads to one of the main oddities in the study of speech perception: Perceptual Constancy (Johnson and Strange 1982, Kuhl 1980). Although for the past five decades or so attempts have been made to address this issue, it is still far from being understood (see Strange 1995 for an overview).

Perceptual constancy or the normalization of variation in the signal is possible because speech perception proceeds categorically; that is, sounds are perceived as belonging to underlying categories and whatever marginal differences exist among these categories are filtered out by our perceptual apparatus. Listeners store perhaps approximate rather than absolute auditory phonetic specifications for different phonemes and use such identifying information in matching phonemes to their relevant acoustic cues encoded in the incoming speech signal (Kuhl 1991, Stevens and Blumstein 1978). Others suggest listeners achieve categorical perception by matching already stored *abstract* motoric movements of the vocal tract (Motor Theory of Liberman and Mattingly 1985) or *actual* articulatory gestures of speech sounds (Direct Realist Approach of Fowler 1989) against information in the signal.

Under a realist approach, listeners store *real* not blueprints of detailed speech sounds gestures, making the units of analysis in perception and production the same. While this explanation circumvents the perceptual constancy problem earlier discussed, it raises serious concerns about speech processing since it presumes that listeners internalize detailed and highly variable gestural information about speech (e.g. contextualized and/or talker-specific). This approach also assumes a very strong link between speech perception and production, although such relationship is not entirely clear (see §3.7).

Having looked at some of the issues in the study of speech perception in general, I now turn to the study of second language speech perception.

3.6. Second Language Speech Perception

L2 speech perception suffers from the same problems that plague the native perception of speech sounds. It is further complicated by the fact that L2 learners are L1 perceivers to begin with, bringing their native-language experience and biases into the perception of L2 sounds and structures. Below I show how this is the case by examining the impact of linguistic experience on the perception of non-native phonemes and consonant strings.

3.6.1. Linguistic Experience and Categorical Perception

The effect of the ambient language sound system on the categorization of new speech sounds has been well-established. Breakthrough research in infant speech perception reemphasizes the role of the native language phonology in shaping the perceptual system of infants and adults alike. During their first year of life infants have been claimed to show the uncanny ability to perceive and categorize both native and non-native speech sounds (Kuhl 2004, Kuhl and Iverson 1995). However, their ability to differentiate non-native speech sounds diminishes radically by the end of the first year (Werker and Tees 1984), while sensitivity to native contrasts becomes increasingly heightened (Kuhl et al. 2006). The loss of sensitivity to foreign contrasts has been attributed to interference from the ambient (native) language and thus, to use Kuhl's (1993a) terminology, the once 'citizen of the world' becomes a 'culture-bound' perceiver.

For adults, continued exposure to the native language, while rendering their auditory system more attuned to and confined to the native language sounds, costs them the ability to detect minor phonetic differences which they encounter as a result of learning a second language sound system. Strange (1995: 19) summarizes the L2 perception experience as follows:

Between early infancy and adulthood, then, children's interactions with their linguistic environment while acquiring their first languages produce significant changes in the perception of speech sounds. There is a loss in the ability to differentiate phonetic categories perceptually that are not phonologically distinctive in the native language, while native contrasts may become more highly differentiated.

The literature is replete with studies showing the effect of linguistic experience on learners' ability to perceive L2 sounds categorically. One classic example is the categorical perception of voicing in stops.[13] Early work by Abramson and Lisker (1970) shows that listeners categorize continua of the acoustic correlate of voicing in stops (i.e. VOT cues) according to the phonological contrasts of their native language: English and Spanish listeners divided stimuli on the voicing continuum into voiced or voiceless, while Thai listeners identified three categories: voiced, voiceless aspirated and voiceless unaspirated. In other words, each listener group relied upon their L1 phonological voice distinctions in categorizing the VOT continuum.

Further, the values along the VOT continuum are interpreted in a language-specific manner. In the Abramson and Lisker (1970) study, English listeners associated relatively short VOT with voiced /b, d, g/ and long VOT with /p, t, k/; however, Spanish listeners displayed the opposite correlation (i.e. voicing was signaled by long VOT while voicelessness by short VOT). Interestingly, this L1 perceptual bias in Spanish listeners carries over to production as demonstrated by Flege and Eefting (1987a), who report on Spanish speakers producing English long

---

[13] Another well-known example is the categorical place distinction of American English /l/-/r/ by Japanese listeners (Best and Strange 1992).

VOT /p, t, k/ in a Spanish-like manner (i.e. with shorter VOT) (see also Flege 1987, in which English speakers produced French /t/ with longer English-VOT values).

3.6.2. Linguistic Experience and Perception of Non-native Phonotactics

The robustness of the native language effect on speech perception is by no means limited to the phenomenon of categorical perception. Research by Jusczyk and colleagues (e.g. Jusczyk et al. 1994) shows infants 9 months of age and younger being influenced by and becoming sensitive to their native language sound patterns (more recently see Kajikawa et al. 2006, Mugitani et al. 2007). By the time adulthood is reached, listeners will have adapted well to the phonotactic constraints of their native language and are likely to demonstrate their L1 perceptual biases when hearing non-native sequences.

In what follows, I discuss the effect of illegal phonotactics and syllable structure on L2 perceptual processing.

3.6.2.1. Perception of Illegal Phonotactics

That adult listeners when perceiving illegal L2 consonant sequences display language-specific biases toward their native language phonotactics is well-documented early on (Brown and Hildum 1956, Polivanov 1931). An early study by Greenberg and Jenkins (1964) argues that listeners' acceptability rating of non-native sequences proceeds as a function of the phonological proximity of these to existing English words. Their study showed that the more changes are introduced in a consonant cluster, the less likely it would be rated as native-like. Massaro and Cohen (1983), replicated by Pitt (1998), observe how their English listeners tend to perceive the illegal onset clusters /sr-/ and /ʃl-/ as legal /sl-/ and /ʃr-/, accordingly. Hallé et al. (1998) report that possible but non-attested French word-initial /tl-/ and /dl-/ are heard as legal /kl-/ and

/gl-/ respectively by French listeners (see Pitt 1998 for different but not necessarily opposing findings).

More recently, Hallé and Best (2007) examine the perception of Hebrew initial /tl-/-/dl-/ and /kl-/-/gl-/ clusters by French and Hebrew listeners. Since French permits only /kl-/ and /gl-/ while all four clusters are allowed in Hebrew, French but not Hebrew participants predictably had difficulty discriminating /tl/-/kl/ and /dl/-/gl/, although the latter pair was relatively less difficult (77% discrimination success compared to 64%). Hallé and Best also observe that, consistent with the discrimination pattern in /tl-/-/kl-/ and /dl-/-/gl-/, French listeners tended to substitute /t/ for /k/ in /tl-/ clusters, and to a lesser extent /d/ for /g/ in /dl-/ clusters during a forced-choice categorization task (see also Moreton 2002 on native English listeners' biases in the perception of non-native sequences).

Berent et al. (2007) investigate the perceptual sensitivity of English speakers to differences in the sonority of onset clusters that are illicit in English. Onset clusters involving small sonority rises (e.g. [bn-]) are believed to be more marked than large sonority rises (e.g. [bl-]); sonority plateaus in the onset (e.g. [bd-]) are more marked than rises; and sonority falls in the onset (e.g. [lb-]) are more marked than plateaus (597). Consequently, it was hypothesized that highly marked onset clusters with falling sonority would incur more perceptual errors than less marked clusters with sonority plateaus, which, in turn, would incur more perceptual errors than onsets with rising sonority.

Using three different tasks – auditory judgment, discrimination, and lexical decision – Berent et al. explore the monotonic relationship between English listeners' ability to perceive vowel epenthesis (anaptyxis) in monosyllabic words containing onset clusters and the sonority-based claims of markedness. Their findings reveal that the examined onset clusters were perceived variably: English listeners were likely to perceive some clusters as disyllabic (i.e. with vowel epenthesis) more often than others. Further, this pattern was directly related to cluster markedness. English listeners were more likely to perceive vowel epenthesis in sonority falling clusters (e.g. lbif → ləbif) than in clusters with sonority plateaus (e.g. bdif → bədif), and in sonority plateaus than in sonority rises (e.g. bnif → bənif). In general, the erroneous percept of vowel epenthesis increased as a function of the cluster sonority profile: more marked clusters triggered the percept of vowel epenthesis more often relative to less marked clusters. Berent et al. conclude that English speakers have knowledge of the relative markedness relations among these onset clusters demonstrated by their perception of vowel epenthesis even though the stimuli contained none.

3.6.2.2. Perception of Illegal Syllable Structure

Studies focusing on the native language experience and its role in the perception of illegal prosodic structures are not that many. One famous cross-linguistic study of illegal syllable-structure perception is Dupoux et al. (1999), which compares French and Japanese listeners on the perception of medial consonant clusters (VCCV). Listeners presented with a six-step continuum ranging from no vowel VCCV (e.g. *ebzo*) to full vowel VC<u>V</u>CV (e.g. *eb<u>u</u>zo*) were asked in an identification task whether they heard a vowel in the middle of the nonce words. Japanese listeners expectedly reported hearing a vowel more often than French listeners, surprisingly even so in stimuli where no vowel was present. In fact, Japanese listeners reported hearing a vowel 70% of the time in the non-vowel condition (i.e. 0 ms):

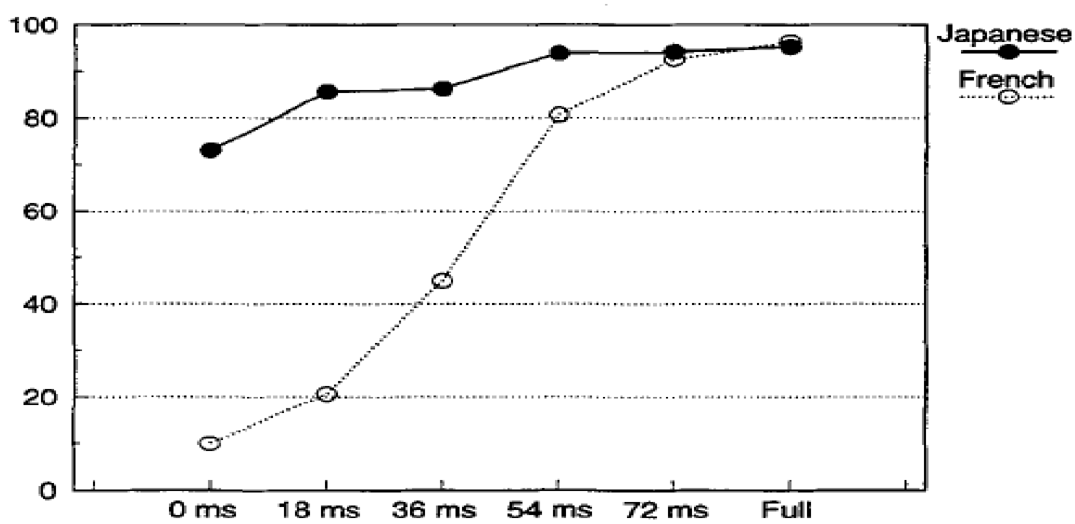


Figure 3.1 Perceptual judgments (identification) of /u/ vowel by Japanese and French listeners as a function of the epenthetic vowel duration (Dupoux et al. 1999: 1570)

Dupoux et al. conduct a follow-up speeded ABX discrimination paradigm with the finding that, consistent with the results of the identification task, Japanese listeners showed considerable difficulty in discriminating VCCV from VCVCV stimuli (e.g. *ebzo-ebuzo*). They attribute Japanese listeners' perception of an epenthetic vowel and their difficulty with discriminating *ebzo* from *ebuzo* to the phonological properties of Japanese: Japanese, unlike French, bans medial consonant clusters such as [-bz-]. More specifically, the perception of an epenthetic vowel is believed to be an artifact of the Japanese perceptual system (see Dehaene-Lambertz et al. 2000 and Dupoux et al. 2001 for similar findings).

Dupoux et al. (1999) is reminiscent of a much earlier study done by Polivanov (1931) that documented Japanese listeners' perception of initial consonant clusters. Polivanov suggests that listeners perceive (and produce) non-native utterances through their L1 phonological system. Based on anecdotal evidence, he notes that Japanese learners of Russian consistently perceive Russian words such as /tak/ 'so' as [ta.ku] and /drama/ 'drama' as [do.rama] (or in some cases [zu.rama]). Crucially, Polivanov argues that the native perceptual grammar introduces the percept of the epenthetic vowels [u] and [o] to avoid the illicit coda [-k] and consonant cluster [dr-]. Polivanov concludes that Japanese listeners perceive syllabic structures within the bounds of their L1 prosodic system.

This phenomenon of vowel perception is not limited to Japanese listeners. The penchant to perceive vowel epenthesis when none exists has also been confirmed in Korean speakers of English who realize English *spike* and *flute* perceptually as [sɨ.p^{h}ai.k^{h}ɨ] and [p^{h}ɨl.lu.tɨ]. Kabak and Idsardi (2007) replicate and extend Dupoux et al. (1999) by examining the perception of English medial consonant clusters by Korean listeners. A word-medial string violation in Korean can be caused by either an illicit type of coda (e.g. [c.] of *[c.m]) or an illicit string of a legal coda and onset (e.g. *[k.m], which surfaces as [ŋ.m]). That is, either syllabic or consonantal contact restrictions can contribute to Korean medial cluster ill-formedness. Confounded in Dupoux et al., Kabak and Idsardi attempt to tease apart these two factors in Korean. Interestingly, their AX discrimination task reveals Korean listeners being more successful in distinguishing pairs where the violation is due to bad contact only as in [p^{h}ákma] vs. [p^{h}ák^{h}ʊma] as opposed to bad coda (and contact) as in [p^{h}ácma] vs. [p^{h}ác^{h}ɪma]. Based on this, they draw the conclusion that epenthesis, at least in Korean, is induced by L1 constraints on syllable structure rather than mere sequential relationships between segments.

3.6.3. Theories of Non-native Speech Perception

Two models of speech perception, namely Best's Perceptual Assimilation Model (1995) and Kuhl's Perceptual Magnet Effect (1991), make claims clearly unique to the perception of non-native speech sounds. These models are concerned, however, with the perception of segments and have little to say about the perception of non-native syllables.

3.6.3.1. Perceptual Assimilation Model (PAM)

PAM seeks to explain the gradient success listeners demonstrate in perceiving and discriminating non-native segments and contrasts. According to PAM, the ability to discriminate various non-native contrasts follows from implicit or explicit assimilation of each contrasting segment to a native category. A regulating factor in determining assimilability is the degree of closeness or discrepancy which native and non-native sounds share, as stated in Best (1995: 139):

> The fundamental premise of the perceptual assimilation model of cross-language speech perception is that non-native segments, nonetheless, tend to be perceived according to their similarities to, and discrepancies from, the native segmental constellations that are in closest proximity to them in native phonological space.

The degree of similarity between native and non-native phonemes is defined by ‘the spatial proximity of constriction locations and active articulators and by similarities in constriction degree and gestural phasing’ (194). A non-native phoneme is more likely to be assimilated to a native one when it is perceived as a good exemplar of its native equivalent. A non-native phoneme can be assimilated as:

(a) an existing native speech sound perceived as being identical, acceptable or deviant exemplar of the native category.

(b) a speech sound within the phonological space, but not representative of any particular native category.

(c) a non-speech sound and therefore outside of the native phonological space.

Given that contrast discriminability in PAM is predictable from the assimilation of each segment in the contrast, the different combinations of (a-c) result in the following pairwise assimilation types, each with its predicted level of discriminatory accuracy (Best 1995: 195):

1. *Two-Category Assimilation (TC Type)*

   Each non-native segment is assimilated to a different native category, and discrimination is expected to be excellent. An example of this type is the Tigrinya ejective contrast between the voiceless alveolar /t'/ and bilabial /p'/, assimilated in English as the alveolar-bilabial contrast in /t/-/p/, respectively (Best 1993).

2. *Category-Goodness Difference (CG Type)*

   Both non-native sounds are assimilated to the same native category, but they differ in discrepancy from the native 'ideal' (e.g. one is acceptable, the other deviant). Discrimination is expected to be moderate to very good, depending on the magnitude of difference in category goodness for each of the non-native sounds. The voiceless ejective and non-ejective velars /k'/-/k/ in Zulu are more likely to be treated as voiceless velar /k/ in English, with Zulu /k/ as the good exemplar and ejective /k'/ as the deviant (Best 1994).

3. *Single-Category Assimilation (SC Type)*

   Both non-native sounds are assimilated to the same native category, but are equally discrepant from the native 'ideal'; that is, both are equally acceptable or both equally deviant. Discrimination is expected to be poor (although it may be somewhat above chance level). The Thompson Salish contrast in ejective velar /k'/ and uvular /q'/ is a SC assimilation type where both sounds are likely to be perceived as deviant exemplars of prototypical English velar /k/ (Best 1994).

4. *Both Uncategorizable (UU Type)*

   Both non-native sounds fall within phonetic space, but outside of any particular native category, and can vary in their discriminability as uncategorizable speech sounds. Discrimination is expected to range from poor to very good, depending upon their proximity to each other and to native categories within native phonological space. The well-known difficulty Japanese speakers have in distinguishing the English /l/-/r/ contrast (Best and Strange 1992) can be an example of this type; neither liquid is assimilated to a good Japanese equivalent.

5. *Uncategorized vs. Categorized (UC Type)*

   One non-native sound is assimilated to a native category, and the other falls in phonetic space outside native categories. Discrimination is expected to be very good. The English /r/-/w/ distinction for Japanese listeners fits this type where, unlike English /w/ which is assimilated as Japanese /w/, English /r/ is not assimilable to any Japanese category (Guion et al. 2000).

6. *Nonassimilable (NA Type)*

   Both non-native categories fall outside of speech domain being heard as non-speech sounds, and the pair can vary in their discriminability as non-speech sounds; discrimination is expected to be good to very good. English speakers' discrimination of the Zulu clicks, which for an English listener do not resemble any speech sound, falls into this category of assimilation (Best et al. 1988).

Several hypotheses of the PAM have been tested. Best et al. (1988) examine the discrimination of Zulu (a Bantu language) place and voicing click contrasts by native speakers of American English. As predicted, English listeners showed NA assimilation when discriminating the Zulu click contrasts. Best et al. argue that since clicks are gesturally very distant from, and therefore cannot assimilate to, any English phoneme as their articulation involves ingressive suction followed by loud release, they were most likely perceived by English listeners as non-speech sounds that do not belong to any native category, in which case listeners were more reliant on auditory and phonetic properties of clicks, which make them highly discriminable.

More recently, Best et al. (2001) evaluate the predictions of PAM for the TC, CG and SC assimilation types by examining English listeners' perception of Zulu and Tigrinya consonant contrasts. For native English listeners the contrasts between the voiceless and voiced Zulu lateral fricatives /ɬ/-/lʒ/ as well as the ejective bilabial and alveolar stops in Tigrinya /p'/-/t'/ were expected to fit a TC assimilation type,[14] in which case Zulu fricatives would be equated with English voiceless vs. voiced apical fricatives and Tigrinya /p'/-/t'/ with non-ejective bilabial vs. alveolar stops. Moreover, English listeners were predicted to show a category goodness (CG) type in their assimilation of Zulu voiceless aspirated and ejective velar stops /k^h/-/k'/ to American English /k/. Finally, a discrimination pattern consistent with single category (SC) was believed to emerge in the Zulu contrast between plosive and implosive voiced bilabial stops /b/-/ɓ/, given that both sounds would most likely be perceived as English /b/. In general, English listeners' performance on the non-native Zulu and

[14] In determining what sound contrasts belong to which assimilation types, gestural similarities and differences among English, Zulu, and Tigrinya sounds were discussed by Best et al. (2001: 778-9).

Tigrinya contrasts followed straightforwardly from these expectations, along with their predicted discrimination levels, confirming the TC > CG > SC discriminability ranking suggested by PAM (also see Best and Strange 1992 and Guion et al. 2000).

3.6.3.2. Perceptual Magnet Effect (PME)

Kuhl (1991) demonstrates that human adults and infants alike perceive speech sounds as being close to or distant from prototypical categories that exist for each sound. She reveals how stimuli resembling prototypic /i/ were perceived as excellent instances of the /i/ vowel, while prototypic-dissimilar stimuli had poor goodness ratings. Kuhl (93) further shows that the use of the prototype vs. non-prototype of a category as a referent affects how listeners perceive speech sounds:

> When the prototype of the category served as the referent, the other members of the category were perceived as being more similar to it. The prototype perceptually assimilated near neighbors in the category, effectively reducing the perceptual distance between it and the other members of the category …

She continues stating that:

> The prototype of the category functioned like a perceptual magnet for other category members; it assimilated neighboring stimuli, effectively pulling them toward the prototype (104).

This magnet effect, Kuhl speculates, can explain why a non-native phoneme that is similar (but not identical) to a native one is often assimilated to the native language phoneme. That is, L2 listeners in evaluating foreign L2 sounds make reference to established L1 prototypic categories. This is due to what Kuhl in later work calls the Native Language Magnet (NLM) effect (Kuhl 1991, 1993a) according to which:

> … exposure to language early in life produces a change in perceived distances in the acoustic space underlying phonetic distinctions, and this subsequently

alters both the perception of spoken language and its production (Kuhl and Iverson 1995: 122).

In Kuhl and Iverson (1995), the ramifications of NLM/PME as they apply to the process of L2 speech perception are investigated on evidence from vowel perception by adults and infants, with motherese vowels being rated as good category exemplars.

To sum up, although, as mentioned above, Best's (1995) PAM and Kuhl's (1991) PME stand out as being models of perceptual learning, they make no predictions for the perception of non-native syllables. Further, theories of L2 acquisition in general make direct claims about production with the assumption that production and perception are isomorphically related. The following section bears on this relationship.

3.7. Relationship between L2 Perception and Production

The relationship between perception and production in L2 speech sounds is not entirely clear. Some of the long-standing lingering questions central to L2 speech research include: is production preceded by perception, or vice versa? Are perception and production of speech sounds closely correlated? And is there ample evidence for a link between speech perception and production.

Precedence of speech perception over production seems to be partially motivated by studies that attribute production errors by L2 speakers to inaccurate perception of L2 sounds (Borden et al. 1983, Grassegger 1991, Rochet 1995, Yeon 2003). These studies suggest that difficulty in the production of L2 speech sounds is precipitated by faulty perception, which presupposes precedence of the perceptual process over the production one (see §3.7.1.1). In addition, the priority of speech perception over production is endorsed by early research on child language acquisition (e.g. Strange and Broen 1981).

However, a study by Goto (1971) on the acquisition of American English /r/ and /l/ reports how Japanese learners of English failed to perceptually distinguish the liquids in their speech and in the speech of native American speakers although they were able to produce them distinctively. This somehow striking discovery was later reconfirmed in a replication study by Sheldon and Strange (1982) in which Japanese speakers' production of AE /l/ and /r/, as judged by American raters, was much better than their perception. Sheldon and Strange believe that perception in the Japanese case was far less developed than production, and take this as evidence for the independence of perception and production. Their findings are consistent not only with Goto but with even earlier work by Briére (1966) demonstrating a similar perception-production relationship in acquiring non-native stop and fricative contrasts by American learners (see also Flege and Eefting 1987a).

3.7.1. Evidence for an L2 Perception-Production Link

The discussion below is confined to studies that established (1) a correlation between production errors and poor perception and vice versa, and (2) perceptual training benefits on production.

3.7.1.1. Correlation between Perception and Production

That perception and production of L2 speech sounds can be correlative has been established. Barry (1989: 160) carries out two experiments on the perception and production of English vowels by L2 German learners and concludes that 'well-established perceptual categories are more likely to be accompanied by more acceptable production'. Similarly, Grassegger (1991) examines the perception and production of Italian plosive consonants by Austrian German learners and remarks how perceptual accuracy of phonetic categories leads to improved production.

In evaluating the role of perception in foreign accent, Rochet (1995: 385) speculates that 'accented pronunciations of L2 sounds by untrained speakers may be perceptually motivated'. A significant relationship between production errors and self-perception is also found by Borden et al. (1983) for Korean L2 learners acquiring American /l/ and /r/, and more recently by Yeon (2003) who reveals a startling similarity in the perception-production error pattern of English palatal codas by Korean speakers.

3.7.1.2. Perceptual Training

In the literature, studies showing speech perception training effects on production abound. Yamada et al. (1996) trained their 23 Japanese speakers on the perception of AE liquids for 15 days after which improvement in both perception and production was evident. Further, two follow-up production tests (after 3 months and 6 months), conducted to unmask any long term training benefit, showed improved production still. In other words, the retention tests indicated that the trained group still performed better in their production of /l/ and /r/ than they did prior to receiving any perceptual training. Similarly, Bradlow et al. (1997) report progress in the perception and production of English /l/ and /r/ after perceptually training their 11 Japanese speakers on the liquid distinction. Logan et al. (1991) suggest that benefits of perceptual training may generalize to novel stimuli and are not limited to stimuli used in the training. Their study shows how laboratory training of Japanese speakers proved useful in perceptually enhancing their ability to perceive the English liquid distinction in stimuli never before heard (see Borden et al. 1983 and Hardison 2003 for similar findings).

To conclude, the correlation between perception and production as well as the advancement in production owing to perceptual training reported on in these studies provide clear evidence of a link between the perceptual and productive modules of speech processing. In this dissertation, I maintain that a relationship between the perception and production of L2 onset clusters exists (see §5.5), but make no specific precedence or causality claims about the nature of this relationship, as evidence is conflicting in this regard.

3.8. Summary

This chapter has looked at some of the issues common to the analysis of speech perception. The role of auditory cues in differentiating speech segments perceptually as well as some speech perception characteristics like variation and perceptual constancy were considered. Next, it was argued that a concomitant factor in the perception of non-native sounds, phonological contrasts, and syllable structure is exposure to the native language phonology. Effects of the listener's native language on the perceptual system included L1-based biases and tendencies clearly observed in the perception of non-native segments and syllables. Two models of (L2) segmental perception were briefly discussed: PAM (Best 1995) and PME (Kuhl 1991). The chapter ends with scrutiny of the relationship between speech perception and production.

CHAPTER 4

MARKEDNESS IN THE PERCEPTION OF L2 ENGLISH ONSET CLUSTERS

4.1. Introduction

In Chapter 3, we saw evidence of how influence from the native language phonology is exerted on the listener. The L1-habituated perceptual system can seem to recognize elements in the speech code that are not there (Chang et al. 2007, Dupoux et al. 1999, Kabak and Idsardi 2007, Moreton 2002, Polivanov 1931); force us to ignore elements of the signal that are there (e.g. Dupoux et al. 1997); or alter the segmental identity of the non-native string we are to decode (Hallé and Best 2007, Hallé et al. 1998, Massaro and Cohen 1983, Pitt 1998). For L2 listeners the impact of the native language is both real and influential; their ability to perceive non-native consonant clusters can vary in difficulty depending on the phonotactics of their native language.

This chapter addresses the following question: is perceptual accuracy of L2 consonant clusters regulated by other cluster-dependent factors, besides the L1? Research on L2 consonant cluster production has concluded that markedness, the relative difficulty in producing clusters natively, is influenced by the type of cluster that is being acquired; the varying degree of the cluster's sonority, frequency, typological distribution, as well as the phonetic salience of its elements have all been argued to determine how well L2 clusters are produced (see Chapter 2). For L2 cluster perception, however, this question remains vaguely addressed.

This chapter explores empirically the relative difficulty with which English onset clusters are perceived by L2 learners whose L1 sanctions no initial clusters. A priori assumption is that listeners' ability to methodically alternate between native-like and non-native-like perceptions of the cluster is typical of L2 perception. That being said, this chapter investigates from a perceptual stance how different English onset clusters vary in perceptibility and whether this variation can systematically be a function of the type of cluster in question.

The chapter is organized as follows. Section 4.2 outlines an exploratory experiment which reveals how auditory difficulty in the perception of vowel epenthesis in English onset clusters is relative. Based on this, §4.2.4.2 proposes a hierarchy of markedness or difficulty in the perception of onset clusters, and §4.3 and §4.4 present phonetic as well as empirical justification in support of this hierarchy. The chapter closes in §4.5 with a reprisal of the main points.

4.2. Experiment 1: Perception of English Onset Clusters by Saudi Listeners

This section details an aural perception experiment, the purpose of which is to test how Saudi Arabian listeners generally perceive American English onset consonant clusters with respect to vowel epenthesis. The experiment is strictly exploratory in that it addresses questions such as whether and how overall error rates in perception are related to different types of consonant clusters, whether a cluster-induced bias toward hearing anaptyxis vs. prothesis exists, and whether English proficiency has an overall effect on error rates in perceiving English clusters. Before delving into the hypotheses and experimental design, a quick overview of relevant aspects of Saudi Arabic phonology ensues.

4.2.1. Language Background

Saudi or Arabian Arabic (SA) is the variety of Arabic spoken in the Kingdom of Saudi Arabia. It subsumes a number of geographically closely related dialects, mainly Hijazi in the west (Bakalla 1979, Omar 1975), Najdi in the north and center (Ingham 1994), Gulf in the east (Ingham 1982, Johnstone 1967b) and Janubi in the south (Al-Shahrani 1988). These dialects, however, are confined to home use and everyday conversations while standard Arabic arises as the official language of media, science and instruction.

Of particular interest is Urban Hijazi Arabic (HA), the native dialect of the L2 population in this experiment (see §4.2.3.1 for more on participants). HA is similar to the Arabic dialects of Khartoum in Sudan and Upper Egypt (Ingham 1971), and although there are at least two main sub-varieties of HA, namely Urban and Bedouin, the term Hijazi usually refers to Urban Hijazi. HA shares essentially the consonant inventory of standard Arabic. The following table shows the different consonants of HA along with their places and manners of articulation:

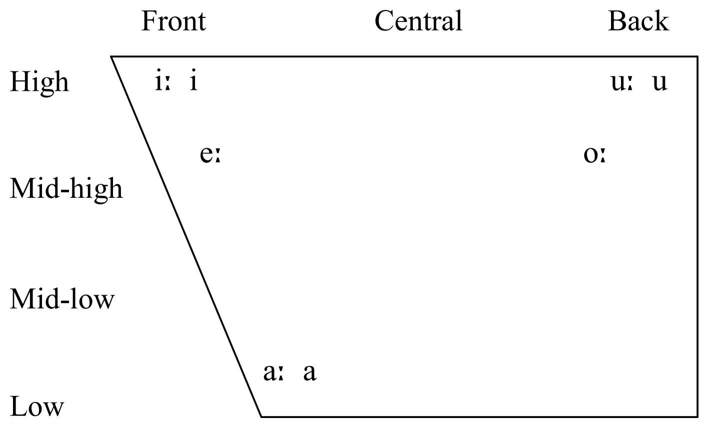
	Stop	Affricate	Fricative	Trill	Continuant	Nasal
Bilabial	b				w	m
Labiodental			f			
Dental	t d					
	t̲ d̲					
Alveolar			s z	r	l	n
			s̲ z̲			
Palatoalveolar		dʒ	ʃ			
Palatal					j	
Velar	k g[15]		χ ɣ			
Pharyngeal			ħ ʕ			
Glottal	ʔ		h			

*Underlining represents velarized (emphatic) consonants.
*Phonemes to the left in pairs are voiceless.

Table 4.1 HA consonant phonemes

A main characteristic of HA is the substitution of dental stops /t/ and /d/ for standard Arabic /θ/ and /ð/. Another point of divergence from standard Arabic is the retention of the uvular stop /q/ in very few words and the tendency to often replace it with the voiced velar /g/ (see Ingham 1971 for further discussion).

In addition to the three canonical vowels of standard Arabic and their long counterparts, HA employs two more vowels, namely /eː/ and /oː/ as illustrated in the following quadrilateral vowel chart:



*Vowels to the right in pairs are short.

Figure 4.1 Vowel phonemes of HA

With regard to syllable structure, HA, like standard Arabic, prohibits consonant clusters word-initially, but not medially or finally. However, in HA speech the short central vowel /ə/ is optionally inserted in medial -CC- clusters. According to Ingham (1971: 279), speakers of Hijazi are often not cognizant of this extra syllable, but it does surface in their speech under certain environments:

1. Clusters containing /ʕ/ or /ɣ/ as the first or second element, e.g.

   [ˈgit̲əʕa] < /ˈgit̲ʕa/ ‘piece’

   [ˈs̲abəɣi] < /ˈs̲abɣi/ ‘my dye’

2. When /r/, /dʒ/, /ħ/ are followed by a voiced consonant in the cluster, with the exception of /dʒ/ followd by /l/ and /z/, e.g.

   [ˈχurəmu] < /ˈχurmu/ ‘its hole’

   [ˈliħəgu] < /ˈliħgu/ ‘they followed’

   [ˈʔadʒəri] < /ˈʔadʒri/ ‘I run’

Finally, the default glottal stop /ʔ/ used in standard Arabic as an obligatory onset for all VC syllables is often elided in HA initial syllables, c.f. Hijazi /ma aruːħ/ with standard /ma ʔaruːħ/ 'I do not go'.

4.2.2. Hypotheses

The following hypotheses are tested in Experiment 1:

1. The likelihood of SA listeners reporting vowel epenthesis will vary as a function of cluster type.
2. In reporting vowel epenthesis, the choice between anaptyxis or prothesis will depend on the type of cluster.
3. Beginning SA listeners will incur more overall errors than advanced listeners in their perception of English consonant clusters.

Our focus on vowel epenthesis versus consonant deletion or phonemic substitution as evidence for perception of onset clusters is justified on the basis of epenthesis being the preferred method of resolution in the production of L2 onset clusters (Anderson 1987, Carlisle 1998, Hancin-Bhatt and Bhatt 1997, Silverman 1992, Yoo 2004), assuming perception and production share a single grammar (Smolensky 1996). Previous research has also investigated the perception of onset clusters from the perspective of vowel epenthesis (e.g. Berent et al 2007). Hypothesis 1 is inspired by similar work done on the production of consonant clusters in which L2 learners' ability to produce consonant clusters in a native manner varies by cluster type (Broselow and Finer 1991, Chan 2006, Chen 2003, Eckman and Iverson 1993, Kim 2000, Kwon 2006, Major 1996). This claim of cluster type as a factor, although found to be true of cluster production, remains largely untested in cluster perception. Instead, similar work on cluster perception has investigated other factors such as the phonotactics of the listener's native language and the role they play in the perception of illegal clusters, but has not examined cluster type as a variable (e.g. Dehaene-Lambertz et al. 2000, Dupoux et al. 1999, Dupoux et al. 2001, Kabak and Idsardi 2007, see §3.6.2.2). An exception, however, is Berent et al. (2007), who report on the likelihood of vowel epenthesis being heard by English listeners as a function of the different sonority profiles that characterize illegal onset clusters (see §3.6.2.1). Thus, Hypothesis 1 attempts to explore the claim that as a function of cluster type some clusters more than others are subject to a perception of epenthesis. An a priori assumption here is that there is variable difficulty in the perception of vowel epenthesis in different English consonant clusters.

Hypothesis 2 is formed on the basis of research suggesting an asymmetry in the placement of vowel epenthesis in consonant clusters (Fleischhacker 2005, Gouskova 2001, Singh 1985). For example, Fleischhacker (2005) presents perceptual evidence for a cluster-dependant asymmetry of vowel epenthesis site in loanwords. She argues that preference in the assignment of anaptyxis vs. prothesis serves to maximize auditory similarity between the cluster and its epenthetic form: while anaptyxis is perceptually favored over prothesis in [obstruent+sonorant] clusters, prothesis over anaptyxis is favored in [sibilant+stop] clusters (see §5.3.2.5).

Finally, Hypothesis 3 follows from research that establishes linguistic proficiency as a factor in L2 learning (e.g. Abrahamsson 2003, Riney and Flege 1998, Weinberger 1987). Bradlow et al. (1997) and Flege (1995), among others, demonstrate how higher proficiency in the target language can lead to improved L2 pronunciation. If in fact a relationship between linguistic proficiency and L2 production holds, it is not unreasonable to assume, therefore, a similar ameliorative role of proficiency in L2 perception. Broadly speaking, Hypothesis 3 is consistent with counter-CPH claims that set age limitations on the mastery of non-native skills (e.g. Birdsong 1992, Bongaerts 1999, Flege et al. 1995).

4.2.3. Methodology

4.2.3.1. Participants

Forty-eight male SA learners of English ranging in age from 20-35 took part in this study. In addition to being among the least commonly tested learner groups, SA learners of English present a good sample since Arabic, as mentioned earlier, is a strictly simple-onset language. Participants were recruited from different universities in Michigan and were stratified into two groups: beginning learners enrolled in ESL English courses with a length of residence (LOR) equal to or less than a year, and advanced learners who at the time of the study were junior and senior college students with LOR of 2-6 years.

Of particular concern was the native dialect of the participants; efforts were made to recruit people born and raised in the Hijaz region of Saudi Arabia (HA remains faithful to standard Arabic in its prohibition of onset clusters). In some cases, however, people who have resided in the Hijaz region since childhood although born somewhere else in the country were also included. Subjects who received any specialized training in phonetics or pronunciation, attended a foreign school where English was the primary language of instruction, and those exposed to English at an early age were excluded. In order to minimize the number of exclusions, subjects were screened prior to their recruitment during a short informal interview to ensure the above requirements are met (see Appendix A for the oral questionnaire). Of the 48 who participated in the study, the results of 5 beginning and 3 advanced subjects were eliminated due to their failure to complete all task requirements. According to self-report, no participant had any hearing or speaking difficulties.

4.2.3.2. Materials

The materials comprised words with 26 English onset clusters which fall under the following 7 categories:

[stop+liquid]: br-, bl-, tr-, dr-, kr-, kl-, gr-, gl-

[stop+glide]: bj-, kj-, mj-, tw-, dw-, kw-

[fricative+glide]: hj-, fj-, sw-

[non-sibilant+liquid]: fr-, fl-, θr-

[sibilant+stop]: st-, sk-

[sibilant+nasal]: sn-, sm-

[sibilant+liquid]: sl-, ʃr-

Two monosyllabic nonsense English words for each consonant cluster were constructed. One of the vowels /i/, /a/ or /u/ served as syllable nuclei while the coronals /t/ or /s/ served as simple word-final codas (see Appendix B for the wordlist).[15] The cardinal vowels /i/, /a/ and /u/ are selected because they represent the three canonical vowels in classical, standard and dialectal Arabic (Holes 2004: 59, Watson 2002: 21), and are among the most widely used vowels cross-linguistically (Ladefoged 2001: 25). Coronal codas are chosen because they occur in HA and because of their relatively unmarked status in the world's languages as stated in Rice 2007: 82 (see also Paradis and Prunet 1991). For each of the 52 words (26 clusters X 2 words), two tokens were added, one with anaptyxis (C<u>V</u>CVC), and one with prothesis (<u>V</u>CCVC). Thus, each cluster had three conditions: the (original) unepenthesized condition (e.g. [klit]), the anaptyctic condition (e.g. [kəlit]), and the prothetic condition (e.g. [əklit]).

---

[15] In some cases, the use of real English words could not be avoided as in /trus/ and /gris/.

A male native speaker of American English trained in linguistics recorded the materials twice (total of 156 test words X 2; see Appendix C for wordlist recorded by talker). When producing the epenthetic forms, the talker was instructed to articulate a central mid schwa vowel with stress placed on the second syllable (e.g. [kə.ˈlit] and [ək.ˈlit]). The recording took place in a quiet library lab room using Audacity Software (v.1.2.6) and a clip-on PRO 7 Electret condenser microphone. The talker had to read on a computer screen the target words which were inscribed in the International Phonetic Alphabet (IPA) to ensure a more accurate pronunciation.

Concern arose as to whether the duration of the epenthetic vowel varied drastically in the talker's production of C<u>V</u>CVC and <u>V</u>CCVC stimuli. Inadvertently, discrepancies in the epenthesized vowel duration could significantly bias listeners' responses on the aural task; the respondent may report hearing anaptyxis or prothesis based on vowel duration, not cluster type; see Dupoux et al. (1999) for a discussion of durational effects on the perceptibility of vowel epenthesis, and Fleischhacker (2001) on how unmonitored vowel epenthesis duration can lead to skewed results. In an attempt to hold the duration of the epenthetic vowel relatively constant, duration measurements of the inserted vowel in all anaptyctic and prothetic tokens were calculated using Praat (version 5.0.32, Boersma and Weenink 2005), and submitted to a paired sample *t*-test. A highly significant effect of vowel location was found: $t(51) = -0.458$, $p < .001$, with prothetic tokens being on average longer in duration than anaptyctic ones (58.38 ms vs. 50.08 ms).

Therefore, vowel durations were synthetically modified and a 50 ms median was set with +/- 3 ms discrepancy tolerance. In other words, the largest possible difference allowable between any given stimuli was 6 ms, and anything below or above the 47-53 ms range was synthetically lengthened or shortened to 50 ms. The established criterion resulted in the adjustment of 32 tokens out of 104. The 50 ms threshold was selected on the basis of being close to pre-synthesis anaptyxis and prothesis mean durations (i.e. 50.08 ms and 58.38 ms, respectively). Moreover, similar vowel durations are reported in the literature. Davidson (2007), for instance, measures 48 ms for vowel epenthesis in American speakers' production of unfamiliar clusters, while a vowel duration mean of 60 ms is reported for Moroccan Arabic speakers by Ali et al. (2008).

Manipulation of epenthetic vowel duration proceeded as follows. First, the vocalic portion to be synthesized was selected and a manipulation object with an empty duration tier was produced in Praat. Second, a new duration tier with new lengthened or shortened duration points was created using Pitch-Synchronous Overlap-and-Add (PSOLA) in Praat.[16] Next, the PSOLA-generated duration tier was applied to the original manipulated sound object and the final output was reproduced separately using the Publish-Synthesis function. A post-modification paired *t*-test revealed duration means of 50.05 ms for anaptyxis and 50.64 ms for prothesis with no significant effect of vowel location: $t(51) = -2.03$, *ns*.

4.2.3.3. Task and Procedure

A forced-choice identification paradigm was used to elicit SA listeners' judgments of vowel epenthesis in English onset clusters. The identification task was designed and run using the Experiment MFC function in Praat. Presented aurally with 156 randomly ordered stimuli, subjects were informed they would hear possible English words and instructed to report whether they heard a vowel or not by listening intently to the word-initial margin. Subjects were instructed as follows:

---

[16] PSOLA is a technique in acoustic speech analysis for manipulating the duration and/or pitch of a speech sample. PSOLA allows for manipulating duration without affecting the original characteristics of pitch, and vice versa.

You are about to hear a number of possible English words. Your task is to indicate whether the word has:

(A) a vowel at the beginning in which case you should click the box labeled *Beginning*.

(B) a vowel following the first consonant in which case you should click the box *Between*.

(C) none of the above in which case you should click the box *None*.

Each of the three boxes were displayed on a computer screen, with the next test item starting a soon as a choice has been made. This way the inter-trial interval (ITI) that separates each trial from another is controlled by the test taker. The experiment was administered over two sessions (78 trials each), with an optional 5-minute intermission. Prior to the experiment, a 5-item practice test ensured subjects understood the instructions. The experiment was conducted in a quiet room and lasted an average of 20 minutes for each subject, which seemed reasonably to have limited the problem of task-fatigue.

4.2.3.4. Results

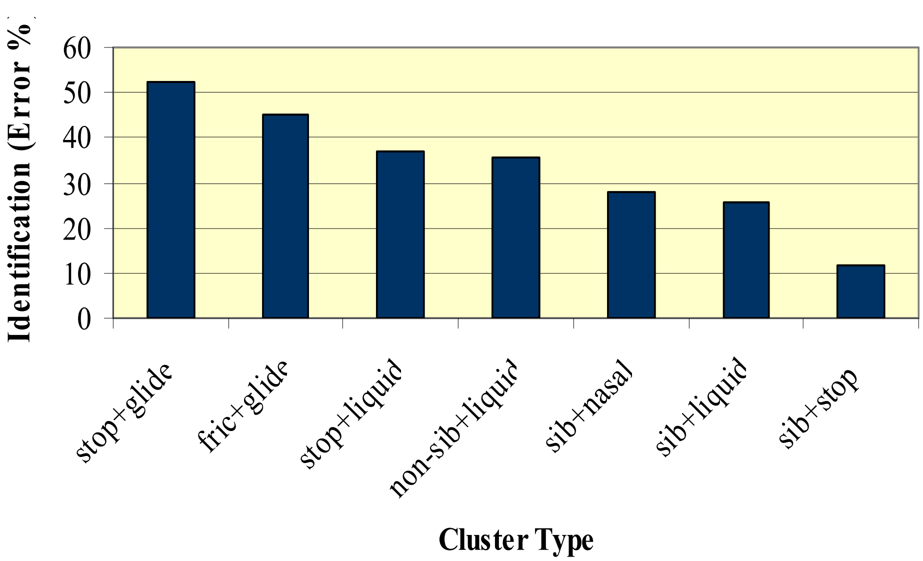
Individual subject identification responses for each of the 156 words were coded A (prothesis), B (anaptyxis), and C (no epenthesis), computed, and tallied in Excel. Identification error rates were calculated, pooled and averaged across the 7 cluster types. Responding 'A' or 'B' to unepenthesized stimuli or 'C' to either anaptyctic or prothetic stimuli was considered a vowel identification error. A prothesis error was counted as such when the participant responded 'A' (prothesis) to a word such as [klit], or 'C' (no epenthesis) to [əklit]. An anaptyxis error was counted as such when the participant answered 'B' (anaptyxis) to [klit], or 'C' (no epenthesis) to [kəlit]. Prothesis responses to anaptyctic stimuli (i.e. 'A' responses to 'B') as well as anaptyxis responses to prothetic stimuli (i.e. ‘B’ responses to ‘A’) were not counted toward vowel identification errors, as the number of such responses was extremely small (see Appendix D for the number of responses on the identification task).

Proficiency

To determine whether proficiency had an effect on listeners’ ability to indentify vowel epenthesis in onset clusters, overall error rates from each subject group were submitted to an independent sample *t*-test. The differences between lower and higher proficiency subjects in identification scores were insignificant: $t(38) = 1.092$, *ns*. Therefore, results for the two subject groups are reported aggregately.

Cluster

To find out if listeners’ responses on the vowel identification task were significantly biased by cluster type, pairwise tests of proportions compared incorrect judgments across the 7 cluster types, i.e. responding ‘yes’ to vowel epenthesis (anaptyxis and prothesis) when there is none, and ‘no’ when there is. The results are shown graphically in Figure 4.2 and summed up in Table 4.2:



For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.

Figure 4.2 Epenthesis: 7-category error rates

	No.	%	fric+glid	stop+liq	non-sib+liq	sib+nas	sib+liq	sib+stop
stop+glide	753	52.29	*ns*	*	*	*	*	*
fric+glide	324	45		*	*	*	*	*
stop+liquid	708	36.87			*ns*	*	*	*
non-sib+liq	258	35.83				*	*	*
sib+nasal	126	26.25					*ns*	*
sib+liquid	123	25.62						*
sib+stop	57	11.87						

* = sig at the level of .05 ($p < .05$)
ns = non-significant ($p > .05$)

Table 4.2 Epenthesis: 7-category error rates and significance (pairwise comparison)

The leftmost column ranks cluster types in a decreasing order of identification errors. [Stop+glide] clusters have the highest rate of vowel epenthesis errors. Just a little over 50% of the time listeners failed to hear the vowel in [stop+glide] vowelled stimuli, as well as reported hearing one in voweless stimuli. Other clusters show improved identification of vowel epenthesis as error rates decline.

Pairwise tests of proportions revealed significant differences among all types of clusters, $p < .05$, with the exception of [stop+glide] vs. [fricative+glide], [stop+liquid] vs. [non-sibilant+liquid], and [sibilant+nasal] vs. [sibilant+liquid] for which differences did not reach the level of significance. Thus, these cluster types were collapsed into broader categories: [stop+glide] and [fricative+glide] types are merged as [obstruent+glide], [stop+liquid] and [non-sibilant+liquid] types are merged as [obstruent+liquid], and [sibilant+nasal] and [sibilant+liquid] types are merged as [sibilant+sonorant]. The a priori division between the cluster types in each pair, although phonologically justifiable, may have been phonetically indistinctive for subjects. Already in the classification of these cluster types we see a possibility for a less refined rearrangement that is both sustainable in phonological terms and necessitated by the results: stops and fricatives in [stop+glide], [fricative+glide], [stop+liquid], [non-sibilant+liquid] clusters belong to the natural class OBSTRUENT, whereas nasals and liquids in [sibilant+nasal] and [sibilant+liquid] clusters belong to the natural class SONORANT. Further statistical analyses on the now four cluster categories revealed all differences were significant. This is graphed in Figure 4.3:

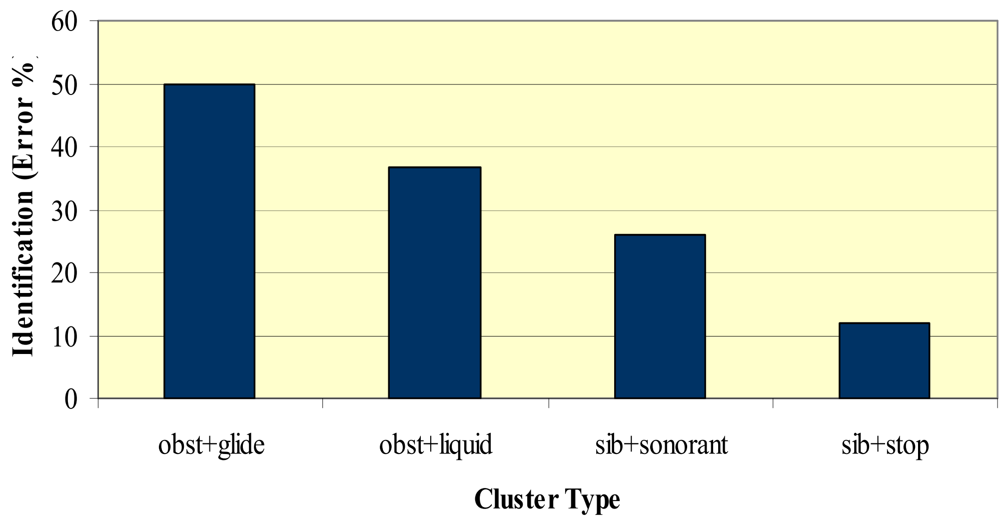


Figure 4.3 Epenthesis: 4-category error rates

This new division among the now four cluster types is significant and ranks [obstruent+glide] as being the cluster in which listeners were least successful and [sibilant+stop] as the one in which listeners were most successful in perceiving vowel epenthesis, as summarized in Table 4.3:

	No.	%	obstruent+liquid	sibilant+sonorant	sibilant+stop
obstruent+glide	1077	49.86	*	*	*
obstruent+liquid	966	36.6		*	*
sibilant+sonorant	249	25.94			*
sibilant+stop	57	11.87			

* = $p < .05$

Table 4.3 Epenthesis: 4-category error rates and significance (pairwise comparison)

Next is the question of whether cluster type as a variable affected listeners' judgments of either anaptyxis or prothesis on the identification task, as indicated by the difference between the two error rates. An anaptyxis error is counted when the listener wrongly records anaptyxis or fails to do so when anaptyxis occurs. By the same token, a prothesis error is counted when the listener wrongly records prothesis or fails to do so when prothesis occurs. Figure 4.4 is a graph of the difference between anaptyxis and prothesis error rates for the four cluster types:

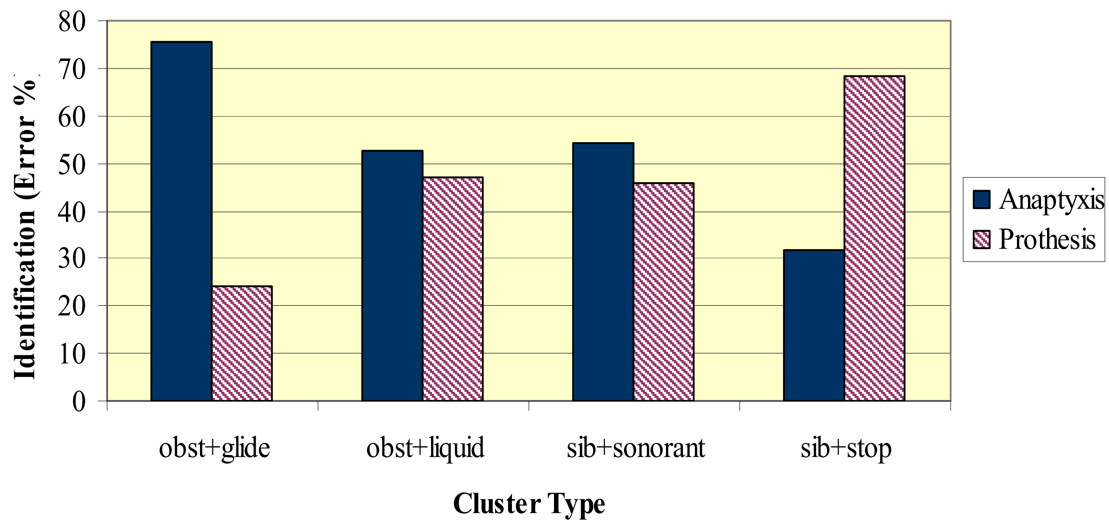


Figure 4.4 Anaptyxis and prothesis error rates for the 4 cluster types

Anaptyxis and prothesis error responses on each cluster type were submitted to a series of individual one variable Chi-square analyses. The following table presents the anaptyxis and prothesis error rates for the 4 cluster types and the Chi-square significance of each comparison:

Cluster Type	Anaptyxis		Prothesis		*sig.*
	No.	%	No.	%	
obstruent+glide	816	75.77	261	24.23	* ($\chi2 = 26.56$)
obstruent+liquid	510	52.8	456	47.2	*ns*
sibilant+sonorant	135	54.22	114	45.78	*ns*
sibilant+stop	18	31.58	39	68.42	* ($\chi2 = 13.56$)

* = $p < .05$

Table 4.4 Anaptyxis and prothesis error rates for the 4 cluster types (Chi-square values in parentheses)

The numbers given in Table 4.4 are erroneous responses incurred by subjects in the identification task. The difference between anaptyxis and prothesis error rates varies in the four cluster types, but is significant for [obstruent+glide] and [sibilant+stop] clusters. No such significance is found in [obstruent+liquid] or [sibilant+sonorant] clusters.

4.2.4. Discussion

This section briefly recapitulates the findings in light of the hypotheses, and then follows this by a more detailed discussion of the results.

Hypothesis 1 predicted that the likelihood of SA listeners reporting vowel epenthesis would vary as a function of cluster type. As clear from the results of the identification task in Table 4.2, subjects' ability to correctly perceive vowel epenthesis in vocalic stimuli (i.e. C<u>V</u>CVC and <u>V</u>CCVC) and reject its presence in non-vocalic stimuli (i.e. CCVC) was best in [sibilant+stop], but poorest in [stop+glide] clusters as indicated by error rates (11.87% vs. 52.29%). This pattern of differential success in which subjects fare better on vowel epenthesis perception depending on the type of cluster obtains significantly except in the pairings [stop+glide] vs. [fricative+glide], [stop+liquid] vs. [non-sibilant+liquid], and [sibilant+nasal] vs. [sibilant+liquid]. In each of these pairings, subjects' identification of vowel epenthesis did not vary significantly. This has led to the re-grouping seen in Table 4.3 whereby [stop+glide] and [fricative+glide] were merged as [obstruent+glide], [stop+liquid] and [non-sibilant+liquid] as [obstruent+liquid], and [sibilant+nasal] and [sibilant+liquid] as [sibilant+sonorant].

Hypothesis 2 stated that in subjects' reports of vowel epenthesis, the choice between anaptyxis or prothesis would vary according to the type of cluster. Table 4.4 shows how this is upheld in two of the four cluster types examined: [obstruent+glide] and [sibilant+stop] clusters demonstrate a clear contrast in the location of the epenthetic vowel. In [obstruent+glide] onsets, there was a preference for anaptyctic vowels vs. a preference for prothetic vowels in [sibilant+stop] onsets.

Finally, Hypothesis 3 is not supported; the ability to correctly identify vowel epenthesis in onset clusters did not differ significantly from lower to higher proficiency subjects. Beginning listeners incorrectly identified vowel epenthesis 39.42% of the time, compared to the slightly lower error rate of 35.86% of advanced listeners. This finding is rather unexpected given the large differences in proficiency and length of residence between the two groups. Recall that beginning subjects were ESL English learners who have stayed in the US for less than a year while advanced subjects were last-two-year college students whose length of residence extended up to 6 yrs. Therefore, inconsistent with the findings of Abrahamsson (2003), Bradlow et al. (1997), Flege (1995), Riney and Flege (1998), the current experiment shows little improvement owing to higher proficiency. One thing to note, however, is that these studies report a proficiency role in production, but not in perception per se. Additionally, findings by other studies (e.g. Altenberg 2005, Dupoux et al. 1999, Duyck et al. 2004, Zsiga 2003) suggest that proficiency may not always be a decisive factor in the L2 learning process.

I discuss below the phenomenon of incorrect perceptions of vowel epenthesis and its possible sources.

4.2.4.1. Source of Vowel Epenthesis: L1 Transfer or CV?

That subjects indicated hearing an epenthetic vowel even though in the stimuli none was physically present is not unprecedented. Japanese, Korean, and Fijian speakers of English report vowel epenthesis when perceiving non-native consonant clusters as they transfer their L1 phonotactic restrictions and syllable well-formedness constraints over to the acquired language (see §3.6.2.2). Thus, it is likely that the percept of vowel epenthesis which subjects in this study reported when encountering unepenthesized clusters originated from constraints of the native language phonology. In particular, the ban on initial consonant clusters found in HA, and more generally in standard Arabic, may prejudice listeners' predisposition to perceive vowel epenthesis when in fact none resides in the input.

Thus, following Dehaene-Lambertz et al. (2000), Dupoux et al. (1999), Dupoux et al. (2001), Kabak and Idsardi (2007), Polivanov (1931), and Schütz (1978), I would like to propose that SA listeners' perception of vowel epenthesis is L1-triggered. In other words, when the listener in this study provided an anaptyxis or prothesis response to a word like /klit/, for example, he relied on the fact that only initial CV- onsets (e.g. HA [ka.ˈraː.si] 'chairs') and VC- (e.g. HA [(ʔ)ɪk.ˈraːm] 'hospitality') are sanctioned in HA. Figure 4.5 illustrates how the percept of vowel epenthesis originates from L1-enforced syllable restrictions:

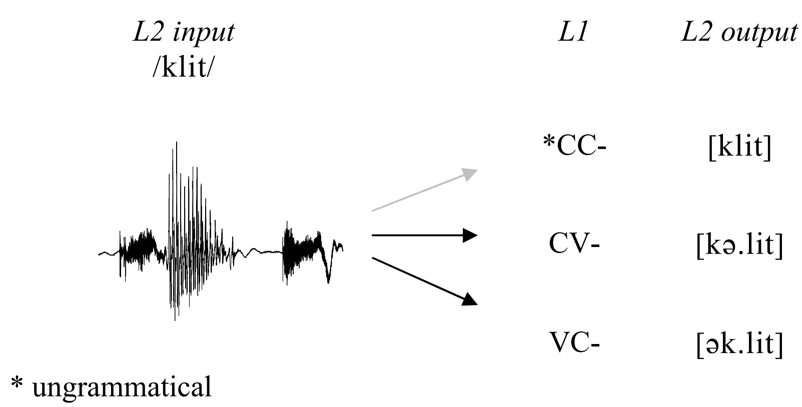


Figure 4.5 L1-driven vowel epenthesis

In Figure 4.5, the acoustic speech signal represented here in waveform is the native American English production of the word /klit/. No evidence of vowel epenthesis exists in the signal per se. In other words, the illusion of vowel epenthesis does not occur in the native acoustic signal; instead it takes place in the L2 perceptual system. The SA listener receives this acoustically raw information and perceptually maps it unto one of the native language possible onset forms, here anaptyxis [kə.lit] or prothesis [ək.lit]. The mapping in Figure 4.5 reflects the influence exerted by the L1 phonotactic system on the listener's ability to perceive non-native input. The question arises then as to what precipitates such bias?

One answer has to do with lexical frequency; it is possible that large lexical frequency differences that naturally exist between legal CV or VC and illegal *CCV- sequences in Arabic biased SA listeners toward hearing (L1-legal) anaptyxis or prothesis in onset clusters (Pitt 1998: 942). This view of the native language intermediating L2 perception is not new. In fact, the L1 has been described in the literature early on as the perceptual filter through which L2 sounds and structures are processed (Polivanov 1931). Thus, Arabic phonotactics provide two possible epenthetic solutions to the onset violation,[17] the choice between either is arguably determined by the perceptual similarity each holds to the input (see §4.4.5).

Alternatively, epenthesis can be attributed to the universal preference for open syllables (CV) as argued for in Hodne (1985), Sato (1984), and Tarone (1987), among others. The sequence CV is not only the most common syllable type among the world's languages (Cairns and Feinstein 1982, Clements 1990, Clements and Keyser 1983, Greenberg 1978), but is perceptually more optimal than CC- or VC- (Borden et al. 2003, Ohala 1990, Wright 2004, see §3.3.2). Therefore, it is reasonable to assume

[17] Another onset cluster resolution in L2 acquisition can theoretically be the deletion of the first or second elements in a cluster although vowel epenthesis is the most common strategy of repair in onset clusters (Carlisle 1998, Hancin-Bhatt and Bhatt 1997, Yoo 2004).

that based on this universal tendency SA listeners were coaxed into believing an anaptyctic vowel existed when encountering CC- onsets.

There is no easy way of telling unfortunately whether L1 transfer, the universal CV-preference or both has led to the percept of vowel epenthesis in this experiment. However, the fact that prothesis responses too were prevalent in some of the results is contradictory to a CV-only explanation. A strong preference for CV- would have given rise to an all anaptyxis and no prothesis pattern. However, such pattern is not borne out by the results as there was no large number of anaptyxis as opposed to prothesis responses; in fact, the distribution in [sibilant+stop] clusters was significantly in favor of prothesis as shown in Table 4.4.

4.2.4.2. Markedness Hierarchy

Based on the findings of Experiment 1, this section introduces a hierarchy of markedness or perceptual difficulty[18] in onset clusters for which evidence is cultivated on phonetic as well as empirical grounds.

A common interpretation for the role of markedness in second language acquisition is the relative difficulty with which learners produce and perceive non-native utterances. Eckman (1987: 61) takes markedness to correspond to difficulty:

> The relative degree of difficulty of the areas of the target language which are more marked than the native language will correspond to the relative degree of markedness.

In the acquisition of L2 consonant clusters, Eckman and Iverson (1993: 241) employ this measure of markedness by assuming that more marked clusters would be more difficult to acquire than less marked clusters. This difficulty in the acquisition of L2 is

[18] Throughout the dissertation, I use the terms ‘perceptual markedness’ and ‘perceptual difficulty’ interchangeably.

often assessed by error rates (Davidson 2003, Eckman and Iverson 1993, Kwon 2006) or by the rate of acquisition (Eckman 1987, Major and Kim 1996).

Following Eckman (1987) and Eckman and Iverson (1993), I take markedness in the perception of onset clusters to correspond to the error rates of SA learners in their identification of vowel epenthesis: a more marked cluster is one in which listeners incurred more identification errors, and a less marked cluster is one in which listeners incurred fewer identification errors. In terms of perceptibility of epenthesis, a more marked cluster is one in which vowel epenthesis is less salient, and a less marked cluster is one in which vowel epenthesis is more salient. The rational for this is phonetic: vowel insertion is more noticeable in a vowel-resistant context, but less so in a context receptive to it (see §4.3). This explanation is consistent with research that associates markedness with perceptual and/or articulatory factors (e.g. Davidson 2003, Kwon 2006).

Given this interpretation of markedness and based on the findings of Experiment 1, the following hierarchy of markedness in the perception of English onset clusters is derived:

(1) Perceptual hierarchy of markedness among onset clusters

less marked ←——————→ *more marked*

[sibilant+stop] >> [sibilant+sonorant] >> [obstruent+liquid] >> [obstruent+glide]

The scale in (1) is established on the basis of identification error rates listeners incurred in Experiment 1, and corresponds therefore to the relative degree of difficulty posed by each cluster type. What contributes most in weight to the relative markedness of [obstruent+glide] clusters is anaptyxis (75.77%); and what contributes most to the relative unmarkedness of [sibilant+stop] clusters is prothesis (68.42%). The markedness of [obstruent+liquid] and [sibilant+sonorant] clusters seems to be sustained almost equally by anaptyxis and prothesis errors, although there was a tendency for anaptyxis errors to be dominant (see Figure 4.4).

Two key questions follow logically: why are anaptyxis errors greatest in [obstruent+glide] clusters, and why, on the other hand, are prothesis errors greatest in [sibilant+stop] clusters? Following I attempt to answer these two questions.

4.3. Phonetic Evidence

The argument here is based on the fact of the number of anaptyxis errors being greater in [obstruent+glide] clusters and fewer in [sibilant+stop] clusters. Because in words having epenthesis listeners had to evaluate two possible outputs, anaptyctic or prothetic, on the basis of the input, in principle preference or dispreference for one or the other has given rise to the observed difference. Recall that duration of the epenthetic vowel was controlled for in Experiment 1 (see §4.2.3.2) and therefore is ruled out as a possible motive for the asymmetry. That being said, this section aims to explain this asymmetrical behavior by examining intrinsic acoustic properties of the consonants in the cluster as well as talker-induced gestural effects on the speech signal.

One thing that may have led to fewer anaptyxis errors in [sibilant+stop] clusters compared to [obstruent+glide] clusters is the voicing status of the segments in the cluster. The [sibilant+stop] clusters examined in this study consist of a voiceless sibilant /s/ followed by one of two voiceless stops: /t/ or /k/. In comparison, the first member in [obstruent+glide] clusters is either voiced or voiceless whereas the second is always one of the voiced glides /j/ or /w/. Given that vowels are by nature voiced, the introduction of a vowel should go less noticed in a context fully voiced compared to a context partly voiceless. The reason for this is that voicing in consonants and vowels is cued by phonation, among other things, which involves the excitation of the vocal folds (glottal pulsing) that results in wave periodicity (see §3.3.3 and §3.3.4.1). In the context of a vowel interposed in a voiced cluster, phonation is initiated for the first consonant (or for the vowel if C_1 is voiceless) and is maintained throughout the epenthetic vowel and the second consonant; that is the whole C_1VC_2 sequence is phonated. By contrast, for epenthesis in a voiceless cluster, phonation is only active during the vocalic portion and ends as soon as the second consonant commences. It is this short period of voicing onset and its abrupt cessation amid two voiceless elements that makes a vowel stand out acoustically. In other words, the alternation between voiceless and voiced in [sibilant+vowel+stop] sequences where the vowel is flanked by two voiceless consonants signals the presence of the vowel. No such alternation is found in [voiced obstruent+vowel+glide] or even in [voiceless obstruent+vowel+glide] sequences.

In addition to voicing, a glide in [obstruent+glide] clusters provides an ideal environment for masking vowel epenthesis. The glides /j/ or /w/, aside from being voiced, are very vowel-like in having internal formant frequencies or transitions, hence the term semi-vowels. The transitions of the second formant (F2), and sometimes the first and third formants (F1, F3), that reflect changes in the vocal tract resonances brought about by movements of the articulators, are essential in the identification of the glide (Borden et al. 2003: 162, Ladefoged 2005: 54). Vowels too are characterized by internal formant structure and can be distinguished from other consonants (and vowels) by their F1, F2 and in some cases F3 values. O'Shaughnessy (2000: 62) argues that like vowels the glides /j/ and /w/ have intense periodic waveforms and 'can be viewed as brief high vowels of greater constriction than corresponding vowels (/j/:/i/, /w/:/u/).' In contrast to glides, voiceless stops in [sibilant+stop] clusters are non-sonorant, do not possess internal formant structure of their own (although they can alter the formant transitions of adjacent vowels), and lack the vocalic quality of periodicity (see §3.3.4).

Similarly, Fleischhacker (2005: 72) proposes that the sequence [obstruent+sonorant] (of which [obstruent+glide] is a subset) involves a stronger cluster-internal 'perceptual break' than does the sequence of [sibilant+stop]. She defines a perceptual break as 'a perceptual event coinciding with the onset of vowel-like formant structure … enhanced both by relatively high-intensity formant structure, and by the presence of a stop closure preceding the onset of formant structure.' According to this definition, the most intense perceptual break or hiatus exists in [stop+sonorant] clusters, the greatest of which is that of [stop+glide] clusters.

[Fricative+glide] as well as [sibilant+glide] clusters also contain a perceptual break, albeit weaker. For a sibilant followed by a stop, however, there is no such perceptual break. Fleischhacker suggests that a vowel inserted where no perceptual break exists, as in [sibilant+stop] clusters, can be far more noticeable – hence dispreferred – than a vowel in [obstruent+sonorant] clusters, which already contain a perceptual break. Fleischhacker (2001: 14-16) goes on to explain how the transition from obstruents to sonorants (e.g. glides) is acoustically similar to the juncture between obstruents and vowels: 'both are characterized by offset of aperiodic noise, onset of formant structure, and a relatively rapid rise in intensity.' By contrast, the transition from sibilant /s/ to a voiceless stop is characterized by 'a rapid decrease in intensity, moving from noise to silence, [and] no formant structure', leading to a juncture distinct from that of /s/ followed by a vowel.

Another explanation that may have led to anaptyxis being perceptually less prominent in [stop+glide] clusters than in [sibilant+stop] clusters has to do with the type of gestures involved in the proper articulation of these clusters and their coordination. As previously discussed, the perception of an epenthetic vowel under an articulatory approach (e.g. Browman and Goldstein's 1989 Articulatory Phonology) can derive from a temporal lag in the alignment of adjacent speech gestures (Davidson 2003, Gafos 2002, Smorodinsky 2002). A temporal lag between C_1 and C_2 can be defined as 'the interval between the end of the constriction plateau of C_1 and the moment in time at which the constriction for C_2 is reached' (Kühnert et al. 2006: 4). This misalignment in intergestural distances can be caused by a myriad of factors such as increased talker's rate of transmission, syllable type (complex vs. simple), prosodic position in the word (margin vs. medial), or gesture type (Byrd 1996, Zsiga 2003). Browman and Goldstein (1989) advocate the notion that cross-linguistically the preferred type of intergestural coordination is one which involves as little overlap as possible (Gafos 2002, Smorodinsky 2002). In clusters consonants are usually shorter in duration and more coarticulated than in prevocalic or postvocalic positions; within clusters the degree of overlap between consonants can depend on the gesture type: a closure gesture temporally overlaps more with a following consonant than does a fricative gesture, as empirically evidenced in Byrd 1996: 233. The motive for this, Byrd suggests, is perceptual. The sudden spike in amplitude and short release burst following the occlusive period bestow crucial cues for the identification of stops. Fricatives, on the other hand, rely on the duration of random noise and the distribution of frequency throughout frication in order to be discernable (see §3.3.4). Because stops can be cued only by the short burst (Malécot 1958, Winitz et al. 1972), they can afford to overlap more with a following consonant than fricatives, which run the risk of imperceptibility if a large portion of the frication noise is masked in extensive overlapping. The result is less articulatory overlapping and more lagging coordination in [sibilant+stop] clusters, relative to the tighter gestural coordination in [stop+glide]. I suggest that a vocalic gesture is far more likely to be audible in a less overlapped loosely coordinated cluster than a more overlapped closely coordinated one.

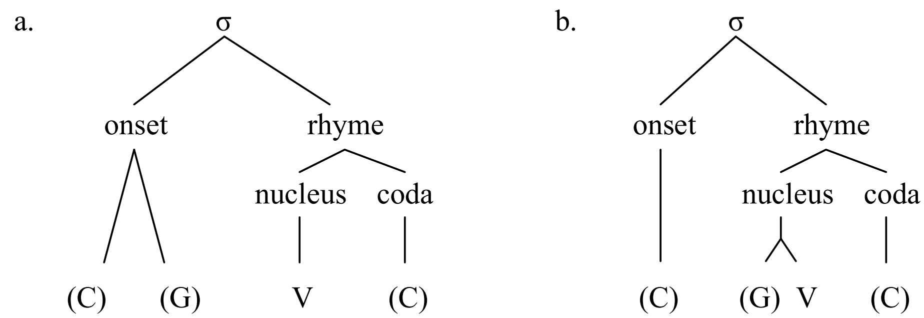
To sum up, acoustic phonetic evidence related to voicing, the vocalic quality of glides, and transition from C_1 to C_2 as well as evidence from gestural timing and overlap are supportive of the empirical findings in Experiment 1 with regard to anaptyxis being more problematic in [obstruent+glide] clusters than in [sibilant+stop] clusters.

4.4. Empirical Evidence

There is a dearth of L2 perceptual studies that deal with variable success as a function of the cluster itself in the learning of non-native consonant clusters. However, there exists some indirect empirical evidence from research on loanword phonology as well as consonant cluster production in L2 and child speech for the relative markedness of onset clusters stated in (1). The following briefly reviews some of these findings.

4.4.1. [obstruent+glide] vs. [obstruent+liquid]

Empirical evidence examining [obstruent+glide] clusters with regard to other non-glide onset clusters is very scarce. The reason for this is perhaps the highly debatable status of glides in the syllable onset: glides, as a natural class of sounds, have been argued to belong either to vowels, analyzable as part of the syllable nucleus (Spencer 1996) as illustrated in Figure 4.6a, or to consonants and therefore analyzed as part of the onset in a cluster (Borden et al. 2003) as in Figure 4.6b:



C= consonant, G= glide, V= vowel)

Figure 4.6 Glide syllable structure: (a) part of the onset, (b) part of the nucleus

As a result, glides are often excluded altogether from the investigation of onset clusters (e.g. Kreitman 2006, Kwon 2006) or agglomerated into the larger category [obstruent+sonorant] that subsumes both [obstruent+glide] and [obstruent+liquid] clusters (e.g. Fleischhacker 2005, Kim 2000).

Despite the dubious status of glides and the tendency of researchers to preclude them in the analysis of onset clusters, there exists some evidence that sustains the markedness of [obstruent+glide] over [obstruent+liquid] clusters. The first piece of evidence comes from Zuraw (2007) who examines the infixation of the infinitival and realis aspect markers /-um-/ and /-in-/ in a corpus of Tagalog loanwords that contain consonant clusters. In comparing the frequency of infixation in [stop+glide] and [stop+liquid] clusters, Zuraw concludes that loanwords with [stop+glide] clusters are significantly more likely to undergo vowel epenthesis splitting (anaptyxis) than loanwords with [stop+liquid] clusters, which appear to be immune to this type of infixation. In addition, Zuraw demonstrates that in Tagalog partial reduplication [stop+glide] clusters are reduced more frequently than are [stop+liquid] clusters. Finally, a survey of the corpus also reveals, consistent with previous generalizations made in Fleischhacker (2005), how [sibilant+glide] (e.g. [sw-]) clusters attracted the greatest number of anaptyctic insertions.

Further proof of the relative markedness of [obstruent+glide] over [obstruent+liquid] clusters comes from Eckman and Iverson's 1993 comparative study of English onset clusters as produced by Japanese, Korean and Cantonese speakers. Eckman and Iverson note that [obstruent+glide] clusters were consistently more difficult for subjects to acquire than [obstruent+liquid] clusters (with acquisition threshold set at 80%, after Eckman et al. 1989). That onset [obstruent+glide] clusters turned out to be most difficult for Japanese, Korean and Cantonese speakers is rather surprising since these clusters are arguably sanctioned in each of these languages. While native (or near-native) production of [C+glide] clusters is expected due to positive L1 transfer in this case, the sheer markedness of [C+glide] clusters could possibly have offset any benefit bestowed by the native language, giving rise to their attested difficulty in the Eckman and Iverson study. In fact, theoretical basis for the relative markedness of [obstruent+glide] clusters over [obstruent+liquid] clusters is found in Clements' (1990: 305) Sonority Dispersion Principle, which states that sonority should increase sharply and steadily from onset to nucleus, as observed in Eckman and Iverson (1993: 248):

> … obstruent-glide onsets emerge as more complex, or more marked, than obstruent-liquid onsets, because the rise in sonority from obstruent to liquid to vowel makes a steadier progression than does the initially sharp, then nearly flat, increase from obstruent to glide to vowel.

Eckman and Iverson (1993) was a rebuttal to earlier claims made by Broselow and Finer (1991) that tested their Minimal Sonority Distance (MSD) parameter by the production of English onset clusters by Japanese and Korean speakers of English (see §2.3.2 for detailed discussion). Broselow and Finer predicted on MSD grounds that the smaller sonority distance between obstruents and liquids, when compared with obstruents and glides, would render [obstruent+liquid] clusters more complex than [obstruent+glide] clusters. They claim that their counter-Clementsian hypothesis is confirmed when 32 subjects reportedly incurred more production errors on [obstruent+liquid] than [obstruent+glide] clusters.

However, it is possible to take issue with this finding on a number of points. First, it is not clear if the relative ease subjects experienced in the production of [obstruent+glide] clusters is in fact caused by higher MSD values or transfer from the L1, since, as Broselow and Finer admit, [obstruent+glide] clusters are the only possible kind of onset clusters in Japanese and Korean. On the contrary, for subjects to show any difficulty with [obstruent+glide] clusters can be indicative of the high markedness status of these clusters. As mentioned earlier, even though aided by positive L1 transfer, subjects still experienced some difficulty in producing [obstruent+glide] clusters.

Broselow and Finer's (1991: 41-42) argument for the lesser markedness of [obstruent+glide] clusters can be further vitiated if we consider the number and type of errors subjects made on these clusters. On the basis of cluster simplification either via vowel insertion or consonant deletion, and excluding phonemic substitution, the [obstruent+glide] clusters [py-] and [fy-] actually had a slightly greater number of errors than [pr-] and [fr-], respectively (3 vs. 2 and 7 vs. 6), with the exception of [by-] which significantly was harder than [br-] (15 vs. 5). Thus, it is clear that the reported markedness of [obstruent+liquid] over [obstruent+glide] clusters is largely motivated by non-syllabic modifications, which still render the cluster prosodically intact. Finally, the epenthesis/deletion error rates by 11 Hindi speakers producing the same set of English clusters cited in Broselow and Finer (48) suggest greater difficulty with [obstruent+glide] than with [obstruent+liquid] clusters (4 vs. 1 total errors).

4.4.2. [obstruent+liquid] vs. [sibilant+sonorant]

Although there are not many studies that examine the markedness between these two cluster types, some of the findings in the acquisition of first and second language point to the markedness of [obstruent+liquid] clusters relative to [sibilant+sonorant] clusters. For example, Barlow (2001) monitors developmentally the acquisition of consonant clusters in her subject KR who suffers a phonological disorder. She reports delayed acquisition of [obstruent+liquid] clusters compared to [sibilant+sonorant] clusters, which were acquired much earlier. Her results are consistent with findings from an earlier study (Barlow 1997) that showed [sibilant+C] clusters being produced more correctly (i.e. native-like) than [non-sibilant+C] clusters, which underwent reduction to singletons (similar conclusions can be found in Fikkert 1994, Freitas and Rodrigues 2003, Gierut 1999).

Looking at the acquisition of English consonant clusters by Korean speakers, Kim (2000) finds partial support for the unmarkedness of [sibilant+sonorant] over [stop+liquid]. This relationship, however, obtained only in lower proficiency subjects (higher proficiency subjects in fact displayed the opposite trend). Similarly, Kwon's (2006) examination of 30 Korean native speakers' production of various English onset clusters provides further evidence for the claimed markedness relation between [obstruent+liquid] and [sibilant+sonorant] clusters. Vowel epenthesis errors comprised most of the modifications (there were very few cases of consonant deletion), which were more common in [stop+liquid] rather than [sibilant+sonorant] clusters: 2.5% error rate in [sibilant+sonorant] clusters compared to 8.2% in [stop+liquid] clusters.

Finally, more solid evidence comes from Chan (2006), who evaluates the production of English onset clusters by 12 native speakers of Cantonese in a multi-tasked experiment that consisted of wordlist reading, picture description, passage reading, and conversational interviews. In all four tasks used to elicit onset cluster production, the percentage of correct consonant clusters obtained by subjects was consistently lower for [obstruent+liquid] clusters than for [sibilant+sonorant] clusters (69.4% vs. 95.2%). The success rate for each individual cluster is broken down as follows:

obstruent+liquid	[br-]	[bl-]	[dr-]	[fr-]	[fl-]	[gr-]	[gl-]
%	82.7	65.5	51.9	68.6	74.7	81.2	66.1
sibilant+sonorant	[sn-]	[sm-]	[sl-]				
%	92.6	99.9	88.7				

Table 4.5 Production of English onset clusters by native speakers of Cantonese (Chan 2006: 338)

Chan's results show that not only are correct [sibilant+sonorant] clusters more than correct [obstruent+liquid] clusters, but every [obstruent+liquid] cluster is outstripped by every [sibilant+sonorant] cluster. The fact that this refined division between [obstruent+liquid] and [sibilant+sonorant] is maintained by subjects across multiple tasks lends more credence to the claimed markedness hierarchy.

4.4.3. [obstruent+liquid] vs. [sibilant+stop]

There is plentiful evidence for the claim that [sibilant+stop] clusters occupy a lower place on the markedness scale than [obstruent+liquid] clusters. On the order of developmental acquisition, for example, subjects with phonological detriments tend to acquire [sibilant+stop] clusters, among other [sibilant+C] clusters, well in advance of [stop+liquid] clusters (Barlow 1997, 2001). In the speech of children acquiring European Portuguese, Freitas and Rodrigues (2003) report a strong tendency for reducing [obstruent+liquid] clusters by means of anaptyxis, which dwindles in [sibilant+stop] clusters although both clusters exist in Portuguese.

In second language acquisition both Kim (2000) and Chen (2003) independently provide support for the hypothesis that [sibilant+stop] clusters are less complex than [obstruent+liquid] clusters. Kim (2000) examined an array of English consonant clusters as realized by 16 native speakers of Korean to find out that [stop+liquid] clusters undergo vowel epenthesis more often than [sibilant+stop] clusters. A similar pattern in the acquisition of English onset clusters is exhibited by Chinese learners by Chen (2003). In agreement with these two studies is Chan (2006), who tested Cantonese speakers' production of English consonant clusters to reveal how the near-correctness with which onset [sibilant+stop] clusters were produced (93.4%) declined considerably in the production of [obstruent+liquid] clusters (69.4%).

Moreover, Major (1996) reported that his native speakers of Brazilian Portuguese had more difficulty producing English [fricative+liquid] onset clusters relative to [sibilant+stop] clusters. While the error ratio in [fricative+liquid] clusters was 70%, [sibilant+stop] clusters incurred only 37% of the total errors. Major's findings on the production of these two onset clusters are corroborated by Tench's (2003) perceptual study. Tench gauges Korean listeners' ability to perceive English segments and onset clusters accurately by asking them to listen to previously recorded words and write down what they think they heard. In the perception of onset clusters, [fricative+liquid] onsets were perceived correctly 80% of the time, while [sibilant+stop] onsets were perceived 100% correctly. Finally, Kwon's (2006) study of Korean speakers of English has some evidence for the markedness of [obstruent+liquid] clusters over [sibilant+stop] clusters, although error rates were small for both. Korean learners of English produced [stop+liquid] incorrectly 8.2% of time, but their incorrect production of [sibilant+stop] was highly marginal (0.6% incorrect).

4.4.4. [sibilant+sonorant] vs. [sibilant+stop]

The claim that [sibilant+stop] clusters are the least marked among onset clusters, and especially less marked than [sibilant+sonorant] clusters, despite their violation of the SSP (Clements 1990, Selkirk 1982, Steriade 1982) is well-motivated. In first language acquisition, it has been observed that [sibilant+stop] sequences are often acquired by children before other types of clusters, including [sibilant+sonorant] (Gierut 1999). Fikkert (1994) has reported on the emergence of [sibilant+C] prior to [non-sibilant+C] in the speech of Dutch children. Kirk and Demuth (2005) demonstrate that the majority of their child participants acquired [sibilant+stop] clusters before [sibilant+nasal]: [sibilant+stop] clusters were more accurately produced than [sibilant+nasal] clusters, 45% vs. 33% respectively. Abrahamsson (1999) longitudinally investigates the Spanish-speaker production of various Swedish sibilant onset clusters. The data, which consisted of natural speech conversations recorded over a period of time, show a clear markedness of [sibilant+sonorant] over [sibilant+stop] clusters. Vowel epenthesis was present in [sl-] 75% of the time compared to 59% in [sibilant+stop] clusters. Although this significantly large number of vowel-epenthesis errors rates for [sl-] subsided for [sibilant+nasal] clusters (54%), it did not differ significantly from that for [sibilant+stop] clusters (Abrahamsson 1999: 491-92).

Major (1996) tests adult Brazilian Portuguese speakers of English on the production of consonant clusters. The findings reported therein indicate a preference for [sibilant+stop] clusters over [sibilant+sonorant] clusters. In producing English onset clusters, Portuguese subjects altered [sl-] onsets more often than [sp-], [st-], and [sk-] onsets. Korean speakers of English display the same preferential treatment of [sibilant+stop] over [sibilant+sonorant] onset clusters. In Kim (2000), Korean speakers systematically modified [sibilant+nasal] as well as [sibilant+liquid] clusters more frequently than [sibilant+stop] clusters as judged by their vowel epenthesis errors, which were virtually absent in [sibilant+stop] clusters. Similarly, the Korean learners of English in Kwon (2006) demonstrated a pattern when producing onset clusters that is compatible with the claimed unmarkedness of [sibilant+stop] relative to [sibilant+sonorant]. Error analyses for onset clusters show that vowel epenthesis in [sm-] and [sl-] totalled 2.5%; however, the error rate in [sibilant+stop] clusters was highly trivial in comparison (0.6%).

Other studies on consonant clusters point to the relative unmarkedness of [sibilant+stop] clusters. Tench (2003) maintains that Korean learners of English perceive [sibilant+stop] clusters natively (100% accuracy rate), compared to other types of onset clusters, and Morelli (1999, 2003) notes the relative typological unmarkedness of [fricative+stop] and especially [sibilant+stop] clusters. Similar results are found by Chan (2006) and Zuraw (2007), among others.

4.4.5. Anaptyxis-Prothesis Asymmetry

The strongest empirical evidence for the anaptyxis/prothesis contrast (Figure 4.4) observed in this study is found in Fleischhacker (2001, 2005). In her dissertation, which investigates vowel epenthesis and consonant deletion patterns in onset clusters, she argues that the contrast of anaptyxis in [obstruent+sonorant] (which subsumes [obstruent+glide]) and prothesis in [sibilant+stop] observed in reduplication, loanword adaptation, and second language data across a number of languages has a perceptual basis. In the grammar of listeners there is preference for the retention of forms that are perceptually similar to their sources (see Kenstowicz 2003 for a similar proposal). Fleischhacker hypothesizes that anaptyxis in [obstruent+sonorant] clusters but prothesis in [sibilant+stop] clusters provide maximal similarity to the unepenthetic forms of the cluster.

To test this hypothesis, Fleischhacker (2001) aurally presented 49 native English-speaking listeners with anaptyctic and prothetic modifications of [obstruent+sonorant] and [sibilant+stop] clusters and asked them to rate on a 1-7 scale (1 = least similar, 7 = most similar) the similarity of each to the original unmodified cluster. Subjects also had to provide preferential judgments between anaptyxis and prothesis as to how much they liked each epenthetic form over the other. The results showed that higher similarity ratings were given for prothesis in [sibilant+stop] clusters as opposed to anaptyxis which was rated more similar than prothesis in [obstruent+sonorant] clusters. Listeners also preferred anaptyxis over prothesis in [obstruent+sonorant] clusters, but preferred prothesis over anaptyxis in [sibilant+stop] clusters (see Wingstedt and Shulman 1988 for similar findings with English and Swedish subjects). Typologically, Fleischhacker (2005: 40-41) demonstrates how this pattern of anaptyxis vs. prothesis is evident in interlanguage phonology (e.g. loanword adaption of Egyptian Arabic, Sinhalese, Amharic, Bengali, Kirgiz, etc.).

Drawing on a different kind of evidence, Fleischhacker (2005: 110-111) shows how the frequency with which English imperfect puns[19] occur is reflective of the anaptyxis/prothesis contrast in [obstruent+sonorant] and [sibilant+stop] onset clusters. Examining a corpus of 1,964 imperfect puns involving vowel epenthesis, Fleischhacker computes the frequency of [obstruent+sonorant] and [sibilant+stop]

---

[19] A perfect pun can be constructed of two different words that are phonetically identical (e.g. *bark-barque*), while an imperfect pun is made up of two words which are phonologically similar (e.g. *Clare-éclair*). In both cases, the words are lexically distinct (Fleischhacker 2005: 80).

clusters in anaptyctic as well as prothetic puns. Calculations of the observed over expected frequencies in the corpus revealed that [obstruent+sonorant] clusters are rather frequent in anaptyxis-based puns (e.g. *broke-baroque*), but completely lacking in prothesis-based puns (e.g. *claim-acclaim*). The Observed over Expected frequency O/E was 1.24 vs. 0.[20] In stark contrast, the frequency of [sibilant+stop] clusters was common (O/E= 1.13) in prothetic puns (e.g. *steam-esteem*), but significantly lower (O/E= 0.29) in anaptyctic puns (e.g. *sport-support*). Furthermore, Fleischhacker presents evidence from poetic alliteration in Germanic, Middle English, and Old Irish in support of the contrast between anaptyxis and prothesis in [obstruent+sonorant] and [sibilant+stop] onsets.

Kwon (2006) conducts a perception experiment to assess the auditory similarity between English onset clusters and their epenthetic counterparts as judged by 38 Korean speakers of English. Although Kwon looked at the effects of anaptyxis only on perceptual similarity, the results of her experiment lend strong support to the claims made in this dissertation with regard to cluster type being a determinant of subjects' sensitivity to anaptyxis vs. prothesis. Pairing nonsense consonant cluster words with their epenthetic forms, Kwon presented subjects with two sets of pairs containing different clusters and asked them whether the first or second pair sounded more similar to each other, e.g. [br]/[bər] vs. [st]/[sət]. Pairwise comparisons show that compared to onset sibilant vs. stop clusters, [stop+liquid] clusters were judged more similar to their anaptyctic forms. Below is a summary of the findings by cluster type:

[20] An Observed/Expected frequency value of 1 indicates the pun occurs as often as expected.

Cluster pairing	Similarity rating	**sig.*
st-pr	16% - 84%	*p ≤ .01
st-br	12.6% - 87.4%	*p ≤ .01
sm-pr	28.9% - 71.1%	*p ≤ .01
sm-br	28.9% - 71.1%	*p ≤ .01
sl-pr	39.5% - 60.5%	*p ≤ .1
sl-br	44.7% - 55.3%	*p > .1

Table 4.6 Cluster-dependant similarity ratings of vowel epenthesis (anaptyxis only) for Korean speakers (Kwon 2006: 182)

As clear from Table 4.6, Korean subjects consistently rated [pr] and [br] onsets as being perceptually more similar to their anaptyctic counterparts. In contrast, the similarity ratings for [sibilant+stop] and [sibilant+sonorant] clusters are much less. Kwon notes that as the sonority of the second element in the cluster increases, listeners tend to judge the cluster and its epenthetic form as being less distinct. Thus, differences in similarity are more pronounced for [sibilant+stop] than for [sibilant+sonorant]. This finding is particularly interesting since in the current study anaptyxis errors were significantly fewer in [sibilant+stop] clusters (31.58%) than in [sibilant+sonorant] clusters (54.22%) (see Table 4.4).

Supporting evidence comes also from some consonant cluster production studies which confirm the predominance of prothesis in [sibilant+C] clusters. Broselow (1987), for example, found that Egyptian Arabic speakers of English use prothesis to repair [sibilant+stop] onset clusters, but anaptyxis to repair other types of clusters (also Broselow 1992a). Moreover, Carlisle (1997) concludes that in the acquisition of English onset clusters Spanish speakers apply prothesis to [sibilant+C] clusters only.

Before concluding this section, I would like to note that the vowel epenthesis contrast (anaptyxis vs. prothesis) of Table 4.4 of this study is consistent with claims made outside the realm of phonetics. For instance, it has been argued by Singh

(1985), who investigates Hindi production of English consonant clusters, that the profile of the cluster determines the location of the epenthetic vowel: consonants in an onset cluster with rising sonority as in [obstruent+glide] (e.g. [kw-]) trigger anaptyxis, while consonants in an onset cluster with non-rising sonority such as [sibilant+stop] (e.g. [sk-]) trigger prothesis. Later, Gouskova (2001) formalizes this generalization under the syllable contact law (after Venneman 1988) which stipulates that coda-onset sequences should decrease in sonority. According to Gouskova's prothesis-based look-ahead argument, when prothesis applies to [sibilant+stop] clusters, it constitutes no violation of syllable contact, i.e. sonority declines from coda to onset as in [əs.tek]. However, prothesis in [obstruent+glide] clusters reverses the preferred sonority profile as in [ək.win], which contains a less sonorous coda than the following onset.

4.5. Summary

This chapter has explored empirically how SA non-native learners of English perceive onset consonant clusters with regard to vowel epenthesis. Results from the identification task show listeners are sensitive to the type of cluster both when presented with vocalic and non-vocalic stimuli. Subjects' responses also indicate a clear contrast in the location of vowel epenthesis: listeners exhibit a bias toward hearing anaptyxis in [obstruent+glide] clusters, but prothesis in [sibilant+stop] clusters. Based on the findings of the study, a hierarchy of markedness which ranks English onset clusters in order of perceptual difficulty was derived. Further support was presented for this hierarchy from phonetic and other experimental research.

CHAPTER 5

A TEST OF THE PROPOSED MARKEDNESS HIERARCHY

5.1. Introduction

In Chapter 4, it was concluded that SA listeners perceive different English onset clusters with varying degrees of difficulty. Auditory sensitivity to vowel epenthesis in non-native perception differed as a function of the cluster type. In some onset clusters, SA listeners also demonstrated a bias in perceiving the location of vowel epenthesis: anaptyxis was preferred in [obstruent+glide] clusters and prothesis in [sibilant+stop] clusters. Based on the pattern by which SA listeners perceived different onset clusters, a hierarchy of markedness was motivated and argued for, reiterated here in (1):

(1) Perceptual hierarchy of markedness among onset clusters

less marked ←——————→ *more marked*

[sibilant+stop] >> [sibilant+sonorant] >> [obstruent+liquid] >> [obstruent+glide]

The purpose of this chapter is to evaluate different aspects of this markedness hierarchy empirically. The chapter is divided into three parts. In §5.2 the effect of vowel durational differences on listeners' ability to perceive anaptyxis in marked vs. unmarked onset clusters is investigated. §5.3 presents an auditory similarity judgment experiment that tests the cluster-dependent anaptyxis/prothesis contrast reported earlier in Chapter 4. Finally, a production experiment in §5.4 tests the validity of the perceptual markedness hierarchy in the English speech of SA learners. A general discussion of the results follows in §5.5. The chapter concludes with a summary of the main findings in §5.6.

5.2. Experiment 2: Sensitivity to Vowel Epenthesis Duration in Onset Clusters

This perceptual experiment is designed to assess how manipulation of the duration of an anaptyctic vowel impacts listeners' ability to perceive vowel epenthesis in various onset clusters. The experiment also investigates the interaction between vowel length and cluster markedness: does the degree of markedness of a cluster affect how longer vs. shorter durations of the epenthetic vowel are perceived by SA listeners? In other words, is vowel length related to cluster markedness? Before attempting an answer to these questions, first I lay out the relevant hypotheses, and general design of the experiment.

5.2.1. Hypotheses

Experiment 2 tests the following 3 hypotheses:

1. SA listeners will better distinguish onset clusters from their anaptyctic forms the longer the duration of the epenthetic vowel.
2. Distinguishing more marked clusters from their anaptyctic forms will require *longer* vowel durations than would discrimination in less marked clusters.
3. Overall discriminability of consonant clusters will be better in higher proficiency than in lower proficiency learners.

Hypothesis 1 generally aims to determine just how sensitive SA listeners are to the presence of anaptyxis in onset clusters. It predicts an effect of manipulated vowel duration on listeners' ability to hear vowel epenthesis. More particularly, it predicts that increasing vowel duration incrementally would have a positive effect on listeners' ability to detect anaptyxis, leading to more successful discrimination between a cluster and its anaptyctic form.

The hypothesis is based on similar findings reported in Dupoux et al. (1999), who demonstrate a linear relationship between vowel length and the ability of Japanese and French listeners to perceive vowel epenthesis in medial consonant clusters: the longer the duration of the epenthetic vowel, the more successful Japanese and French listeners were in detecting it. The current hypothesis further tests this in a different prosodic environment (i.e. onset position) and employing a new listener group.

Hypothesis 2 examines the relationship between vowel length and relative markedness of onset clusters with regard to anaptyxis. Based on the results of Experiment 1 in which identifiability of anaptyxis was best in [sibilant+stop] and worst in [obstruent+glide] clusters, it is predicted that listeners would be more sensitive to anaptyxis in the former but not the latter. The prediction for [sibilant+sonorant] and [obstruent+liquid] clusters, however, is less clear since markedness between these two clusters was motivated by overall vowel epenthesis identification errors rather than by anaptyxis or prothesis alone. Nonetheless, given the association of anaptyxis with markedness (i.e. anaptyxis was most common in clusters that incurred the highest number of errors), I predict a tendency for listeners to be more perceptive of anaptyxis in [sibilant+sonorant] than [obstruent+liquid] clusters. In this experiment sensitivity to vowel epenthesis is gauged by varying the duration of the inserted vowel along a continuum, and listeners are expected to detect anaptyxis on the vowel continuum earlier in relatively unmarked clusters than in marked clusters.

Hypothesis 3 tests the effect of proficiency on listeners' ability to correctly discriminate onset clusters from their anaptyctic counterparts. It is founded on research that attributes perceptual as well as productive gain to linguistic proficiency in the L2 learning process (e.g. Abrahamsson 2003, Bradlow et al. 1997, Flege 1995, Riney and Flege 1998).

5.2.2. Methodology

5.2.2.1. Participants

Participants were the same as in Experiment 1.

5.2.2.2. Materials and Task

The materials for this experiment were a continuum of five different vowels for the clusters [kl-], [kw-], [sw-], [fl-], [sk-], [sn-], [sl-], each representative of the 7 cluster categories discussed in Experiment 1.[21] For each cluster, a 5-step continuum of vowel length ranging from 16 to 80 ms was interpolated by synthetically manipulating the duration of the schwa vowel in the anaptyctic cluster condition, taken from the same materials used for Experiment 1. PSOLA[22] in Praat (version 5.0.32) was used to shorten and lengthen the epenthetic vowel by 16 ms intervals. Each of the 5 vowel lengths was paired with no vowel, the unepenthesized condition of the cluster also taken from the materials used for Experiment 1: vowels of 16, 32, 48, 64, and 80 ms each compared with no vowel (i.e. no anaptyxis).

Using a forced choice AXB discrimination task, four 3-word test items (AAB, ABB, BAA, BBA) were generated for each of the 5 pairings, yielding 20 randomly ordered test trials for each cluster. For example, the cluster [kl-] had 5 anaptyctic vowels and for each 4 test trials were generated ([kəlit-kəlit-klit], [klit-klit-kəlit], [kəlit-klit-klit], and [klit-kəlit kəlit]). An AXB discrimination task is believed to provide a better tool in tapping listeners' ability to perceive vowel epenthesis

---

[21] The selection of these clusters was based on several factors such as the frequency and voicing status of the cluster (see Appendix E).

[22] See footnote 16.

durational differences. It provides a reference point (i.e. X) against which the similarity of stimuli can be gauged, as opposed to a simple AX discrimination task where listeners may base their same/different responses on non-linguistic factors (Beddor and Gottfried 1995).

5.2.2.3. Procedure

The experiment was set up in Praat's Experiment MFC which aurally presented listeners with 280 experimental trials in a random fashion (7 clusters X 5 pairs X 4 trials X 2 reps). Each trial proceeded as follows. The participant heard all three words over headphones (Koss R80) and had to indicate whether the first or third word was the same as the second by clicking on one of two boxes labeled *first* and *third* visible on a computer screen. In each triad, an ISI (inter-stimulus interval) of 250 ms separated each token from the other. Shorter ISI is believed to promote phonetic processing as opposed to phonemic perception encouraged by longer ISI (Werker and Logan 1985). The next trial began automatically once a selection was made. The experiment was completed over two sessions (140 trials each), with optional 5-minute breaks every 70 trials. To make sure subjects understood the procedure, a 3-item practice test preceded the actual experiment, which took place in a quite room and lasted less than 30 minutes for each subject.

5.2.2.4. Results

Proficiency

Discrimination scores from the AXB task for each subject group revealed discriminability was better in the advanced subject group than in the beginners, 66.42% vs. 52%. To determine whether differences in discriminability due to proficiency were significant, an independent sample *t*-test compared mean scores from each subject group. Results show the difference was significant: $t(36.85) = -3.29$, $p < .05$ ($p = .002$).

Duration

Discrimination scores were pooled for all cluster types and analyzed by duration for each subject group. Independent one-variable pairwise Chi-square analyses for each subject group were performed for the variable duration. Results are summarized in Table 5.1 and 5.2 (Chi-square values in parentheses):

Duration	Discrimination %	32ms	48ms	64ms	80ms
16 ms	14.79	*($\chi2$ =22)	*($\chi2$ =77.80)	*($\chi2$ =201.82)	*($\chi2$ =248.98)
32 ms	28.96		*($\chi2$ =13)	*($\chi2$ =104)	*($\chi2$ =142)
48 ms	42.92			*($\chi2$ =46.20)	*($\chi2$ =73.72)
64 ms	**76.87**				*ns*
80 ms	87.71				

* = $p < .05$

Table 5.1 Discrimination by vowel epenthesis duration in beginning listeners

Duration	Discrimination %	32ms	48ms	64ms	80ms
16 ms	22.92	*($\chi2$ =47.08)	*($\chi2$ =150.38)	*($\chi2$ =189.62)	*($\chi2$ =218.32)
32 ms	49.58		*($\chi2$ =33.44)	*($\chi2$ =55.18)	*($\chi2$ =72.76)
48 ms	**79.58**			*ns*	*($\chi2$ =7.94)
64 ms	89.50				*ns*
80 ms	96.66				

* = $p < .05$

Table 5.2 Discrimination by vowel epenthesis duration in advanced listeners

As clear from Table 5.1 and 5.2, overall discrimination rates were poor in shorter vowel durations and improved steadily as the duration of the vowel became longer. All pairwise comparisons were significant, except for 64-80 ms in the beginning group, and 48-64 ms and 64-80 ms in the advanced group, for which differences in discriminability were small and insignificant. For beginning listeners, discrimination picked up noticeably in the 64 ms condition (i.e. the point where discrimination rises sharply above 50% chance level). For advanced listeners, the 48 ms condition marked the shift in their ability to discriminate clusters correctly, as illustrated in Figure 5.1:

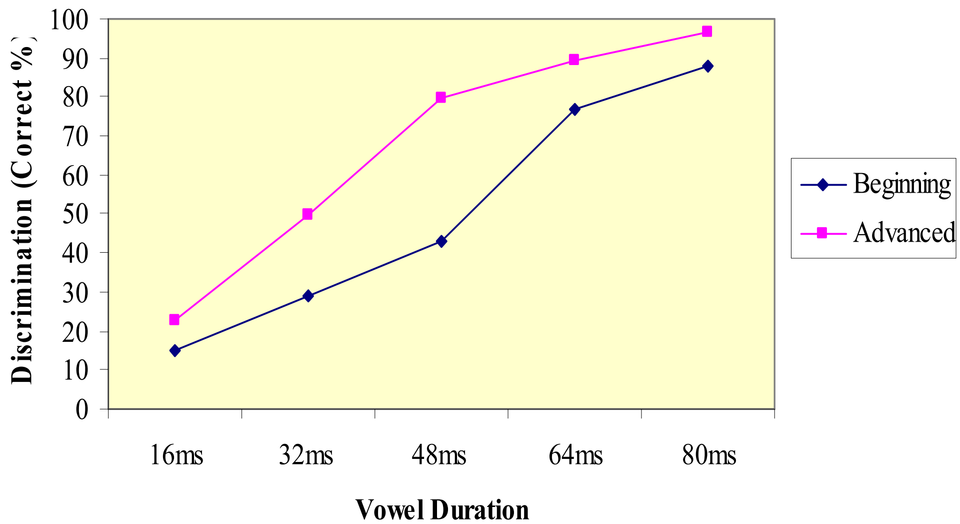


Figure 5.1 Cluster discrimination as a function of vowel duration for beginning and advanced SA listeners

Duration by Cluster

Discrimination scores from the 7 cluster types were aggregated, averaged and grouped according to the four cluster categories [obstruent+glide], [obstruent+liquid], [sibilant+sonorant] and [sibilant+stop]. For each subject group, one variable pairwise Chi-square tests compared discrimination rates between each pair of cluster types on every vowel duration. The following table presents beginners' discrimination rates by cluster category on each of the five steps of the vowel-length continuum:

16 ms	Discrimination %	obst+liquid	sib+son	sib+stop
obstruent+glide	12.5	*ns*	*ns*	*ns*
obstruent+liquid	14.37		*ns*	*ns*
sibilant+sonorant	13.12			*ns*
sibilant+stop	21.25			
32 ms				
obstruent+glide	27.5	*ns*	*ns*	*($\chi2$ =4.22)
obstruent+liquid	22.5		*ns*	*($\chi2$ =7.50)
sibilant+sonorant	28.12			*($\chi2$ =3.90)
sibilant+stop	45			
48 ms				
obstruent+glide	33.75	*ns*	*ns*	*($\chi2$ =11.24)
obstruent+liquid	37.5		*ns*	*($\chi2$ =8.56)
sibilant+sonorant	40.62			*($\chi2$ =6.68)
sibilant+stop	67.5			
64 ms				
obstruent+glide	65	*ns*	*ns*	*($\chi2$ =6.50)
obstruent+liquid	60.6		*($\chi2$ =17.52)	*($\chi2$ =5.30)
sibilant+sonorant	88.75			*($\chi2$ =8.57)
sibilant+stop	97.5			
80 ms				
obstruent+glide	85	*ns*	*ns*	*ns*
obstruent+liquid	82.5		*ns*	*ns*
sibilant+sonorant	90			*ns*
sibilant +stop	93.75			

* = $p < .05$

Table 5.3 Discrimination by cluster type and duration for beginning listeners

As the percentage column in Table 5.3 demonstrates, discriminability varied from one cluster to another. However, these differences were only substantial in the 32 ms, 48 ms, and 64 ms vowel durations, in which [sibilant +stop] was significantly discriminated from its epenthetic form better than [obstruent +glide], [obstruent +liquid], and [sibilant +sonorant] clusters. The 64 ms condition also showed significant increased discrimination for [sibilant +sonorant] clusters. This is plotted graphically in Figure 5.2:

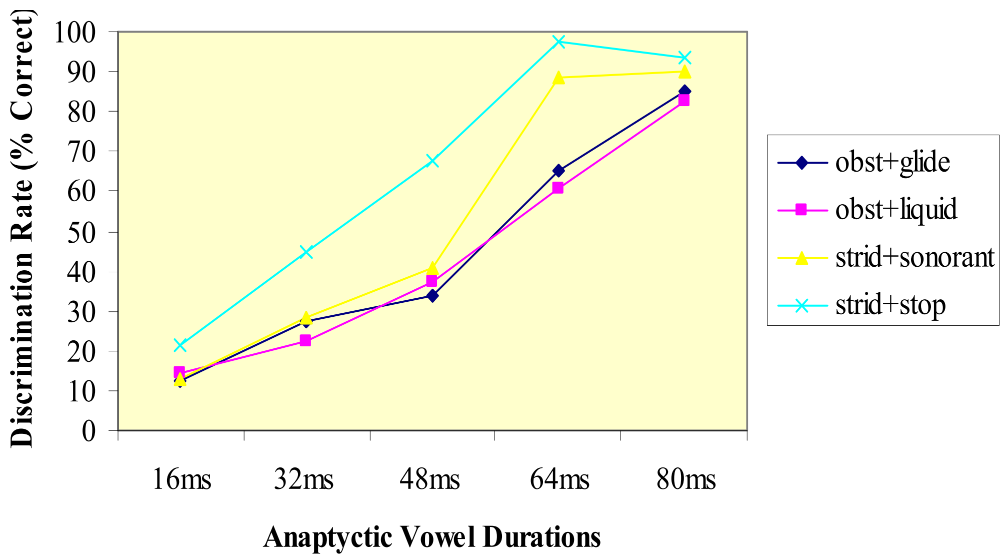


Figure 5.2 Discrimination as a function of cluster type and vowel duration for beginning listeners

For the advanced group, discrimination also followed from cluster type, although significant differences between different clusters occurred only in two conditions: 32 ms and 48 ms. Unlike discrimination in the 48 ms condition in which both [sibilant+stop] and [sibilant+sonorant] clusters were significantly better discriminated from [obstruent+liquid] and [obstruent+glide] clusters, discrimination in the 32 ms condition was only significant between [sibilant+stop] and [obstruent+glide] clusters. Discrimination rates and pairwise comparisons for the advanced group are summarized in Table 5.4, and illustrated in Figure 5.3:

16 ms	Discrimination %	obst+liquid	sib+son	sib+stop
obstruent+glide	21.25	*ns*	*ns*	*ns*
obstruent +liquid	20		*ns*	*ns*
sibilant+ sonorant	23.75			*ns*
sibilant +stop	28.75			
32 ms				
obstruent +glide	37.5	*ns*	*ns*	*($\chi2$=4.68)
obstruent +liquid	45.62		*ns*	*ns*
sibilant + sonorant	55			*ns*
sibilant +stop	58.75			
48 ms				
obstruent +glide	62.25	*ns*	*($\chi2$=6.80)	*($\chi2$=7.29)
obstruent +liquid	64.37		*($\chi2$=5.86)	*($\chi2$=6.32)
sibilant + sonorant	95			*ns*
sibilant +stop	96.25			
64 ms				
obstruent +glide	80	*ns*	*ns*	*ns*
obstruent +liquid	83.75		*ns*	*ns*
sibilant + sonorant	97.5			*ns*
sibilant +stop	95			
80 ms				
obstruent +glide	95	*ns*	*ns*	*ns*
obstruent +liquid	96.25		*ns*	*ns*
sibilant +sonorant	96.87			*ns*
sibilant +stop	98.75			

Table 5.4 Discrimination by cluster type and duration for advanced listeners

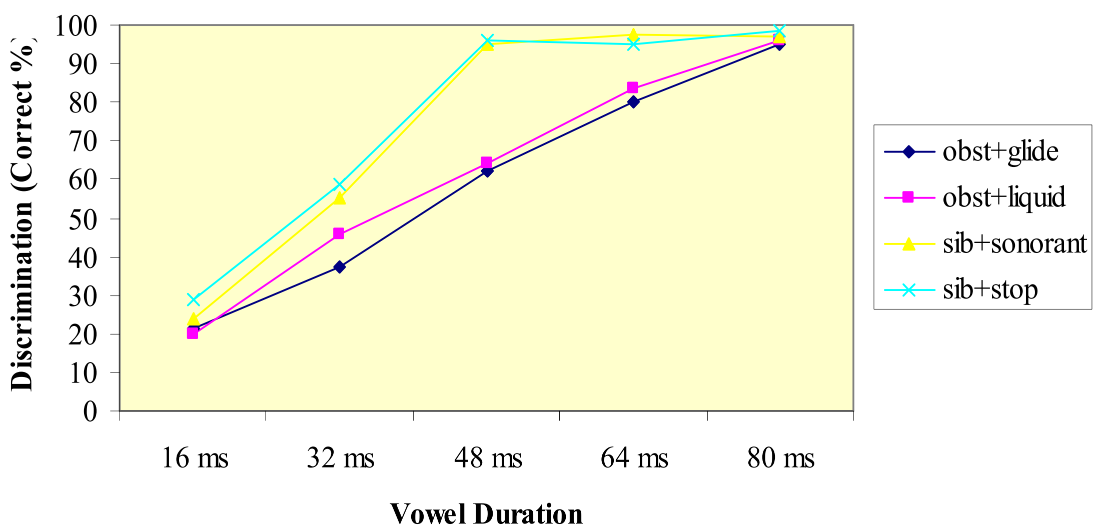


Figure 5.3 Discrimination as a function of cluster type and vowel duration for advanced listeners

5.2.2.5. Discussion

Hypothesis 1 claimed that SA non-native listeners would be better able to detect vowel epenthesis in onset clusters the longer the duration of the inserted vowel. Assuming that listeners in the AXB task discriminated stimuli on the basis of anaptyxis (or its absence), we expect discrimination rates to be greatest in the longest vowel duration condition of 80 ms, and least in the 16 ms condition. This transpires for both subject groups. Looking at Table 5.1 and 5.2, we see that beginning and advanced listeners' ability to correctly discriminate stimuli improved significantly as duration of the inserted vowel increased. This is naturally expected and is in conformity with previous findings.

In Dupoux et al. (1999), for example, native French and non-native Japanese listeners reported hearing vowel epenthesis more often in medial consonant clusters the longer the duration of the inserted vowel:

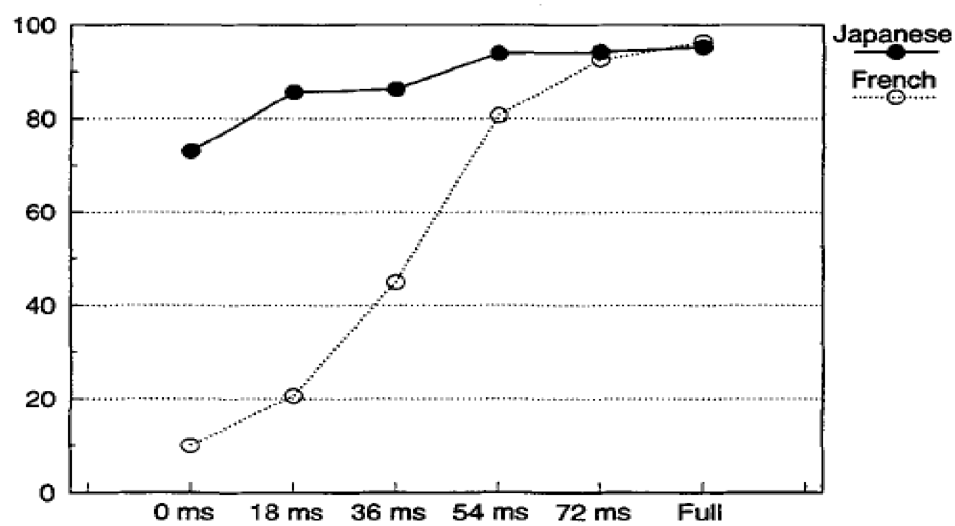


Figure 5.4 Japanese and French identification rates of the vowel /u/ (Dupoux et al. 1999: 1570)

Whereas more accurate /u/ identifiability was associated with longer vowel durations for both Japanese and French listeners, the effect of vowel duration was more pronounced in French listeners as clear from the steep rise in the identification function (see §3.6.2.2).

Similarly, in this study both SA listener groups exhibited a pattern of discrimination compliant with Hypothesis 1: the longer the duration of the anaptyctic vowel the more successful were listeners in discriminating stimuli correctly. For the lower proficiency group, this was very systematic except for the last two vowel conditions, in which the difference in discrimination proved insignificant although it increased from 76.6% in the 64 ms to 87.71% in the 80 ms condition. For higher proficiency learners, the differences between 48 ms vs. 64 ms, and 64 ms vs. 80 ms were insignificant, perhaps due to the ceiling effect of relatively high performance on the 48 ms condition.

An interesting point of difference between the two groups lies in the perceptual boundary which marks the distinction between unepenthesized clusters and their anaptyctic stimuli. Lower proficiency listeners seem to have their boundary set somewhere in the 64 ms condition, where discrimination improved drastically from 42.92% to 76.87%. As the vowel duration became shorter, discriminability on the basis of anaptyxis dwindled as evident in listeners' discriminatory rates on the 48 ms, 32 ms and 16 ms conditions. When the duration was 80 ms listeners became well-aware of anaptyxis and attained a high discrimination rate (87.71%). On the other hand, the perceptual shift for higher proficiency listeners occurred earlier on the vowel continuum, between 32 ms and 48 ms. As the vowel shortened, listeners' ability to discriminate stimuli correctly became noticeably poorer, 49.58% in 32 ms and 22.92% in 16 ms. In contrast, their ability to distinguish vocalic from non-vocalic stimuli above the 32 ms threshold significantly exceeded chance level.

Hypothesis 2 maintained that perceptual sensitivity to anaptyxis is related to the markedness status of the cluster in question: anaptyxis would be less perceptible in more marked clusters, in which case discrimination requires longer vowel duration than in less marked clusters. The results of the lower proficiency group show no significant interaction between duration and markedness in the 16 ms or the 80 ms conditions: all four types of clusters were equally discriminated, inaccurately when the duration was short (16 ms) and accurately when it was long (80 ms). In other words, discrimination under these two conditions was either bad or good for all clusters with no significant differentiation based on cluster type. Thus, Hypothesis 2 is neither supported nor disconfirmed in the 16 ms and 80 ms conditions.

Discrimination results from 32 ms, 48 ms and 64 ms lend partial support to the hypothesis: [sibilant+stop] has significantly higher discrimination rates compared to all other cluster types. That is, this cluster was successfully discriminated from its epenthetic counterpart even when the duration of the anaptyctic vowel was relatively short. This was not the case for other clusters. Further support for the hypothesis is also yielded by [sibilant+sonorant], which was significantly better discriminated than [obstruent+liquid] clusters in the 64 ms condition. For the higher proficiency group, Hypothesis 2 is sustained by [sibilant+stop] clusters, which were significantly better discriminated than [obstruent+glide] in 32 ms durations. However, most of the support for the hypothesis comes from the significant differences found in the 48 ms condition, namely the accurate discrimination by SA listeners of [sibilant+sonorant] and [sibilant+stop] clusters in comparison to [obstruent+liquid] and [obstruent+glide] clusters.

Hypothesis 3 which predicted an effect of proficiency in successfully discriminating clusters from their anaptyctic forms was upheld by the results of the experiment. SA listeners' ability to discriminate stimuli correctly varied as a function of their proficiency in English. The perceptual gain in discrimination skills demonstrated by advanced listeners in this study is in line with research emphasizing the important role of experience in the acquisition of a second language (Abrahamsson 2003, Bradlow et al. 1997, Flege 1995, Riney and Flege 1998).

To recap, Hypothesis 1 is well-supported by the results. The fact that the discriminatory boundary that signals the perceptual transition from no vowel to vowel occurs one step earlier on the continuum for advanced vs. beginning learners can be taken as evidence of their greater auditory sensitivity to anaptyxis. Whereas 48 ms of duration was sufficient for advanced listeners to correctly perceive anaptyxis in the discrimination task, beginners required at least 64 ms.

The predictions made by Hypothesis 2 are partially buttressed by the findings of this experiment. While durations at the lower and higher ends of the vowel continuum yielded no significant differences, due to discrimination being either extremely poor or excellent, none provided counter-evidence for the hypothesis in which more marked clusters significantly differed from those for less marked ones. It is also noteworthy that in each subject group the same vowel-duration conditions which yielded maximum support for Hypothesis 2 (i.e. which showed the greatest number of significant pairwise comparisons) also marked the perceptual boundary in discriminatory ability for that group: 64 ms for beginners and 48 ms for advanced. Discrimination was poorer for lesser durations but generally good in later steps of the continuum. Under both very poor and very good discrimination rates, significant differences are naturally hard to obtain because of ceiling and floor effects.

5.3. Experiment 3: Auditory Similarity Judgment of Anaptyxis vs. Prothesis

The purpose of this experiment is to test another aspect of the markedness hypothesis stated in (1) which pertains to vowel epenthesis location by exploring SA listeners' judgments of perceived similarity between onset clusters and their epenthetic forms.

5.3.1. Hypotheses

Experiment 3 aims at testing the following two hypotheses:

1. In relatively marked clusters, SA listeners will judge a cluster with anaptyxis rather than prothesis as being more similar to the unepenthesized form.
2. In relatively unmarked clusters, SA listeners will judge a cluster with prothesis rather than anaptyxis as being more similar to the unepenthesized form.

The hypotheses here are largely based on the results of Experiment 1 with regard to vowel epenthesis location and cluster markedness: if SA listeners in the identification task of Experiment 1 make more anaptyxis errors in relatively marked clusters and more prothesis errors in less marked clusters, then we would expect similarity judgments to follow this pattern of preference, i.e. judgments of greater similarity of anaptyctic forms in marked clusters and of prothetic forms in unmarked clusters. Given the cluster-markedness scale in (1), the prediction is that, consistent with the identification results of Experiment 1, [obstruent+glide] clusters would be rated more similar to their anaptyctic rather than prothetic forms, whereas [sibilant+stop] clusters would be rated more similar to their prothetic rather than anaptyctic forms. Although not borne out in Experiment 1, the hypotheses also predict a similar pattern between the more marked [obstruent+liquid] and less marked [sibilant+stop] clusters in listeners' judgments of anaptyxis and prothesis.

5.3.2. Methodology

5.3.2.1. Participants

Participants were the same as in Experiments 1 and 2.

5.3.2.2. Materials and Task

The materials for this experiment were drawn from the materials used in Experiment 1. Similarity judgments were elicited from SA listeners using a fixed AXB forced-choice discrimination paradigm. Two experimental trials (AXB, BXA) with an ISI of 250 ms were generated for each of the 52 consonant-cluster words. Every test trial was a triad consisting of three stimuli: the anaptyctic and the prothetic forms interposed by the unepenthesized cluster (e.g. [kəlit-klit-əklit] and [əklit -klit-kəlit] for the test word /klit/), all taken from the 156 wordlist recorded by the native speaker for Experiment 1.

5.3.2.3. Procedure

In Praat's Experiment MFC, each trial was played twice over headphones in a quiet room. After listening to each of the 208 randomly ordered trials (52 clusters X 2 trials X 2 reps), subjects indicated whether the first or the third word sounded more similar to the second by clicking on one of two boxes provided on a computer screen. The next trial followed as soon as a choice has been made. The experiment was administered over two sessions (104 trials per session) with an optional intervening 5-minute break. A 3-item practice test ensured subjects understood the instructions.

5.3.2.4. Results

To determine if the distribution of anaptyxis and prothesis responses for the different types of clusters was significant, total similarity ratings for each cluster type were submitted to tests of proportions. The results of the AXB discrimination task are represented graphically in Figure 5.5 for beginning and in Figure 5.6 for advanced listeners:

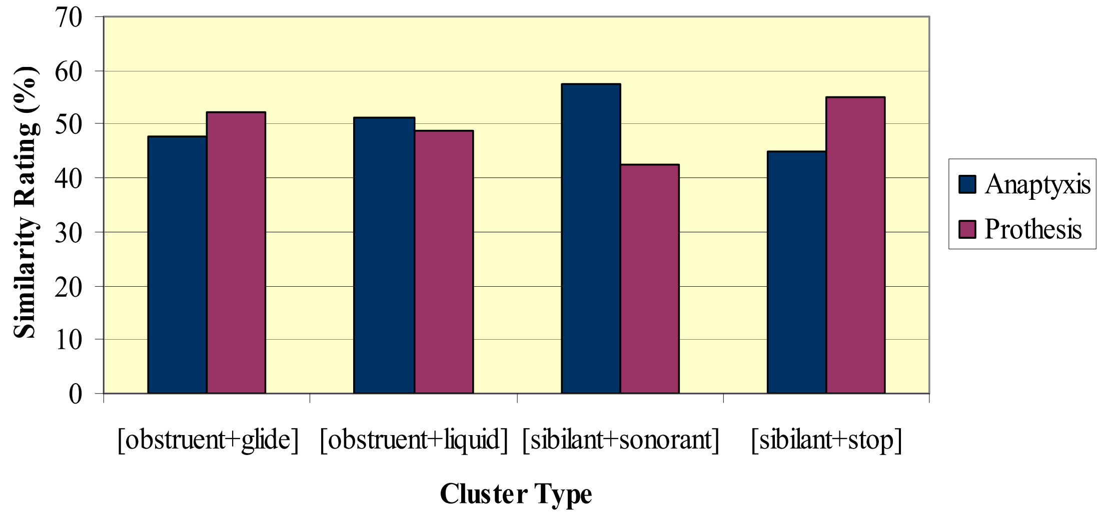


Figure 5.5 Beginning listeners' anaptyxis vs. prothesis similarity ratings by cluster type

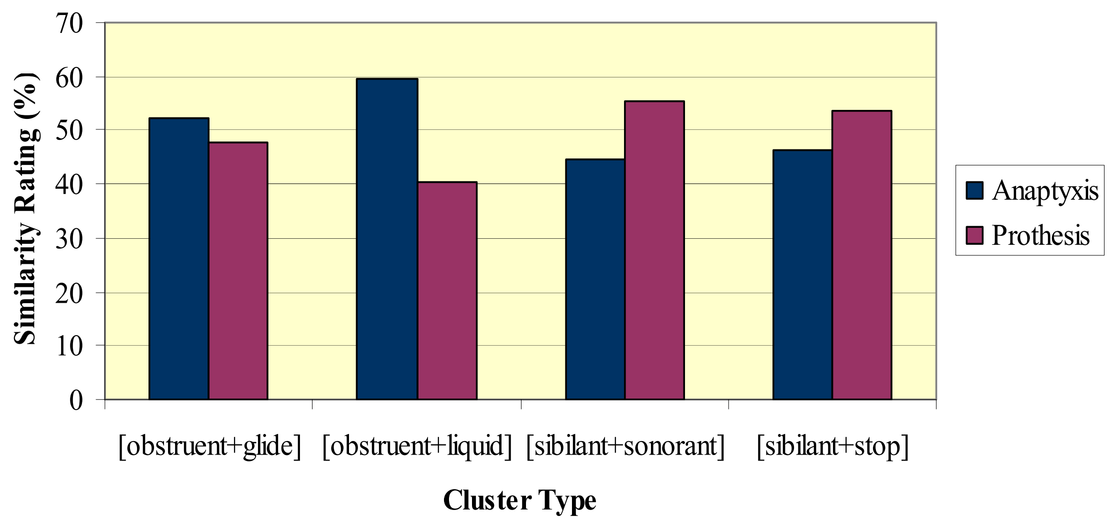


Figure 5.6 Advanced listeners' anaptyxis vs. prothesis similarity ratings by cluster type

Similarity ratings of anaptyxis and prothesis by cluster type are given below in Table 5.5:

	obstruent+glide	obstruent+liquid	sibilant+sonorant	sibilant+stop
Beginning				
Anaptyxis	47.78	51.25	57.5	45
Prothesis	52.22	48.75	42.5	55
sig.	*ns*	*ns*	*	*ns*
Advanced				
Anaptyxis	52.36	59.66	44.69	46.25
Prothesis	47.64	40.34	55.31	53.75
sig.	*ns*	*	*	*ns*

* = $p < .05$

Table 5.5 AXB anaptyxis vs. prothesis similarity ratings by cluster type

The distribution of anaptyctic and prothetic responses in the auditory similarity discrimination task was significant only for [sibilant+sonorant] clusters in both listener groups. However, anaptyxis was judged more similar to the unepenthesized cluster by the beginning group, but prothesis by the advanced group. The distribution was also significant although only in the advanced group for [obstruent+liquid] clusters, for which anaptyxis was judged more similar. The rest of the clusters show a weak tendency toward one type of vowel epenthesis or the other with no significant differences.

5.3.2.5. Discussion

The hypotheses in this experiment predicted ratings of more similarity for anaptyxis with relatively marked clusters, and ratings of more similarity for prothesis with relatively unmarked clusters. A cursory look at the similarity ratings yielded by the beginners in this study reveals that the results from [sibilant+sonorant] clusters actually disconfirm the hypotheses: SA beginners significantly rated anaptyxis as being more similar to the unepenthesized cluster. This counter-hypothesis pattern is also found for [obstruent+glide] clusters: although insignificant, prothesis was judged more similar than anaptyxis. Similarity ratings provided for the remaining clusters seem to be in line with the predictions. The more marked [obstruent+liquid] cluster was similar to that with anaptyxis, whereas the less marked [sibilant+stop] cluster was similar to that with prothesis. However, results for these two cluster types were also insignificant.

On the other hand, the effect of cluster type on listeners' ability to discriminate stimuli in favor of anaptyxis or prothesis was more robust in advanced listeners, who perceived little distinction between anaptyctic modifications and [obstruent+liquid] clusters. That is, they significantly judged the more marked [obstruent+liquid] clusters as perceptually closer to anaptyctic rather than prothetic modifications (59.66% vs. 40.34%). By contrast, they judged the less marked [sibilant+sonorant] clusters as more similar to prothetic than anaptyctic modifications (55.31% vs. 44.69%). [Obstruent+glide] and [sibilant+stop] clusters demonstrated the same contrast as well, but the differences in listeners' responses did not reach significance.

To sum up, preference for anaptyxis in marked clusters and prothesis in unmarked clusters as predicted by Hypotheses 1 and 2 is only confirmed by the results for the advanced listeners for [obstruent+liquid] and [sibilant+sonorant] clusters as indicated by the discrimination task in this study. For [sibilant+sonorant] clusters, the opposite distribution was found for the beginning listeners who discriminated anaptyctic stimuli less accurately. Note that this finding can only be interpreted as counter-hypothetical if results from beginning as well as advanced listeners are considered; it is meaningful only in relation to other (significant) anaptyxis-prothesis comparisons. It is worth noting that the least marked [sibilant+stop] clusters were associated with greater similarity for prothesis in both subject groups (55% in beginning, 53.75% in advanced). Even though this asymmetry represents only a tendency, since statistical significance did no obtain, perhaps due to the relatively small number of clusters in this category, it deserves reporting that both groups concluded that prothetic modifications sounded more like [sibilant+stop] clusters.

The findings of this experiment and in particular the counter-hypothetical results from beginning learners fail to support claims made by Fleischhacker (2001, 2005) who, on articulatory as well as experimental auditory grounds, argues for the perceptual obtrusiveness of anaptyxis in [sibilant+stop] clusters and prothesis in [obstruent+liquid] clusters. In an auditory experiment, Fleischhacker (2001) elicited similarity judgments from her native English listeners by presenting them with anaptyctic and prothetic modifications of [obstruent+sonorant] and [sibilant+stop] clusters. Her findings support a contrast between anaptyxis and prothesis that is dependent on the cluster type: English listeners reported more similarity for prothesis in [sibilant+stop] clusters but more similarity for anaptyxis in [obstruent+sonorant] clusters (see §4.4.5). Similar findings were reported for English and Swedish subjects by Wingstedt and Shulman (1988).

One point of contention between this study and Fleischhacker's (2001) is that the listeners here are non-native speakers of English, as opposed to the native English speakers who participated in Fleischhacker's. It is not unreasonable to think that linguistically the non-native perceptual system is different from the native one. Our intuitions about non-native perception are highly speculative and, therefore, less conclusive simply because, unlike with production, we have no direct way of knowing how non-native learners perceive L2 sounds and structures. But the fact of the matter is that non-native perception can be as divergent from the native norm as can production. Maybe the non-native listeners in this study were not as sensitive to the epenthetic vowels as were the native listeners in Fleischhacker's, because they process unepenthesized clusters differently. For instance, it is possible that when subjects were presented with the unepenthesized stimuli in this study they actually heard a vowel in all stimuli. Given the findings of Experiment 1 where SA listeners consistently reported hearing vowel epenthesis even when none existed in the input, this is not at all implausible. For example, the test trials [bris-bəris-bris] and [stis-sətis-stis], for example, may have sounded like [bəris-bəris-bəris] and [sətis-sətis-sətis], accordingly. If this was the case, then producing judgments of perceptual similarity on the AXB task may not have been triggered by sensitivity to vowel presence per se, but rather due to chance.

Alternatively, subjects may have been insensitive to the presence of the epenthetic vowel due to the vowel being insufficiently long, and hence barely noticeable, in which case listeners may have indiscriminately heard [bris-bəris-bris] and [stis-sətis-stis] as [bris-bris-bris]and [stis-stis-stis], respectively. The results of Experiment 2 support such hypothesis since generally discrimination near the 48 ms phase (the average duration of the epenthetic stimuli in this experiment) was not that great, especially for the beginning group (42.92% correct discrimination). However, the possibility of a bias toward hearing a vowel throughout the stimuli is more likely, given that initial /bər-/ and /sət-/, but not /br-/ and /st-/, sequences occur in Arabic.

To conclude, the non-native SA learners of English in this study demonstrated some aural sensitivity to the location of vowel epenthesis. A difference between anaptyxis and prothesis in onset clusters was significantly upheld in [sibilant+sonorant] clusters by each subject group. For [obstruent+liquid] clusters, on the other hand, the difference was maintained only by advanced listeners. For other onset clusters subjects showed no significant preference for one type of vowel epenthesis or the other. The fact that the anaptyxis/prothesis difference was more apparent in advanced learners implies a developmental role of L2 proficiency on perceptual ability. Not only is the pattern of similarity judgment of advanced learners in the current study statistically more robust compared to those of beginning learners, it is also more in line with the intuitions of the native English-speakers reported by Fleischhacker (2001). This can be taken as evidence of the greater proximity of the perceptual grammar of advanced non-natives to that of natives.

5.4. Experiment 4: Perceptual Markedness in the Production of Onset Clusters

In addition to exploring the relationship between cluster perception and production, this experiment evaluates how well the proposed perceptual markedness of consonant clusters is reflected in the speech of non-native learners. If the auditory biases demonstrated in the perception of onset clusters by SA listeners in Experiment 1 carry over to production, we should see in the production of onset clusters a pattern of modification that is consistent with the markedness hypothesis stated in (1).

5.4.1. Hypotheses

The hypotheses here make direct reference to the perceptually-based markedness hierarchy in (1):

1. SA learners of English will incur more vowel epenthesis errors in producing more marked clusters.
2. SA learners of English will incur fewer vowel epenthesis errors in producing less marked onset clusters.
3. In the production of onset clusters by SA learners, vowel epenthesis is likely to be anaptyctic in more marked clusters and prothetic in less marked clusters.
4. Overall production error rates will be greater for low proficiency learners than for high proficiency learners.

Hypotheses 1 and 2 are formulated on the basis of the results of Experiment 1, and more particularly on the hierarchy of perceptual markedness expressed in (1), with the assumption that markedness translates into difficulty or increased error rate (Anderson 1987, Broselow and Finer 1991, Davidson et al. 2004, Eckman 1991, Eckman and Iverson 1993). They predict a linear relationship between markedness and difficulty in the production of consonant clusters: the more marked a cluster is, the more likely it undergoes vowel epenthesis. More generally, the hypotheses are motivated by research that has linked L2 speakers' ability to produce non-native clusters natively to a myriad of factors including typology (Eckman 1987), sonority (Broselow and Finer 1991), frequency (Levelt et al. 2000, Kirk and Demuth 2005), or prosodic structure (Broselow 1993, Giegerich 1992).

Hypothesis 3 tests the correlation between cluster markedness and vowel epenthesis site. It predicts, following Experiment 1, that vowel epenthesis errors in cluster production will be asymmetrically distributed among onset clusters with anaptyxis commonly found in [obstruent+glide] clusters, and prothesis in [sibilant+stop] clusters. [Sibilant+sonorant] and [obstruent+liquid] clusters for which no significant differences were found in Experiment 1 are, nonetheless, predicted to follow suit: higher frequency of anaptyxis in [obstruent+liquid] and prothesis in [sibilant+sonorant] clusters. The rationale for this prediction is the association of anaptyxis with marked clusters and prothesis with unmarked clusters in general; in Experiment 1 anaptyxis contributed to the greatest number of vowel epenthesis errors whereas prothesis contributed to the least number of vowel epenthesis errors in the perception of onset clusters. Hypothesis 3 also builds on research suggesting a cluster-based difference in the location of vowel epenthesis realized by non-native speakers (Gouskova 2001, Kwon 2005, 2006, Singh 1985).

Hypothesis 4 tests the developmental effect linguistic proficiency may have on cluster production. Akin to Hypothesis 3 in Experiment 1, it finds basis in studies that have reported L2 improvement in production with higher proficiency in the target language (e.g. Bradlow et al. 1997, Flege 1995).

5.4.2. Methodology

5.4.2.1. Participants

Participants were the same as in previous experiments.

5.4.2.2. Materials and Task

The materials for this experiment included 52 test words plus 30 filler words that served as distracters. The 52 words were representative of the 7 cluster categories and were those of the wordlist used in Experiment 1. The filler words were everyday commonly used one and two-syllable English words chosen at random from Kucera and Francis's (1967) *Computational Analysis of Present-day American English* (see Appendix F for the list of fillers). A wordlist reading task was used to elicit the production of onset clusters by SA learners of English. In general, there does not seem to be a clear association between the type of task used to elicit consonant cluster production data and error rate. At least in the production of consonant clusters, Hanson (2004) and Lin (2001), for example, independently maintain that better controlled wordlist or passage reading production tasks, such as the one used in this experiment, tend to yield higher levels of accuracy than more 'spontaneous' tasks such as conversations. Major (1999), on the other hand, demonstrates that his Brazilian subjects produced onset and coda clusters more accurately in reading passages as opposed to citation forms in wordlists.

5.4.2.3. Procedure

The experiment was set up using Praat's Experiment MFC. The randomized 82 target words embedded in the carrier phrase 'This English word is __' were presented to participants individually.[23] They were instructed to click on a box for each sentence to appear on a computer screen and were asked to read each out loud as their voice was being digitally recorded using Audacity Software (v.1.2.6) and a clip-on PRO 7 Electret condenser microphone. When participants finished reading a sentence, the next one started 1000 ms after a click had been made. An optional 5-minute break was offered after the first 41 items. To make sure every participant understood the instructions, a 3-item practice test was conducted prior to the experiment. The recording took place in a quiet room and lasted an average of 25 minutes.

5.4.2.4. Analysis and Results

The production data for each participant were edited by extracting the 52 target words containing consonant clusters and exporting them onto separate WAV (Waveform Audio File Format) sound files. The edited sound files were then rated independently by two native American English graduate linguistics students whose task was to listen to the data and indicate on a provided sheet of paper whether in the production of onset clusters one of the following criteria was met:

1. Modified by prothesis: if a vowel is inserted before the cluster as in /skul/ → [əskul] 'school'.
2. Modified by anaptyxis: if the cluster was produced with a vowel in the middle as in /klem/ → [kəlem] 'claim'.
3. Modified by deletion: if either C_1 or C_2 of the cluster is deleted as in /drim/ →

[23] Care was taken not to use a carrier phrase whose final segment can potentially form a coda cluster with the ensuing onset. Almost all of the onset clusters examined in this dissertation cannot combine with /z/ and be resyllabified as coda clusters, except for /mj/ for which the coda /-zm/ is possible. Davidson (2003: 48) discusses this possibility of resyllabification as an after-the-fact confound in her dissertation.

[rim] or [dim] ‘dream’, respectively.

4. Modified by phonemic substitution: if either C_1 or C_2 in the cluster is replaced by a different phoneme.
5. Modified in another way, such as by metathesis, a combination of prothesis and anaptyxis, total cluster deletion, etc.

The raters were instructed to record no response if cluster production did not conform to any of the 5 categories stated above. This included, among other things, native-like production of the cluster.

Inter-rater reliability

Inter-rater reliability was computed by comparing the initial rating scores provided by the two judges. A Pearson’s coefficient revealed a significant albeit moderate correlation between the two raters, $r = .87$, $p < 001$, with an overlap of r^2 (.76). Differences between the two raters were settled by employing waveform and spectrogram representations as an additional acoustic analysis measure in determining the presence or absence of the epenthetic vowel and its whereabouts. In particular, presence of clear formant structure F1 and F2 in the spectral representation and periodicity in the waveform were taken as visual evidence for vowel epenthesis (Imai 2004: 58). The adjudicated ratings were submitted again to a Pearson’s coefficient which yielded a stronger correlation, $r = .96$, $p < 001$. In other words, the two ratings overlapped to the extent of r^2 (.92), which is a strong relationship.

The following table sums up error rates for beginning and advanced subjects on the wordlist reading task as judged by the two raters:

Error type	Beginning		Advanced		Total	
	No. errors	%	No. errors	%	No. errors	%
V-epenthesis	173	16.6	125	12	298	14.3
C-deletion	132	12.7	38	3.6	169	8.1
P-substitution	89	8.56	48	4.6	137	6.6
Other	74	7.11	75	7.21	150	7.2
Total	468	45	286	27.5	754	36.25

Table 5.6 Error rates in the production of onset clusters by SA speakers

Out of the 1040 consonant cluster tokens, SA beginning subjects incurred 173 vowel epenthesis errors compared to 125 produced by advanced subjects. Although vowel epenthesis was the most common error type, it still constituted a small percentage of the overall data (only 14.3%). Other repair strategies were even less frequently reported (8.1% deletion and 6.6% phonemic substitution). The largest difference in error rates between the two groups involved *C-deletion* (12.7% vs. 3.6%), and the smallest difference was in the *Other* category (7.11% vs. 7.21%). This can be clearly observed in Figure 5.7:

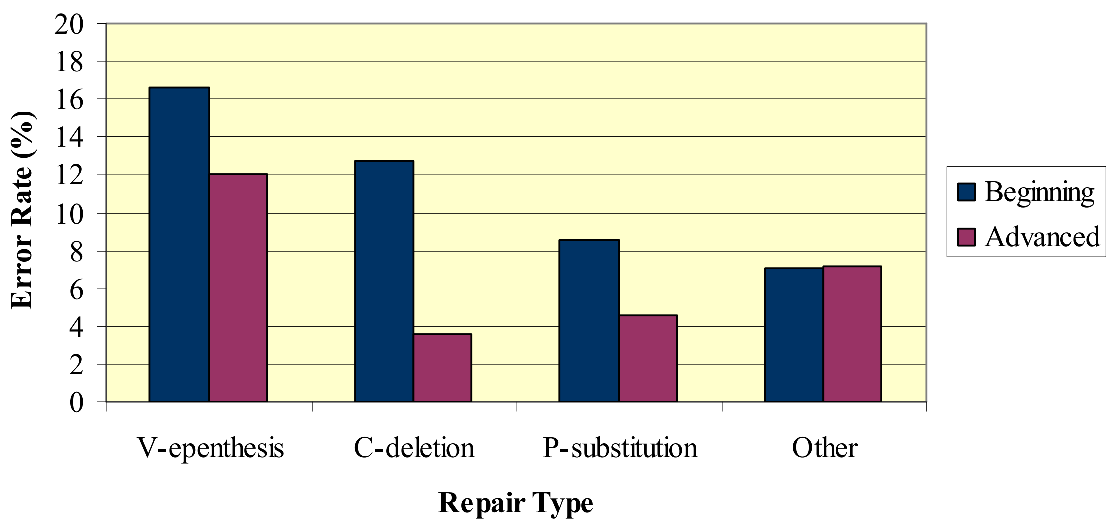


Figure 5.7 Error rates in producing English onset clusters as judged by raters

Proficiency

To find out if proficiency significantly affected subjects' ability to produce onset clusters correctly, total error rates were submitted to an independent sample *t*-test. The results showed a significant difference between beginning and advanced subjects: $t(35.85) = 4.27$, $p < .001$. However, differences involving vowel epenthesis only, as opposed to all error types, did not yield a significant difference between the two subject groups: $t(38) = 1.62$, *ns*. Since Hypotheses 1, 2 and 3 are concerned only with vowel epenthesis errors, data were pooled across the two subject groups and treated aggregately.

Cluster

Vowel epenthesis error rates were then tallied up and averaged for each cluster type. Pairwise tests of proportions were conducted to explore the statistical significance of the cluster variable. Results are as follows:

	V-epenthesis	%	obst+liquid	sibilant+sonorant	sibilant+stop
obst+glide	107	14.86	*ns*	*	*
obst+liquid	153	17.39		*	*
sib+son	24	7.5			*ns*
sib+stop	14	8.75			

* = $p < .05$

Table 5.7 Vowel epenthesis errors in the production of each cluster type

The above table shows numbers of instances of vowel epenthesis judged by the two raters to be present in the speech of the subjects, not numbers of overall errors as this is irrelevant to the hypotheses. The percentages are those of vowel epenthesis relative to the total number of tokens subjects had to produce in each cluster type. For [obstruent+glide] clusters, subjects made 107 vowel epenthesis errors, and this was 14.86% of the 720 tokens of this cluster. For [obstruent+liquid] clusters, 153 vowel epenthesis errors was 17.39% of the 880 total cluster tokens. For [sibilant+sonorant] clusters 24 errors was 7.5% of 320 tokens, and for [sibilant+stop] clusters 14 errors was 8.75% of 160 tokens. This is graphed in Figure 5.8:

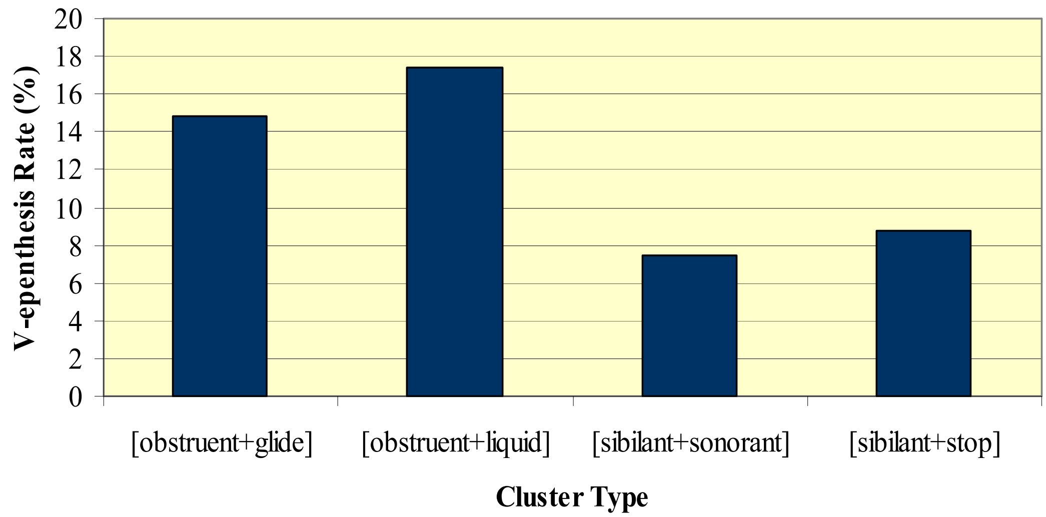


Figure 5.8 Vowel epenthesis error rates by cluster type

Pairwise comparisons show the differences in vowel epenthesis error rates among cluster types to be significant, except for the pairs [obstruent+glide] vs. [obstruent+liquid] and [sibilant+sonorant] vs. [sibilant+stop], for which differences were small.

Cluster by Site

The frequency of anaptyxis to prothesis in the epenthesized production of different onset clusters was examined next. Individual one variable Chi-square analyses that compared the proportions of anaptyxis vs. prothesis were carried out for each cluster type, as summarized in Table 5.8:

	Prothesis %	Anaptyxis %	*sig.*
obstruent+glide	41.83	58.17	*ns*
obstruent+liquid	18.7	81.3	* ($\chi 2 = 39.18$)
sibilant+sonorant	40.4	59.6	*ns*
sibilant+stop	64.28	35.72	* ($\chi 2 = 51$)

* = $p < .05$

Table 5.8 Percentage of anaptyxis and prothesis errors according to cluster type

The results in Table 5.8 demonstrate a significant effect of cluster type on the propensity of SA speakers to employ anaptyxis vs. prothesis when producing [obstruent+liquid] and [sibilant+stop] clusters. While anaptyxis was by far the preferred choice of repair in the production of [obstruent+liquid] clusters (81.3%), prothesis, by contrast, was more prevalent in [sibilant+stop] clusters (64.28%). Such difference in the location of vowel epenthesis was not upheld either in [obstruent+glide] nor [sibilant+sonorant] clusters although anaptyxis error rates in these clusters were greater. Figure 5.9 illustrates this:

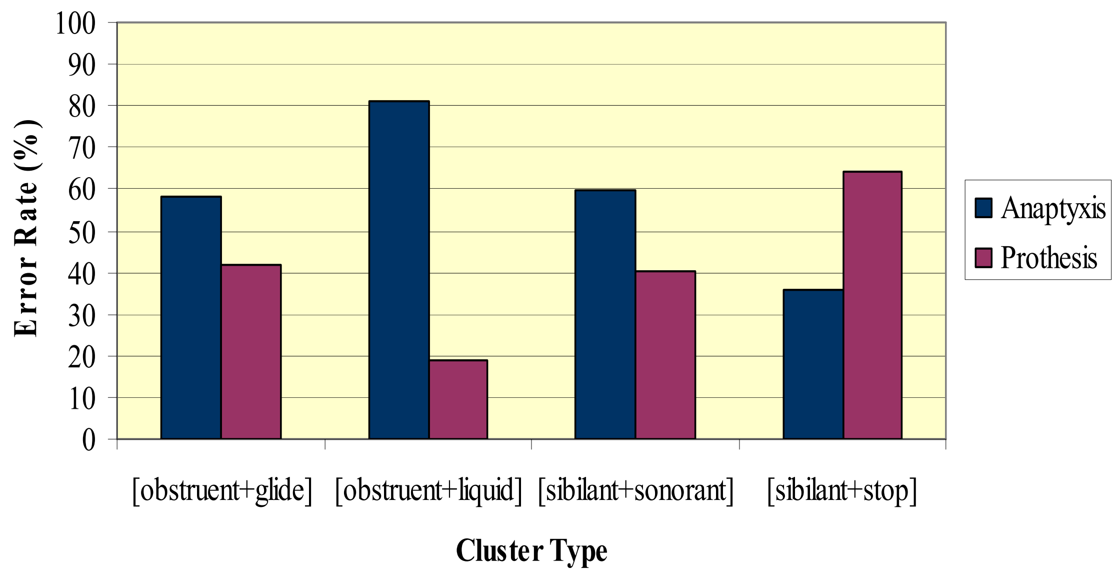


Figure 5.9 Anaptyxis vs. prothesis production errors in onset clusters

As can be seen from Figure 5.9, anaptyctic modifications appear to be the dominant repair strategy utilized by SA speakers in realizing all types of onset clusters, except [sibilant+stop] in which the inverse pattern is observed as mentioned above.

5.4.2.5. Discussion

Looking at the results, Hypotheses 1 and 2 which state that SA learners of English will incur more vowel epenthesis errors in marked vs. unmarked onset clusters seem to be partially supported. There were significantly more errors in the more marked [obstruent+glide] and [obstruent+liquid] clusters than in the relatively less marked [sibilant+sonorant] and [sibilant+stop] clusters. On the other hand, differences in [obstruent+C] and [sibilant+C] clusters were not significant and therefore not supportive of the hypotheses. Importantly, no pattern counter to the hypotheses was observed, however. That is, no greater vowel epenthesis error rates were found for less marked than for more marked cluster types.

These findings are consistent with other studies of L2 consonant cluster acquisition which attribute greater difficulty in the production of [obstruent+glide] and [obstruent+liquid] clusters, compared to [sibilant+sonorant] and [sibilant+stop]. Although studies comparing onset [obstruent+glide] clusters to other types of clusters are rare, perhaps owing to the controversial status of glides in syllable onsets (Davis and Hammond 1995), Eckman and Iverson (1993), who tested Japanese, Korean and Cantonese speakers of English on various English clusters, conclude that [obstruent+glide] clusters were the most problematic type for their subjects. Similar conclusions are found by Fleischhacker (2005) and Zuraw (2007). The relative difficulty of [obstruent+liquid] over [sibilant+sonorant] onsets, on the other hand, was reported in a number of studies. Kim (2000) shows how Korean learners of English produced fewer errors in [sibilant+sonorant] than in [stop+liquid] clusters. Kwon (2006) examined the production of various English onset clusters by 30 Korean native speakers to find out they had more difficulty producing [obstruent+liquid] compared to [sibilant+sonorant] clusters. In Kwon, vowel epenthesis was more commonly found in [stop+liquid] rather than [sibilant+sonorant] clusters, although, as in this study, error rates were not high (2.5% in [sibilant+sonorant] clusters, and 8.2% in [stop+liquid] clusters). In addition, Chan (2006), employing wordlist reading, picture description, passage reading and conversational interviews to elicit production of English onset clusters by 12 native speakers of Cantonese, reports better results for [sibilant+sonorant] clusters compared to [obstruent+liquid] clusters, 95.2% vs. 69.4%.

The likelihood of SA speakers to epenthesize in [obstruent+liquid] but not [sibilant+stop] onsets as exhibited in this study can be corroborated by several other L2 studies. Kim (2000) and Chen (2003) independently maintain that [sibilant+stop] clusters are less complex than [obstruent+liquid] clusters. Kim's Korean learners and Chen's Chinese learners of English both inserted vowels more often when producing [stop+liquid], but not [sibilant+stop] clusters. Chan (2006) demonstrated how Cantonese speakers of English produced [sibilant+stop] clusters in a near native manner (93.4%) whereas [obstruent+liquid] clusters were produced less accurately (69.4%). Similarly, native speakers of Brazilian Portuguese (Major 1996) and Korean learners of English (Kwon 2006) both had more difficulty in producing English [fricative+liquid] onsets relative to [sibilant+stop] onsets.

Contrary to what some consonant cluster studies have found, between [obstruent+C] and [sibilant+C] clusters, SA speakers displayed no significant vowel epenthesis error differences. Zuraw (2007), for example, analyzes Tagalog loanwords with consonant clusters and concludes that [stop+glide] clusters in loanwords are significantly more likely to undergo anaptyxis than [stop+liquid] clusters. The Japanese, Korean and Cantonese speakers of English in Eckman and Iverson 1993 experienced more difficulty in producing [obstruent+glide] than [obstruent+liquid] onsets (although see Broselow and Finer 1991).

Unlike the pattern observed in this study for [sibilant+C] clusters, L2 research on consonant clusters has also reported a difference in non-native speakers' ability to produce [sibilant+sonorant] and [sibilant+stop] clusters natively. Major (1996), for instance, reports [sp-], [st-], and [sk-] onsets less subject to modification by Brazilian Portuguese learners of English than [sl-] onsets. The non-native treatment of [sibilant+stop] onsets compared to [sibilant+sonorant] was also true for Korean speakers as evidenced by Kim (2000) and Kwon (2006), whose non-native English learners were less successful in producing [sibilant+sonorant] relative to [sibilant+stop] onsets (also Chan 2006, Tench 2003).

The fact that most SA speakers overwhelmingly retained the consonants of onset clusters by simplification through vowel epenthesis, as opposed to other means of cluster modification, is consistent with previous reports on consonant cluster production. Davidson (2001), for example, reports that vowel epenthesis was the most frequent error type made by her English speakers in their production of Polish consonant clusters. Hancin-Bhatt and Bhatt's (1997) study of L1 Japanese and Spanish speakers' production of English consonant clusters emphasizes the likelihood of vowel insertion, but not deletion, in onset clusters, while Yoo (2004) concludes that consonant deletion is often associated with coda clusters (see also Anderson 1987, Carlisle 1998, Kwon 2005, Sato 1984).

Vowel epenthesis is also the widely attested method of adaptation in onset cluster loanwords (Silverman 1992). Its pervasiveness can be attributed to a general preservation principle in loanword phonology (Paradis 1996). Brasington (1981) argues that the choice between epenthesis and deletion is universally positionally determined: epenthesis in word-initial and deletion in word-final because the 'initial position is a strengthening position while final position is a weakening one.'

Hypothesis 3 attributed the asymmetry in vowel-epenthesis location to the type of cluster in question, with more marked clusters characterized by anaptyxis and less marked clusters by prothesis. The hypothesis is sustained only by the difference of anaptyxis and prothesis errors in [obstruent+liquid] and [sibilant+stop] clusters. Error rates of anaptyxis were significantly greater in more marked [obstruent+liquid], and fewer in less marked [sibilant+stop]. By contrast, prothesis errors were significantly less in [obstruent+liquid] and more in [sibilant+stop] clusters. However, the hypothesis was neither confirmed nor falsified by the error rate in [obstruent+glide] or [sibilant+sonorant] clusters.

That the SA speakers in this study associated anaptyxis with [obstruent+liquid] and prothesis with [sibilant+stop] clusters is in line with several conclusions made in other studies of the acquisition of onset clusters. Egyptian Arabic speakers of English employ prothesis to repair [sibilant+stop] onsets but are more likely to resort to anaptyxis when dealing with other types of clusters (Broselow 1987, 1992a). Spanish L2 learners of English treat onset clusters differentially, too. Carlisle (1997) found that Spanish speakers used prothesis exclusively on [sibilant+C] clusters. Investigating Hindi speakers' production of English consonant clusters, Singh (1985) noted how rising sonority onset clusters such as [obstruent+liquid] undergo anaptyxis, whereas non-rising sonority clusters such as [sibilant+stop] trigger prothesis (also Gouskova 2001).

Finally, Hypothesis 4, which predicted more overall production errors by beginning subjects was supported. Beginning SA learners of English produced onset clusters non-natively 45% of the time, compared to the much smaller error rate of 27.5% by advanced subjects. This is in line with research showing how general proficiency improves L2 speech as maintained in Bradlow et al. (1997) and Flege (1995). Recall, however, that proficiency did not have a significant effect on vowel epenthesis errors per se, perhaps due to the infrequency of epenthesis errors in both subject groups, only 16.6% and 12% of the consonant clusters produced by beginning and advanced subjects, respectively.

5.5. General Discussion

In general, there was substantial support for the perceptual markedness hypothesis developed in Chapter 4 and reiterated at the outset of this chapter, especially in the results of Experiment 2 and 4. In addition to showing the effect of manipulating the duration of the epenthetic vowel, the results for Hypothesis 2 in Experiment 2 revealed how beginning SA listeners were perceptually oriented consistent with the markedness differences between [sibilant+sonorant] and [obstruent+liquid] clusters and between [sibilant+stop] and the rest of the clusters. Similarly, the advanced listeners differentiated the less marked [sibilant+stop] onset from its epenthesized form more successfully than they differentiated the more marked [obstruent+glide]. Interestingly, this distinction materialized for the relatively short vowel duration of 32 ms. Number of errors by advanced learners in the 48 ms condition followed straightforwardly from the markedness relations among the clusters: the more marked [obstruent+liquid] and [obstruent+glide] onsets showed poor discriminability compared to the less marked [sibilant+sonorant] and [sibilant+stop] onsets. All in all, Experiment 2 yielded findings that were consistent with the markedness hierarchy as well as with the results of Experiment 1.

Experiment 3 was aimed at testing the anaptyxis/prothesis perceptual bias SA listeners had previously shown in Experiment 1. In general, there was mild support for the hypothesis associating location of vowel epenthesis with markedness. For example, preference for anaptyxis in marked clusters and prothesis in unmarked ones emerged only in advanced listeners who judged [obstruent+liquid] as more similar to anaptyctic clusters and [sibilant+sonorant] to prothetic clusters. Weaker evidence for the anaptyxis-prothesis asymmetry was found in the manner by which SA listeners differentiated [sibilant+stop] clusters from their epenthesized tokens: although insignificant, both subject groups experienced difficulty differentiating [sibilant+stop] clusters from their prothetic but not anaptyctic forms. On the whole, results from Experiment 3 showed some consistency with results of the identification task of Experiment 1.

Finally, Experiment 4 measured SA speakers' ability to produce English onset clusters correctly. Its findings lend credence to the markedness hierarchy established in Chapter 4. The more marked [obstruent+glide] and [obstruent+liquid] clusters were consistently produced with more vowel epenthesis errors than were the less marked [sibilant+sonorant] or [sibilant+stop] clusters. The fact that this differential pattern in vowel epenthesis errors also existed in Experiment 1 suggests a link between SA learners' perception and production of English onset clusters. Recall that SA listeners significantly incurred more epenthesis errors in [obstruent+C] onsets than in [sibilant+C] onsets (see §4.2.3.4). However, unlike Experiment 1, Experiment 4 failed to draw a distinction between [obstruent+C] and [sibilant+C] onsets. It seems reasonable to assume, therefore, that SA learners of English were more sensitive to the markedness hierarchy in their perception, rather than production, of onset clusters. Despite this seeming difference between the perceptual and productive grammars of SA speakers, the markedness hierarchy was obeyed in perception and production. That is, although Experiment 4 failed to fully support the results of Experiment 1, there was no clear evidence contrary to the markedness hierarchy.

The differences between Experiments 1 and 4 may not be due after all to lesser sensitivity to cluster markedness in production, but could be task-related. Recall that the task used in Experiment 1 was forced identification while the task used in

Experiment 4 was wordlist reading. In the identification task, listeners were unavoidably more alerted to the purpose of the experiment since they were instructed to focus on the onset cluster of the word and report whether a vowel was present. In other words, it is possible that listeners' attention was directed to epenthesis errors only. In the production task, this was impossible. SA speakers were instructed to read phrases which contained words with onset clusters. The kinds of production errors speakers incurred included not only vowel epenthesis as we saw earlier, but consonant deletion and phonemic substitution, among other things, and the repair strategy speakers used in realizing onset clusters was to a large degree unpredictable.

This procedural difference between the two tasks in the elicitation of onset cluster responses was hardly avoidable and may have in principle contributed to the greater number of epenthesis errors in the perception task compared to the production task. SA learners totalled 783 (36.6%) vowel epenthesis errors in Experiment 1's identification task while the number of vowel epenthesis errors reached only 298 (14.3%) in Experiment 4. It is possible that the statistical insignificance for the [obstruent+C] and [sibilant+C] cluster categories in Experiment 4, as mentioned above, may have been a direct consequence of the relatively small number of vowel epenthesis errors in producing these onset clusters. More care, therefore, has to be taken when comparing SA learners' perception and production of English onset clusters in this study. Nonetheless, it seems reasonable to claim that the SA participants in this dissertation tended to produce onset clusters to a large degree according to their perceptual biases, and vice versa. In other words, they were more likely to hear their own articulatory prejudices when aurally processing English onset clusters.

5.6. Perceptual Account vs. Typological Markedness and Sonority

In this section, I briefly draw a comparison between the perceptual account of consonant cluster markedness in this study and markedness claims determined on the basis of typological and sonority grounds.

5.6.1. Typological Markedness

Recall from §2.3.1 that a typological account bases markedness relations on the frequency and implicational relations of consonant clusters in language. One generalization the typological markedness account makes is that [non-sibilant+liquid] (e.g. [fl-]) and [sibilant+liquid] (e.g. [sl-]) clusters are more marked than [stop+liquid] clusters (e.g. [bl-]), given Eckman and Iverson's (1993: 241) observation that fricatives by themselves are more marked typologically relative to stops, as well as Clements' (1990: 313) SMP in which markedness relations among members of a cluster hold for the entire sequence of consonants (see §2.3.1). A typological account would also render [obstruent+glide] clusters less marked relative to other onset clusters since typologically it is possible for languages that do not allow other onset clusters to have only [obstruent+glide] clusters, as in Japanese, Korean and Chinese.[24]

However, the findings in this dissertation yield generalizations that are different from the ones made by the typological markedness account. According to the perceptual markedness hypothesis, restated here in (2), [sibilant+stop] clusters are considered least marked compared to other clusters, especially [stop+liquid] which are part of [obstruent+liquid] clusters:

[24] Note that this argument is only valid under an account that treats glides as part of the onset, not the nucleus.

(2) Perceptual hierarchy of markedness among onset clusters

less marked ←————→ *more marked*

[sibilant+stop] >> [sibilant+sonorant] >> [obstruent+liquid] >> [obstruent+glide]

Furthermore, the perceptual markedness hierarchy of this dissertation makes the claim that [obstruent+glide] onsets are most marked relative to other clusters. The markedness claims for both [sibilant+stop] and [obstruent+glide] are justified on phonetic grounds (refer to §4.3 for a discussion of the possible phonetic reasons). Another difference between the typological markedness and the perceptual accounts is the more marked status of [non-sibilant+liquid] onsets relative to [stop+liquid], which does not obtain in this dissertation.

Note that even within a typological markedness account, the interpretation of markedness can vary according to whether the distribution of the cluster as a whole is considered, as opposed to the distribution of its members. For example, Morelli (1999) concludes that [fricative+stop], and in particular [sibilant+stop], clusters are least marked among obstruent-onset clusters: the presence of any other obstruent cluster always entails their presence. The following scale summarizes Morelli's findings with regard to obstruent onset clusters:

(3) [fricative+stop] >> [stop+fricative] >> [stop+stop], [fricative+fricative]

The unmarkedness of [fricative+stop] vs. [stop+fricative] and [stop+stop] clusters claimed by Morelli is contrary to the typological observation of Eckman and Iverson (1993: 241) stated above, according to which fricatives are more marked typologically relative to stops. In fact, markedness relations established on the basis of the relative distribution of the individual members of a cluster, and not the whole cluster, would yield, in stark contrast to (3), a scale along the lines of (4):

(4) [stop+stop] >> [stop+fricative], [fricative+stop] >> [fricative+fricative]

Although Morelli's conclusion in (3) is limited to obstruent clusters only, it is worth noting that it is consistent with the perceptual unmarkedness of [sibilant+stop] clusters, evidenced in this dissertation.

5.6.2. Sonority-based Markedness

As discussed in §2.3.2, a sonority-based account such as the MSD of Broselow and Finer (1991) evaluates markedness on the basis of the sonority differences between members of the cluster: the larger the difference in sonority between two members of a cluster, the less marked the cluster would be, and vice versa. The following table sums up the degree of the sonority differences for the set of clusters examined in this dissertation:

Cluster Type	MSD Setting
stop+glide	4
fricative+glide	3
stop+liquid	3
non-sibilant+liquid	2
sibilant+liquid	2
sibilant+nasal	1

Table 5.9 Minimal sonority differences among obstruent-onset clusters

Based on the MSD, the following scale of markedness is derived:[25]

(5) Sonority-based markedness among obstruent-onset clusters

[sibilant+nasal]	*more marked*
[non-sibilant+liquid], [sibilant+liquid]	↑
[fricative+glide], [stop+liquid]	↓
[stop+glide]	*less marked*

The MSD account differs from the perceptual account of this study in a number of ways. First, the MSD makes a distinction between [stop+glide] vs. [fricative+glide],

[25] [Sibilant+stop] clusters are excluded here because they violate the SSP assumption of the MSD.

and [stop+liquid] vs. [fricative+liquid] clusters. Second, the MSD makes no markedness distinction between [fricative+glide] and [stop+liquid] clusters. Aside from the fact that in this set of clusters the MSD setting is the same, it is really not clear why the MSD treats these two clusters indistinctly; that is, there does not seem to be a clear justification, phonological or phonetic, for grouping [fricative+glide] with [stop+liquid]. Third, the MSD considers [stop+glide] and [fricative+glide] clusters less marked than other types of clusters. This is in contrast to the findings here, which report essentially the opposite pattern: [obstruent+glide] clusters (i.e. [fricative+glide] and [stop+glide]) are marked relative to other types of onset clusters. Similarly, [sibilant+nasal] onsets are considered more marked relative to other onset clusters by the MSD. In this study, however, the perceptual hierarchy maintains that [sibilant+nasal], which along with [sibilant+liquid] form the category [sibilant+sonorant], are less marked relative to other onset clusters.

However, it is possible to take issue with the MSD on a number of points (see §2.3.2 for more criticism of the MSD). In addition to over-generating non-occurring clusters in Japanese and Korean (Eckman and Iverson 1993: 239), and failing to account for the English speakers' production of Polish onset clusters with different MSD values in Davidson (2001), the MSD claim for [obstruent+glide] clusters (i.e. [stop+glide] and [fricative+glide]) contradicts the Sonority Dispersion Principle (Clements 1990), which states that sonority should increase sharply and steadily from onset to nucleus. According to Eckman and Iverson's (1993: 248) interpretation of the SDP, [obstruent+glide] onsets are more marked than [obstruent+liquid] onsets because 'the rise in sonority from obstruent to liquid to vowel makes a steadier progression than does the initially sharp, then nearly flat, increase from obstruent to glide to vowel.' In fact, Eckman and Iverson's sonority-based claim that [obstruent+glide] clusters are marked relative to [obstruent+liquid] clusters is consistent with the perceptual markedness hierarchy argued in this dissertation.

To conclude, it is obvious that the typological and sonority accounts yield markedness relations among onset clusters that are quite different from each other as well as from the perceptual markedness hierarchy developed here. However, such differences may be reflective of the various instruments used in defining markedness.

5.7. Summary

In this chapter, three experiments independently have provided empirical testing grounds for the perceptual markedness hierarchy of onset clusters derived from the results of Experiment 1 in Chapter 4. Besides assessing SA listeners' discriminative ability under different durations of anaptyctic vowels, Experiment 2 has established a relationship between onset cluster markedness and length of the vowel in epenthesis. Experiment 3 has explored SA listeners' perceptual biases with regard to the location of vowel epenthesis by assessing similarity judgments of anaptyxis vs. prothesis in various clusters. Overall, there was hardly a discernable pattern although some clusters showed preference for anaptyxis or prothesis, in judgments of the advanced group. Finally, Experiment 4 looked at the manner in which SA learners of English produced various onset clusters. Results were in line with the markedness hypothesis and bore considerable similarity to the findings of Experiment 1. Both experiments consistently have shown that the more marked onset clusters are, the more susceptible to vowel epenthesis errors they become, and vice versa. Neither experiment yielded a pattern that statistically refuted the proposed hierarchy of markedness.

Chapter 6

Formal Analysis

6.1. Introduction

So far the dissertation has focused on the patterns in which SA learners of English perceive and produce vowel epenthesis. In Chapter 5, a series of tests examined the validity of the markedness claim introduced and argued for in Chapter 4 based on the perception of English onset clusters by SA learners of English. In this chapter, the markedness hypothesis proposal is formalized under the framework of Optimality Theory (Prince and Smolensky 1993). The analysis will be limited to the data from Experiment 1, and will show how different individual patterns in the perception of vowel epenthesis can be accounted for in the OT framework.

In §6.2, some background to the theory is given. §6.3 presents an OT analysis of vowel epenthesis in onset clusters as perceived by SA learners of English in this dissertation. §6.4 elaborates on the notion of constraint demotion and its implication for learnability in onset clusters. In §6.5, typological implications of the analysis are discussed. A summary of the main points in §6.6 concludes the chapter.

6.2. Theoretical Background

In this section, I discuss the basic tenets and assumptions of Optimality Theory (Prince and Smolensky 1993). I also show how phonetic knowledge of contrasts in the grammar factors into OT's constraint ranking by reviewing Steriade's (2009) P(erceptibility)-map Theory, a revision of the Correspondence Theory (McCarthy and Prince 1995).

6.2.1. Optimality Theory

OT was first introduced by Prince and Smolensky (1993) as an alternative to more conventional derivational rule-based approaches (e.g. Generative Phonology of Chomsky and Halle 1968, Auto-segmental Theory of Goldsmith 1990). Although originally developed for phonological theory, its application extends to other fields of language such as syntax (e.g. Grimshaw 1997, Keer and Baković 1997). The review of OT here highlights the fundamental principles of the theory and is by no means extensive. It is mainly based on Kager (1999).

In OT, phonological forms are achieved via input-output correspondence relations. The basic idea is that first *(Gen)erator* draws on underlying forms encoded in the *Lexicon* of language to create an unbounded number of outputs for a specific input:

(1) Basic components of OT grammar (Kager 1999: 19)

Lexicon: contains lexical representations (or underlying forms) of morphemes.

Generator: generates output candidates for a given input.

Evaluator: the set of ranked constraints, which evaluates output candidates according to their harmonic values, and selects the optimal candidate.

Next, output forms are assessed through the function of *Eval(uator)* and whichever candidate incurs the least serious violation(s) of a set of violable constraints emerges in the grammar as the *optimal* candidate, the actual output in the language. All candidates are first evaluated on the highest ranked constraint, and any candidate that violates that constraint is eliminated. The same process of evaluation and elimination applies for the rest of the hierarchally ranked constraints until one candidate survives by either satisfying all the constraints or having the lowest ranked violations. A schematic representation of how the assessment and elimination process proceeds in OT is given below:

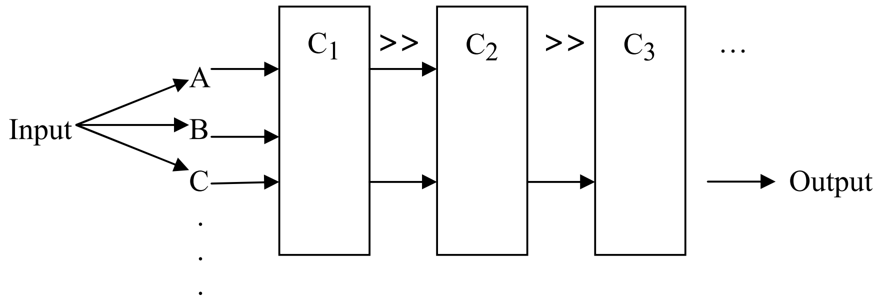


Figure 6.1 Input to output mapping in OT grammar (Kager 1999: 8)

As clear from Figure 6.1, assessment takes place one constraint at a time. Candidate B violates the highest ranked constraint (C_1) and is therefore excluded from further competition; however, candidates A and C make it to the next lower ranked constraint (C_2). Again, both candidates are evaluated at this stage and only candidate C satisfies C_2. Candidate A, on the other hand, incurs a violation of C_2 and is therefore eliminated. At this point, it is clear that candidate C is the winner whether it violates C_3 or not. Hence, C emerges as the actual output attested in the grammar.

There are two types of constraints: markedness and faithfulness. Markedness constraints impose some criterion of structural well-formedness on outputs. They include statements that exert pressure toward unmarked structures over marked ones (Kager 1999: 4, 9). Markedness requirements should, in addition to being typologically motivated, have phonetic grounding.[26] In other words, statements of markedness in the grammar should evolve from (1) cross-linguistic generalizations or universal tendencies, (2) functional motivations that include perceptual, articulatory

[26] It should be noted that not all markedness constraints proposed in OT have been phonetically grounded. There exists a good number of constraints for which no phonetic basis has been established.

and other speech processing factors.[27] An example of a markedness constraint that is both typologically and functionally grounded is NO-CODA which bans closed syllables. Cross-linguistically, the open syllable (CV) is preferred over other types of syllables (Greenberg 1978, Hodne 1985, Sato 1984, Tarone 1987). The sequence of a singleton consonant in the onset, followed by a nucleus vowel is believed to be the unmarked syllable form in language (Jakobson 1962: 526). Perceptually, postvocalic consonants tend to be shorter and unreleased, especially when followed by another consonant, and therefore lack the robust perceptual cues found in prevocalic consonants, which have longer durations in comparison and are often released (Borden et al. 2003, Ohala 1990). In addition, Wright (2004) maintains that CV transitions are more reliable perceptually and provide better consonant cues than VC transitions (see §3.3.2).

Faithfulness constraints, on the other hand, require, as the name suggests, that output candidates be faithful to their lexical input by preserving their input properties (Kager 1999: 10). An example of a faithfulness constraint is DEPENDENCY-IO (DEP-IO) which militates against adding segments in the output that are not present in the input. DEP-IO preserves the integrity of the input by requiring output segments to have input correspondents (Kager 1999: 101). In general, faithfulness constraints assess the output relative to the input. That is, in their application they make direct reference to the correspondence between the input and output forms. This is in contrast to markedness constraints which assess the output only and make no reference whatsoever to the input.

Markedness and faithfulness constraints are intrinsically in conflict with each other: markedness constraints advocate the unmarked by requiring outputs to conform

[27] Yen-Hwei Lin, personal communication, July 5, 2010.

accordingly, whereas diametrically opposing faithfulness constraints strive to retain input-output integrity at the cost of markedness so as to maintain lexical contrasts. This counterbalance of forces in the grammar translates into constraint violability, a major property of OT: satisfying a markedness constraint may come at the expense of a faithfulness constraint, and vice versa. However, violation of constraints must be minimal and justifiable in that it is done in avoidance of violating other higher-ranked constraints (Kager 1999: 12).

Another property of OT is constraint universality. As mentioned above, markedness and faithfulness constraints should reflect cross-linguistically valid generalizations, in addition to having clear functional bases. In other words, markedness and faithfulness constraints should be universal. OT accounts then for different grammars by positing that the hierarchal ranking of the constraints is language-specific. While ranking markedness over faithfulness constraints in one language yields the unmarked form, the inverse ranking in another language ensures, to the detriment of markedness, input faithfulness.

Conflicts between markedness and faithfulness constraints are resolved via domination: the higher ranked member of a pair of conflicting constraints takes precedence over the lower ranked one. A constraint A, for example, is said to dominate a constraint B when A is ranked higher than B in the hierarchy. Violation of a higher ranked constraint is more detrimental than violation of a lower ranked one and it results in the elimination of the candidate violating the higher-ranked constraint. Optimality, therefore, is attained by the output which incurs the least serious violations of a set of constraints (Kager 1999: 13).

For second language acquisition, the input presumably is the attested form in the target language (i.e. the native form) and the interaction between markedness and faithfulness constraints ultimately yields the L2 outcome. Given an input, if the learner's L2 ranking of markedness and faithfulness constraints is parallel to the target language ranking, the optimal output in both will be equal. If, however, the L2 ranking differs from that of the target language, the optimal output will be L2 non-native.

In the acquisition of L2 consonant clusters, for example, modification by vowel epenthesis is achieved in OT terms through the ranking of the markedness constraint *COMPLEX (onsets and codas are simple) and the faithfulness constraint MAX (no deletion) over the faithfulness constraint DEP-V (no vowel epenthesis) (McCarthy and Prince 1995):

(2) OT ranking for vowel epenthesis in onset clusters (vowel epenthesis underlined)

/CCV-/	*COMPLEX	MAX	DEP-V
☞ a. C<u>V</u>CV-			*
b. CCV-	*!		
c. CV-		*!	

Outputs (a-c) are evaluated in accordance with the ranking in tableau (2). The faithful output (b) is eliminated by *COMPLEX, and the deletion output (c) by MAX, as indicated by the notational exclamation mark (!) next to the violation mark (*). Despite the violation of DEP-V, the vowel epenthesis candidate (a) emerges as the optimal output since it incurs the least costly violation of low ranking DEP-V, relative to high-ranking *COMPLEX and MAX. As long as DEP-V is ranked lowest in this particular example, the ranking between *COMPLEX and MAX is irrelevant, as indicated by the dotted line. While strict domination of DEP-V by *COMPLEX and MAX ensures epenthesis in consonant clusters as the L2 optimal output, re-ranking the constraints in (2) yields other possible outputs. For example, ranking DEP-V and MAX above *COMPLEX allows the faithful output (b) to be the winner, while deletion can be derived by low ranking MAX.

The location of the epenthetic vowel in a modified consonant cluster can also be accounted for in OT. It has been argued in the literature that the interaction between the sonority sequencing constraint SYLLABLE CONTACT (Davis 1998, Murray and Vennemann 1983) and other faithfulness and markedness constraints can give rise to the location asymmetry of vowel epenthesis. Gouskova (2001), for example, demonstrates how anaptyxis vs. prothesis asymmetry in Kirgiz loanwords in Russian can be modeled in OT by ranking SYLLABLE CONTACT and faithfulness CONTIGUITY over the markedness constraints ONSET (see also Morelli 1999):

(3) SYLLABLE CONTACT: sonority must not rise across a syllable boundary.

CONTIGUITY: input segments should be contiguous at output.

ONSET: no onsetless syllables.

*COMPLEX: onsets and codas are simple.

Anaptyxis is enforced by high ranking *COMPLEX and SYLLABLE CONTACT over CONTIGUITY in clusters that do not obey SYLLABLE CONTACT. On the other hand, in clusters that satisfy SYLLABLE CONTACT, prothesis is enforced by high-ranking *COMPLEX and CONTIGUITY over ONSET. This is exemplified in tableau (4):

(4) Anaptyxis in /brek/ 'brake' and prothesis in /stek/ 'steak'

	/brek/			/stek/		
	☞ a. bə.rek	b. əb.rek	c. brek	a. sə.tek	☞ b. əs.tek	c. stek
*COMPLEX			*!			*!
SYLL-CONTACT		*!				
CONTIGUITY	*			*!		
ONSET		*			*	

Note that in (4) *Eval* proceeds vertically, as opposed to horizontally in (2). In (4), both faithful outputs that retain the cluster violate *COMPLEX fatally and are, therefore, excluded from further evaluation. In /brek/, the prothetic candidate (b) is in violation of SYLLABLE CONTACT since inserting a vowel at the beginning of the cluster has resulted in the re-syllabification of the onset: /b/ is no longer part of the onset, but is part of the preceding coda. The re-syllabification of /brek/ into [əb.rek] yields a context in which the onset /r/ of the second syllable is more sonorous than the preceding coda /b/. In other words, a rise in sonority across the syllable boundary is the outcome of prothesis in /brek/, thus in violation of SYLLABLE CONTACT. By contrast, the anaptyctic candidate (a) satisfies SYLLABLE CONTACT since the transition from the first to the second syllables in [bə.rek] involves a sonority decline (vowel /ə/ is more sonorous than liquid /r/). At this stage, candidates (b) and (c) have been eliminated from the competition leaving candidate (a) as the winner.

In /stek/, the anaptyctic candidate (a) is penalized and thus ruled out by the constraint CONTIGUITY since the insertion of a vowel in the middle of the cluster disrupts the adjacency of the input segments. No such violation of CONTIGUITY is incurred by affixing the vowel to the beginning of the cluster as in (b). Because of the irrelevance of SYLLABLE CONTACT to prothetic [əs.tek], which, unlike [əb.rek], results in no sonority increase, candidate (b) surfaces as the optimal one, even though it disobeys low ranking ONSET.

6.2.2. P-map Theory

Steriade (2009) notes that the phonological systems of the world's languages show less typological variation than originally predicted by the free interaction of markedness and faithfulness constraints in the OT grammar (Prince and Smolensky 1993). To illustrate this point, she cites the case of final obstruent devoicing as an example. Even though the free ranking of correspondence constraints in OT provides nasalization, devoicing, deletion, metathesis and vowel epenthesis all as possible solutions to a constraint on final voiced obstruents *[+VOICE]/__ , only devoicing is typologically attested (Steriade 2009: 152-154). An input violating the markedness constraint *[+VOICE]/__ , such as /tab/, is often mapped unto [tap], as opposed to nasalization *[tam], deletion *[ta], metathesis *[bat], or vowel epenthesis *[tabə]. This phenomenon, which Steriade calls the 'too many solutions' conundrum, arises when a constraint-based theory such as OT yields multiple possible solutions for a single phonotactic violation. The P-map attempts to eschew this multiplicity in the resolution of such violations.

The P-map is defined as a set of statements about the degree of perceived distinctiveness among different contrasts under various contexts. Two basic tenets of the P-map are the notion of contrast and positional effect. Contrast is defined as 'for *a* to be more perceptible than *b* is for *a* to be more reliably distinguished from a reference term *x* than *b* is distinguished from *x*' (Steriade 2009: 157). Positional effect states that segmental distinctiveness varies according to the syntagmatic context. It describes the fact that contrast between segments may be strengthened or weakened depending on the phonetic context in which they occur (e.g. the strong prevocalic and weak postvocalic voicing contrast in obstruents cited above).

A general guiding principle of the P-map when complying with a certain constraint is the preference for faithfulness: the output that is most similar to the input, among different other outputs, is chosen by the grammar. In the case of the final obstruent voicing constraint stated above, the P-map claims that *only* devoicing renders the contrast [tab]-[tap] most similar, while other repairs such as nasalization [tab]→[tam], deletion [tab]→[ta], metathesis [tab]→[bat] and vowel epenthesis [tab]→[tabə] bring about more perceptual distinctiveness from the input /tab/ and are therefore ultimately avoided by the grammar. The P-map bases this claim of perceptual closeness in [tab]-[tap] on the fact that voicing contrasts are less perceivable after vowels (c.f. prevocalic contrast in [bat]-[pat]), leading to the confusability of the pair [tab]-[tap] (Borden et al. 2003, Ohala 1990, Wright 2004). The P-map makes available this type of detailed phonetic knowledge and phonological similarity to the speakers, enabling them to assess the relative similarity of pairs in various contexts of the language.

One source for the degree of distinctiveness in contrasts that is derivable from the P-map is found in the robust perceptual representations intrinsic to segments. This type of knowledge results from established phonetic experiments that test the perceptibility of consonants and vowels under different phonetic environments. We know, for instance, that vowels are far more perceptible than consonants, and that stops in general are inherently less salient than fricatives perceptually. Within fricatives, sibilants are more audible relative to non-sibilants (Borden et al. 2003, Wright 2004, among many others). Consonants also vary in their perceptibility along the syntagmatic dimension with the prevocalic position being perceptually advantageous over the postvocalic one (see §3.2.2).

Relative similarity among contrasts can also be attained by direct similarity judgments elicited from speakers. Judgments of perceived similarity often manifest the degree with which contrasts are differentiated and can offer a crucial insight into the speaker's P-map. Examples of this include Walden and Montgomery (1975) who report that voicing contrasts (e.g. [p]-[b] as in [pa]-[ba]) are rated more similar than other types of oral vs. nasal or continuant vs. non-continuant contrasts. In one of their experiments, Greenberg and Jenkins (1964) tested listeners' intuition of the relative similarity non-English words have to English words. They provided participants aurally with nonsense words and asked them to write down what they believe to be the closest English cognate word. Responses involving voicing substitutions were more predominant than other possible changes (e.g. nasalization). When participants were presented with the stimulus [klæb], for example, they wrote down [klæp] and to a lesser extent *hand*, a clear associate of *clap*, while other responses such as [klæm] were very infrequent in comparison. More recent examples can be found by Fleischhacker (2001, 2005) and Kwon (2006) who independently rely on perceived auditory similarity judgments collected from subjects in explaining consonant cluster modification patterns.

Steriade (2009) proposes the inclusion of the P-map component into the OT grammar (Prince and Smolensky 1993) and Theory of Correspondence (McCarthy and Prince 1995). She suggests that ranking correspondence constraints in OT results from a perceptually determined hierarchy of distinctiveness that is accessible via the P-map. Therefore, not only do constraints have to be well-grounded in phonetics, their relative ranking should proceed as a function of the perceptibility gradient as well, with the bias to preserve least distinctive contrasts. Thus, if the contrastive pair *x-y* is more distinctive than the pair *x-z* in the context A__, the P-map will rank correspondence constraints higher for the pair *x-y* in order to retain the less distinctive *x-z* pair. Returning to the previously cited voicing prohibition on final obstruents, we can now construct the following perceptual distinctiveness relation between devoicing and nasalization, along with the relevant constraint ranking (Steriade 2009: 153):

A. Perceptibility ranking (P-map)	*More distinctive*	*Less distinctive*
	[tab]-[tam]	[tab]-[tap]
B. Constraint ranking (OT)	*Higher ranked*	*Lower ranked*
	IDENT [±nas]/V__	IDENT [±voice]/V__

Table 6.1 Projection of perceptual distinctiveness into OT ranking

The greater distinctiveness of the nasal contrast in [tab]-[tam] over voicing in [tab]-[tap] gives rise to the ranking of IDENT [±nas]/V__ over IDENT [±voice]/V__ as schematized in Table 6.1. This ranking specifies that a violation of IDENT [±nas]/V__ in the grammar is preclusive while a violation of IDENT [±voice]/V__ is tolerable. The interaction between these two constraints with regard to the markedness constraint banning final obstruent voicing *VOICED-CODA (Kager 1999: 14) is illustrated in the following tableau:

(5) Final obstruent devoicing in /tab/

/tab/	*VOICED-CODA	IDENT [±nas]/V__	IDENT [±voice]/V__
a. tab	*!		
b. tam		*!	
☞c. tap			*

In (5), the faithful output (a) ends in a voiced obstruent and therefore is penalized by *VOICED-CODA. Output (b) is disqualified because of a mismatch in the nasality feature: the output contains a nasal that is not present in the input. The final devoicing output (c) wins since it incurs the least serious violation of faithfulness to the input in voicing.

To sum up, the P-map serves as a repertoire of statements about the perceived distinctiveness of various contrasts under varying contexts. It possesses the necessary machinery that relates the ranking of OT correspondence constraints to speakers' unconscious knowledge of relative perceptibility in any given contrast.

6.3. Perception of Onset Clusters: An OT Analysis

In this section, I present a formal analysis of the vowel epenthesis error patterns SA learners of English demonstrated in their perception of English onset consonant clusters. Recall that SA listeners perceived vowel epenthesis with a varying degree of accuracy, depending on the type of cluster involved. Here, I attempt to elucidate how this differential success can be modeled under an optimality theoretic framework which reflects L2 learners' perceptual grammar. The analysis is limited mainly to the perceptual difficulty hierarchy stated in §4.2.4.2, derived from the results of Experiment 1 and independently motivated by the phonetic argument there. The analysis is situated in the general principles of OT (Prince and Smolensky 1993) as well as Steriade's (2009) P-map.

6.3.1. Transforming Perceptual Markedness into OT Ranking

A point of departure is the perceptual markedness hierarchy for onset clusters, developed in Chapter 4 and repeated here in (6) for convenience sake:

(6) Perceptual hierarchy of markedness among onset clusters

less marked ⟷ *more marked*

[sibilant+stop] >> [sibilant+sonorant] >> [obstruent+liquid] >> [obstruent+glide]

The scale in (6) can be interpreted to reflect the following vowel epenthesis scale of distinctiveness in onset clusters:

(7) Relative perceptual distinctiveness of onset clusters and their epenthesized alternatives

Vowel Epenthesis

more distinctive ⟷ *less distinctive*

[sibilant+stop] >> [sibilant+sonorant] >> [obstruent+liquid] >> [obstruent+glide]

The continuum in (7) indicates, for example, that the contrast between [sibilant+stop] clusters and their epenthetic counterparts is most distinctive, and this distinctiveness lessens in other clusters as we proceed rightward on the scale.

Two assumptions are made. The first is that detailed knowledge of the relative similarity between clusters and their epenthesized forms as illustrated in the distinctiveness scale in (7) is encoded in the P-map. The second is that L2 learners have direct access to the P-map where this type of knowledge is readily available.

Steriade (2001a: 21) maintains that OT input-output 'correspondence constraints are ranked as a function of the relative distinctiveness of the contrasts they refer to'. Thus, we can capture the pattern in which onset clusters were perceived by SA learners by projecting the distinctiveness scale in (7) as a family of context-sensitive constraints DEP-V/A+B:

(8) Ranking among DEP-V/A+B constraints

higher ranked ⟷ *lower ranked*

DEP-ə/sib+stop >> DEP-ə/sib+son >> DEP-ə/obst+liquid >> DEP-ə/obst+glide

The faithfulness constraint DEP-ə/A+B is here considered to be context-sensitive in the sense that it penalizes all schwa insertions, anaptyctic or prothetic, in clusters made up of A and B segments only. It is different from the more general context-free faithfulness constraint DEP-V (McCarthy and Prince 1995), which bans vowel epenthesis irrespective of the cluster segmental identity. The vowel in the family of DEP-V/A+B constraints is specified here as schwa since the quality of the epenthetic vowel in Experiment 1 was a schwa. It is, however, presumed that the constraint should prohibit any vowel, not just schwa.

The hierarchy in (8) ascribes more weight to vowel epenthesis violations in higher ranked clusters and less to epenthesis in lower ranked clusters, thus a violation of top-ranked DEP-ə/sib+stop is most serious and a violation of lower ranking DEP-ə/sib+son is less critical. More tolerable even is failure to comply with DEP-ə/obst+liquid. Finally, unfaithfulness to DEP-ə/obst+glide should be least costly as it is ranked at the bottom of the scale. The ranking in (8) reflects the varying degree of markedness in the perception of onset clusters demonstrated by SA learners of

English in this study, and is therefore believed to be a major guiding force in shaping their perceptual system. Following, I discuss in more detail how the ranking in (8) plays a vital role in constructing the perceptual grammar of SA learners.

6.3.2. Patterns in the Perception of Onset Clusters

Recall that in OT faithfulness and markedness constraints are often in conflict with each other. That being said, the faithfulness constraints in (8) which preserve the cluster by prohibiting vowel epenthesis need to be counterbalanced by the following markedness constraint that requires cluster resolution:

(9)

$*\text{COMPLEX}^{\text{Ons}}$

$*[_{\sigma}$ CC Onsets are simple (Kager 1999: 97)

The constraint $*\text{COMPLEX}^{\text{Ons}}$ bars onset clusters. As defined here it places a restriction on all onsets comprised of two consonants. Grounding for $*\text{COMPLEX}^{\text{Ons}}$ comes from the typological tendency for languages to have simple onsets in the syllable (Cairns and Feinstein 1982, Clements 1990, Clements and Keyser 1983, Greenberg 1978). It is also well-grounded in phonetics. Aside from the articulatory feat in the production of CC- sequences, CV- is perceptually more robust than CC- since release burst, an important cue of consonants, especially plosives, is aided by the following vowel; the transition into a following vowel provides excellent acoustic cues for the identity of a prevocalic consonant. By contrast, the release burst of a consonant in preconsonantal position is impeded by the subsequent consonant, which masks its auditory effects (e.g. Borden et al. 2003, Wright 2004).

The relative ranking of $*\text{COMPLEX}^{\text{Ons}}$ with regard to the DEP-V/A+B faithfulness constraints yields multiple different grammars which can capture SA learners' patterns of consonant cluster perception. For example, ranking *COMPLEX^{Ons} lower than any of the DEP-ə/A+B constraints will yield an L2 grammar in which all clusters are perceived natively, i.e. with no vowel epenthesis. Having *COMPLEX^{Ons} dominate all DEP-ə/A+B constraints, on the other hand, gives rise to perception of vowel epenthesis in all onset clusters regardless of their type. Below is a summary of the possible L2 grammars which arise as a function of the ranking relationship which *COMPLEX^{Ons} holds relative to the family of DEP-ə/A+B constraints:

(10) Possible vowel epenthesis patterns in OT

A. Vowel epenthesis in all clusters

*COMPLEX^{Ons} >> DEP-ə/sib+stop >> DEP-ə/sib+son >> DEP-ə/obst+liquid >> DEP-ə/obst+glide

B. Vowel epenthesis in all clusters except [sibilant+stop]

DEP-ə/sib+stop >> *COMPLEX^{Ons} >> DEP-ə/sib+son >> DEP-ə/obst+liquid >> DEP-ə/obst+glide

C. Vowel epenthesis in [obstruent+liquid] and [obstruent+glide] only

DEP-ə/sib+stop >> DEP-ə/sib+son >> *COMPLEX^{Ons} >> DEP-ə/obst+liquid >> DEP-ə/obst+glide

D. Vowel epenthesis in [obstruent+glide] only

DEP-ə/sib+stop >> DEP-ə/sib+son >> DEP-ə/obst+liquid >> *COMPLEX^{Ons} >> DEP-ə/obst+glide

E. No vowel epenthesis

DEP-ə/sib+stop >> DEP-ə/sib+son >> DEP-ə/obst+liquid >> DEP-ə/obst+glide >> *COMPLEX^{Ons}

The ranking schemata in (10) follow from the position *COMPLEX^{Ons} holds relative to the set of faithfulness constraints. In Grammar A, *COMPLEX^{Ons} dominates all faithfulness constraints. The result is an all-onset cluster prohibition. Grammar A would be equivalent to that of a language which permits no clusters in the onset position. Following proposals for L1 learning (e.g. Demuth 1995, Prince and Tesar 1999), Grammar A characterizes the initial state for L2 learning in which markedness constraints outrank all faithfulness constraints. Grammar A was by far the most common type for the SA learner population tested in this study. It was maintained by 18 out of the 40 individual learners who took part in Experiment 1, who made perceptual errors on all onset clusters when asked to identify vowel epenthesis, although their error rates fluctuated by cluster type.

Grammar B describes a pattern in which vowel epenthesis applies to all except [sibilant+stop] clusters. This is accomplished by ranking *COMPLEX^{Ons} lower than DEP-ə/sib+stop, but higher than the rest of the context-sensitive constraints. This type of grammar was that of 12 participants who managed to correctly identify all instances of vowel epenthesis in [sibilant+stop] clusters.

In Grammar C vowel epenthesis is only perceivable in [obstruent+liquid] and [obstruent+glide] clusters, while perception on the other two cluster types is native-like. This is possible because DEP-ə/obst+liquid and DEP-ə/obst+glide are immediately dominated by *COMPLEX^{Ons}. Grammar C was that of 1 participant, who was able to accurately identify all instances of vowel epenthesis in [sibilant+sonorant] and [sibilant+stop] clusters. However, there were three more participants whose error rates for [obstruent+C] clusters were consistently greater than those for [sibilant+C] clusters.

Grammar D ascribes vowel epenthesis to [obstruent+glide] clusters only. This takes place when *COMPLEX^{Ons} is ranked lower than all other constraints save DEP-ə/obst+glide. While no participant showed vowel epenthesis errors in [obstruent+glide] only and none in other clusters, there were 3 who consistently had greater error rates in [obstruent+glide], compared to other types of clusters.

Finally, Grammar E characterizes the final or end state of the learner's grammar: it exemplifies the native English grammar in which perception in all onset clusters of Experiment 1 is epenthesis-free. It is the opposite of Grammar A, which is concerned with the initial (native L1) state. In Grammar E, *COMPLEX^{Ons} has no significant role as it is dominated by all faithfulness constraints, and therefore, exerts no type of force in restricting onset cluster formation. After studying individual vowel epenthesis patterns, I found no participant whose responses on the identification task evidence this type of grammar. Put differently, native-like perception on all onset clusters was not found in Experiment 1.

The remaining 3 participants exhibited a pattern that failed to adhere to the grammatical repertoire in (10), and hence did not follow straightforwardly from the perceptual distinctiveness scale in (8). In this aberrant pattern, participants incurred more vowel epenthesis errors in [sibilant+stop] onsets, but fewer in [sibilant+sonorant] onsets. Figure 6.2 summarizes the number of subjects for each of the perceptual patterns in (10):

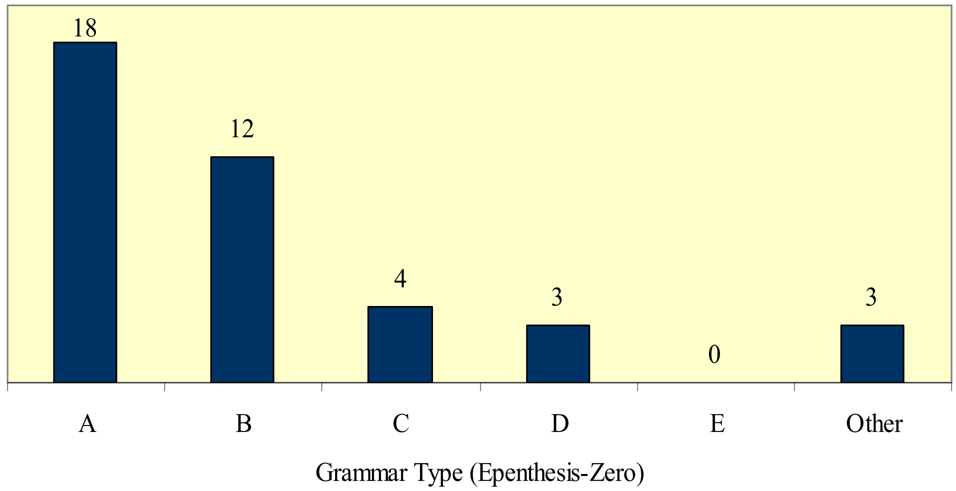


Figure 6.2 SA listeners' zero-epenthesis asymmetries in onset clusters

The following tableaux show how the grammar types A-E derive from the relative ranking between *COMPLEX^{Ons} and the family of context-sensitive DEP-ə/A+B:

(11) Epenthesis and zero-epenthesis in onset clusters

A. Vowel epenthesis in all clusters

	/dwit/		/drit/		/smit/		/stis/	
	dwit	☞dəwit	drit	☞dərit	smit	☞səmit	stis	☞sətis
*COMPLEX^{Ons}	*!		*!		*!		*!	
DEP-ə/sib+stop								*
DEP-ə/sib+son						*		
DEP-ə/obst+liquid				*				
DEP-ə/obst+glide		*						

B. Vowel epenthesis in all clusters except [sibilant+stop]

	/dwit/		/drit/		/smit/		/stis/	
	dwit	☞dəwit	drit	☞dərit	smit	☞səmit	☞stis	sətis
DEP-ə/sib+stop								*!
*COMPLEX^{Ons}	*!		*!		*!		*	
DEP-ə/sib+son						*		
DEP-ə/obst+liquid				*				
DEP-ə/obst+glide		*						

C. Vowel epenthesis in [obstruent+liquid] and [obstruent+glide] only

	/dwit/		/drit/		/smit/		/stis/	
	dwit	☞dəwit	drit	☞dərit	☞smit	səmit	☞stis	sətis
DEP-ə/sib+stop								*!
DEP-ə/sib+son						*!		
*COMPLEX^{Ons}	*!		*!		*		*	
DEP-ə/obst+liquid				*				
DEP-ə/obst+glide		*						

D. Vowel epenthesis in [obstruent+glide] only

	/dwit/		/drit/		/smit/		/stis/	
	dwit	☞dəwit	☞drit	dərit	☞smit	səmit	☞stis	sətis
DEP-ə/sib+stop								*!
DEP-ə/sib+son						*!		
DEP-ə/obst+liquid				*!				
*COMPLEX^{Ons}	*!		*		*		*	
DEP-ə/obst+glide		*						

E. No vowel epenthesis

	/dwit/		/drit/		/smit/		/stis/	
	☞dwit	dəwit	☞drit	dərit	☞smit	səmit	☞stis	sətis
DEP-ə/sib+stop								*!
DEP-ə/sib+son						*!		
DEP-ə/obst+liquid				*!				
DEP-ə/obst+glide		*!						
*COMPLEX^{Ons}	*		*		*		*	

To sum up, the SA individual perceptual patterns in Experiment 1 bore the hallmarks of Grammars (A-D). While the majority fell under Grammar A, Grammar B was also relatively common. Less attested were Grammars C and D.

6.3.3. Anaptyxis-Prothesis Asymmetry

Note how the distinctiveness scale in (8) lacks specification of vowel epenthesis site. While it determines the segmental context in which insertion is to be avoided, it makes no distinction between anaptyctic or prothetic insertions. To specify the contrastive relationship between onset clusters and their epenthetic forms, and to formally account for the difference in vowel location observed in SA listeners' perception of onset clusters, we must incorporate the context in which vowel epenthesis occurs. Thus, I adopt the following anaptyxis-based scale of distinctiveness in onset clusters:[28]

(12) Relative perceptual distinctiveness among onset clusters with regard to anaptyxis

more distinctive ←——→ *less distinctive*

[sib+V+stop] >> [sib+V+son] >> [obst+V+liquid] >> [obst+V+glide]

In (12), the distinctiveness of clusters with anaptyxis becomes less as clusters increase in markedness. The scale here is built on the markedness hierarchy proposed in Chapter 4 (see §4.2.4.2), but is modified to reflect the relative distinctiveness between [sibilant+sonorant] and [obstruent+liquid] clusters and their anaptyctic counterparts.

Recall that based on the results of Experiment 1 alone, no ranking could be established between [sibilant+sonorant] and [obstruent+liquid] onsets with regard to the location of vowel epenthesis. Nonetheless, the fact that anaptyxis was associated with more marked clusters and prothesis with less marked clusters offers some rationale for this modification: the highest rate of vowel epenthesis errors in more marked clusters was contributed by anaptyxis, and the highest rate of vowel epenthesis in less marked clusters by prothesis errors. Partial evidence from subsequent experiments support this division: Experiment 2 found better discrimination on the basis of anaptyxis in [sibilant+sonorant] relative to [obstruent+liquid]; advanced listeners in Experiment 3 ascribed higher similarity to

[28] Note that the scale could equally be constructed to reflect prothesis instead, in which case the distinctiveness values would be reversed.

anaptyxis in [obstruent+liquid], but to prothesis in [sibilant+sonorant]; and finally in Experiment 4 there were more anaptyctic than prothetic insertions in [obstruent+liquid] clusters. Based on the diametric relationship between anaptyxis and prothesis in motivating cluster markedness as well as on evidence from previous experiments, the scale in (12) is proposed, in which greater distinctiveness is assigned to [sibilant+sonorant] than [obstruent+liquid] clusters.

Thus, following similar proposals in the field (e.g. Fleischhacker 2001, 2005, Kwon 2005, 2006), this difference of locations in which vowel epenthesis was perceived in onset clusters can be captured by projecting the distinctiveness scale in (12) onto a family of context-sensitive vowel epenthesis constraints of the type DEP-V/A_B:

(13) Ranking among DEP-V/A_B constraints

higher ranked ⟷ *lower ranked*

DEP-ə/sib_stop >> DEP-ə/sib_son >> DEP-ə/obst_liquid >> DEP-ə/obst_ glide

The faithfulness constraints in (13) differ from those of (8) in that they specify the location of schwa epenthesis: only anaptyctic insertions in clusters made up of A and B are penalized; prothetic insertions go unpenalized. Their interpretation is similar to that in (8) in that anaptyctic violations are more serious in higher ranked clusters.

While the set of constraints in (13) ban anaptyxis, the following structural well-formedness constraint is needed to ensure that $*\text{COMPLEX}^{\text{Ons}}$ is not satisfied by prothesis:

(14)

ONSET

$*[_\sigma$ V   Syllables must have onsets   (Kager 1999: 93)

The markedness constraint ONSET requires each syllable to begin with a consonant. It militates against prothesis just as the family of DEP-ə/A_B constraints in (13) militates against anaptyxis. Similar to *COMPLEX^{Ons}, ONSET is grounded in the typology of syllable structure in which the sequence CV is more common cross-linguistically than VC (Greenberg 1978, Jakobson 1962, see Blevins 1995 for an overview).

The interaction between ONSET and the rest of the faithfulness constraints outlined in (13) accounts for contrastive patterns of anaptyxis and prothesis in a formal way. If ranked higher, ONSET imposes anaptyxis on the dominated faithfulness DEP-ə/A_B constraints. Ranking ONSET below faithfulness DEP-ə/A_B, however, forces prothesis to emerge. The different locations of vowel epenthesis that result from the ranking of ONSET relative to context-sensitive DEP-ə constraints are summarized below:

(15) Anaptyxis/prothesis contrast in onset clusters

A. Anaptyxis in all clusters

*COMPLEX^{Ons} >> ONSET >> DEP-ə/sib_stop >> DEP-ə/sib_son >> DEP-ə/obst_liquid >> DEP-ə/obst_ glide

B. Prothesis in [sibilant+stop] / anaptyxis in all other clusters

*COMPLEX^{Ons} >> DEP-ə/sib_stop >> ONSET >> DEP-ə/sib_son >> DEP-ə/obst_liquid >> DEP-ə/obst_ glide

C. Prothesis in [sibilant+stop] and [sibilant+sonorant] / anaptyxis in [obstruent+liquid]and [obstruent+glide]

*COMPLEX^{Ons} >> DEP-ə/sib_stop >> DEP-ə/sib_son >> ONSET >> DEP-ə/obst_liquid >> DEP-ə/obst_ glide

D. Anaptyxis in [obstruent+glide] / prothesis in all other clusters

*COMPLEX^{Ons} >> DEP-ə/sib_stop >> DEP-ə/sib_son >> DEP-ə/obst_liquid >> ONSET >> DEP-ə/obst_ glide

E. Prothesis in all clusters

*COMPLEX$^{\text{Ons}}$ >> DEP-ə/sib_stop >> DEP-ə/sib_son >> DEP-ə/obst_liquid >> DEP-ə/obst_ glide >> ONSET

In each of the ranking schemata (A-E) in (15), it is important to always have *COMPLEX$^{\text{Ons}}$ ranked at the top of the hierarchy to allow for only epenthetic outputs to surface. Grammars A and E display symmetries of vowel epenthesis location: in Grammar A, ONSET dominates all the constraints forcing anaptyxis in all clusters by eliminating prothesis, while in Grammar E prothesis is the dominant strategy by virtue of low ranking ONSET. Neither grammar was exemplified by subjects, however. No SA participant demonstrated across-the-board anaptyxis (i.e. zero prothesis errors) or across-the-board prothesis (i.e. zero anaptyxis errors). In other words, perception of anaptyxis or prothesis always conformed to one of the patterns B, C, or D. That is not to say that there was no tendency to have more epenthesis errors with one cluster or the other. In fact, two participants did show a tendency for more anaptyxis errors in all clusters somehow conforming to Grammar A, while one participant had more across-the-board prothesis, consistent with Grammar E.

Grammar B stipulates prothesis in [sibilant+stop] clusters only, and anaptyxis in all other clusters by ranking ONSET below DEP-ə/sib_stop and above the rest of the faithfulness constraints. While only 2 participants demonstrated strict adherence to Grammar B, the majority of the SA participants (17 in total) exhibited a tendency toward this pattern. Their perceptual patterns associated prothesis more commonly with [sibilant+stop] clusters, but anaptyxis with [sibilant+sonorant], [obstruent+liquid] and [obstruent+glide] clusters.

Grammars C and D may be upheld by considering the ratio of anaptyxis to prothesis errors in subjects' responses. As for Grammar C no listener, strictly speaking, produced zero anaptyxis errors in [sibilant+C] clusters, and zero prothesis errors in [obstruent+C] clusters. Rather, there was preponderance of prothesis in [sibilant+C] and of anaptyxis in [obstruent+C] for 12 participants. Grammar D which maintains anaptyxis in [obstruent+glide] and prothesis in the remaining types of clusters fit the perceptual patterns of 5 participants only. These individuals methodically incurred more anaptyxis errors in [obstruent+glide] clusters compared to other types of clusters, which were prothesis-dominant instead.

There remains 1 participant whose perceptual pattern of vowel epenthesis location did not fit the repertoire in (15). This participant expressed a strong bias toward anaptyxis in [obstruent+liquid] clusters, while prothesis very much dominated his responses for other types of clusters. Figure 6.3 sums up the number of subjects in this study who conformed to the patterns in (15):

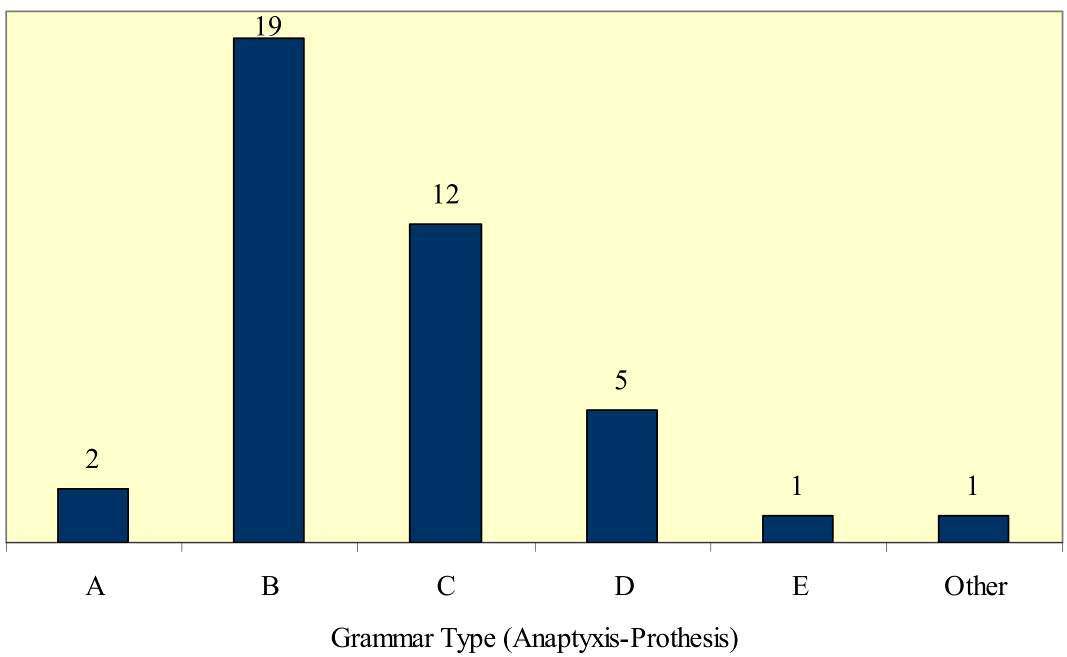


Figure 6.3 SA listeners' anaptyxis and prothesis in onset clusters

The tableaux in (16) represent a schematization of the differences of anaptyxis and prothesis in (15):

(16) Anaptyxis and prothesis in onset clusters

A. Anaptyxis in all clusters

	/dwit/			/drit/			/smit/			/stis/		
	dwit	☞ dəwit	ədwit	drit	☞ dərit	ədrit	smit	☞ səmit	əsmit	stis	☞ sətis	əstis
*COMPLEX^{Ons}	*!			*!			*!			*!		
ONSET			*!			*!			*!			*!
DEP-ə/sib_stop											*	
DEP-ə/sib_son								*				
DEP-ə/obst_liquid					*							
DEP-ə/obst_glide		*										

B. Prothesis in [sibilant+stop] / anaptyxis in all other clusters

	/dwit/			/drit/			/smit/			/stis/		
	dwit	☞ dəwit	ədwit	drit	☞ dərit	ədrit	smit	☞ səmit	əsmit	stis	sətis	☞ əstis
*COMPLEX^{Ons}	*!			*!			*!			*!		
DEP-ə/sib_stop											*!	
ONSET			*!			*!			*!			*
DEP-ə/sib_son								*				
DEP-ə/obst_liquid					*							
DEP-ə/obst_glide		*										

C. Prothesis in [sibilant+stop] and [sibilant+sonorant] / anaptyxis in [obstruent+liquid] and [obstruent+glide]

	/dwit/			/drit/			/smit/			/stis/		
	dwit	☞ dəwit	ədwit	drit	☞ dərit	ədrit	smit	səmit	☞ əsmit	stis	sətis	☞ əstis
*COMPLEX^{Ons}	*!			*!			*!			*!		
DEP-ə/sib_stop											*!	
DEP-ə/sib_son								*!				
ONSET			*!			*!			*			*
DEP-ə/obst_liquid					*							
DEP-ə/obst_glide		*										

D. Anaptyxis in [obstruent+glide] / prothesis in all other clusters

	/dwit/			/drit/			/smit/			/stis/		
	dwit	☞ dəwit	ədwit	drit	dərit	☞ ədrit	smit	səmit	☞ əsmit	stis	sətis	☞ əstis
*COMPLEX^{Ons}	*!			*!			*!			*!		
DEP-ə/sib_stop											*!	
DEP-ə/sib_son								*!				
DEP-ə/obst_liquid					*!							
ONSET			*!			*			*			*
DEP-ə/obst_glide		*										

E. Prothesis in all clusters

	/dwit/			/drit/			/smit/			/stis/		
	dwit	dəwit	☞ ədwit	drit	dərit	☞ ədrit	smit	səmit	☞ əsmit	stis	sətis	☞ əstis
*COMPLEX^{Ons}	*!			*!			*!			*!		
DEP-ə/sib_stop											*!	
DEP-ə/sib_son								*!				
DEP-ə/obst_liquid					*!							
DEP-ə/obst_glide		*!										
ONSET			*			*			*			*

To sum up, the analysis states that, other things being equal, the determining factor in restructuring initial consonant clusters via anaptyxis or prothesis is the relative perceptibility scale in (13). If the learner's perceptual grammar prioritizes a constraint restricting vowel-initial syllables (e.g. ONSET), across-the-board anaptyxis would emerge. In contrast, if the learner's grammar entirely obeys a constraint that preserves the contiguity of the cluster segments, across-the-board prothesis would arise. However, when neither constraint exerts its power, the perceptual account predicts that anaptyxis modifies more marked clusters, while prothesis modifies comparatively less marked ones.

6.4. Learnability and Constraint Demotion

The previous section discussed how SA listeners' different perceptual patterns of vowel epenthesis and its location in onset clusters can be accounted for in OT by varying the ranking of markedness *COMPLEX^{Ons} and ONSET with a hierarchically-fixed set of context-sensitive faithfulness constraints. In this section, I briefly discuss, based on Kager (1999), the precise mechanism which allows for these variant patterns to emerge and point out implications it may have for second language development and learnability. For the sake of brevity, I limit the discussion to the patterns in (10).

Recall the assumption made earlier that the L2-initial state for the learners in this study is their native language grammar; the SA learners' native-language syllable restrictions which sanction no clusters in the onset are transferred over to English. Following the learning theory proposed by Tesar and Smolensky (1998) and its application to L2 acquisition (Broselow 2004, Davidson 1997), after Broselow et al. (1998) I suggest that the SA learners in this study start off with their L1 (Arabic) constraint ranking in which the markedness constraint *COMPLEX^{Ons} is undominated at first, but later gets re-ranked to a lower position on the scale in order to permit for more L2 native-like renditions of the cluster. This process of *Constraint Demotion*, by which the constraint demoted to a lower rank along the hierarchy loses its ability to exert any pressure on other constraints, is driven by the learner's deductions of the right ranking based on positive L2 input.

For the example at hand, the learner has access to the hierarchy of fixed-ranked DEP-V/A+B constraints spelled out in (8). This ranking schema is supplied through the P-map and is available to all SA learners of English. Since their native language bans onset clusters, SA listeners come into the L2 English learning process with their L1 perceptual biases that mandate simple onsets. As such, *COMPLEX^{Ons} is undominated as in Grammar A in (10). As the learner recurrently encounters positive

L2 input via either corrective feedback or the learner's own observation, he or she will be able to recognize the target form and consequently deduce the ranking requisite for the optimal output, given that the target form and the optimal output are, by definition, one.

Let us describe how this proceeds for the 12 participants whose grammar fits Type B in (10). These learners start off with the ranking schema of Grammar A, which mandates, in agreement with the L1 native language, vowel epenthesis in all clusters:

(17) Across-the-board vowel epenthesis (Grammar A-initial state)

*COMPLEX^{Ons} >> DEP-ə/sib+stop >> DEP-ə/sib+son >> DEP-ə/obst+liquid >> DEP ə/obst+glide

As they perceive more L2 positive evidence suggesting the non-existence of vowel epenthesis in [sibilant+stop] clusters, they conclude that [sətis], for example, is no longer the surface form; instead [stis] is. On this basis the learner is now in a position to deduce the appropriate ranking schema necessary to derive the optimal non-epenthetic output, and accordingly demotes the cluster-banning constraint *COMPLEX^{Ons} one stratum down the hierarchy, banning vowel epenthesis only in [sibilant+stop] clusters as in Grammar B:

(18) Vowel epenthesis in all clusters except [sibilant+stop] (Grammar B-medial state)

DEP-ə/sib+stop >> *COMPLEX^{Ons} >> DEP-ə/sib+son >> DEP-ə/obst+liquid >> DEP-ə/obst+glide

The new ranking in (18) is now one step closer to the L2 native one in which *COMPLEX^{Ons} occupies the lowest position. Similarly, when learners become cognizant of the actual non-epenthetic surface form for [sibilant+sonorant] clusters, they demote the constraint *COMPLEX^{Ons} even one stratum lower (Grammar C).

More demotions of *COMPLEX^{Ons} continue to take place until *COMPLEX^{Ons} has been relegated to the lowest stratum on the hierarchy where no further demotions are possible, and the learner has arrived at the target language ranking (Grammar E). While in approximating the target language these different re-ranking schemata or sub-grammars are active in the learners' grammar at one point or another, they are completely suppressed in the final state due to epenthesis being banned in all clusters.

The various perceptual patterns attested by the SA learners in this study constitute the different developmental stages learners go through in their perception of English onset clusters. Constraint demotion provides the necessary tool to account for individual subject variation as well as L2 perceptual development.

6.5. Typological Implications

We saw how the re-ordering of *COMPLEX^{Ons} and ONSET in previous sections generated multiple possible rankings, the majority of which fit the SA learners in this study. Here, I discuss the typological implications of the analysis by reviewing the patterns of vowel epenthesis in second language and loanword phonology of onset clusters surveyed in Fleischhacker (2001).

One logical conclusion that follows from the analysis in §6.3.2 is that for many L2 learners vowel epenthesis is the default in marked clusters, unless abolished by a faithfulness constraint, whereas absence of vowel epenthesis is the default in unmarked clusters, unless enforced by a structural well-formedness constraint. The perceptibility schema in (8) yields patterns that are compatible with this generalization: all-cluster vowel epenthesis (Grammar A), all-cluster zero vowel epenthesis (Grammar E), and epenthesis in marked/no epenthesis in unmarked clusters (Grammars B-D). An implication of this is that a pattern of no vowel epenthesis in marked and epenthesis in unmarked clusters should not exist and even if it does should be very uncommon. Nonetheless, this implication is weakened by the fact that 3 participants in Experiment 1 exhibited a counter-hypothetical pattern: they made more vowel epenthesis errors on the less marked [sibilant+stop], than on the more marked [sibilant+sonorant] clusters. Typological evidence also indicates the existence of a pattern that is incompatible with the implication above. Fleischhacker (2001: 5) reports a pattern in Haitian Creole, Catalan and Spanish loanwords in which prothesis applies only to [sibilant+stop] clusters, while [obstruent+liquid] clusters are produced with no vowel epenthesis.

Concerning the site of vowel epenthesis, the typological implications of the analysis are better upheld. The analysis in §6.3.3 makes the general prediction that, ceteris paribus, anaptyxis should occur in more marked clusters, while less marked clusters should attract prothesis. Across-the-board anaptyxis or prothesis obtain when imposed by some structural constraint (i.e. a prothesis blocking or anaptyxis blocking constraint, respectively). This implies that when less marked clusters undergo anaptyxis, more marked clusters must also undergo anaptyxis, and when prothesis applies to more marked clusters, it must apply to less marked clusters as well. This means that a pattern of anaptyxis in less marked clusters and prothesis in more marked clusters should in theory be non-existent. However, this claim was contradicted in this study by one participant only who had more anaptyxis errors in the less marked [obstruent+liquid], and more prothesis errors in the more marked [obstruent+glide] onsets.

Fleischhacker (2001: 2-10) reports that numerous patterns of vowel epenthesis asymmetries in different onset clusters are attested in one language or another. She reports that in realizing English onset clusters, across-the-board anaptyxis is observed in Korean, Japanese and Punjabi, while across-the-board prothesis is observed in Iraqi

Arabic and Central Siberian Yupik. In addition to these two symmetrical patterns, there were differences of vowel epenthesis location as follows: anaptyxis in [obstruent+sonorant] but prothesis in [sibilant+sonorant] and [sibilant+stop] onsets (e.g. Farsi), anaptyxis in [obstruent+sonorant] and [sibilant+liquid] but prothesis in [sibilant+nasal] and [sibilant+stop] onsets (e.g. Kazakh), and finally anaptyxis in [obstruent+sonorant] and [sibilant+sonorant] but prothesis in [sibilant+stop] onsets (e.g. Egyptian Arabic). Other languages that display the general asymmetrical pattern of anaptyxis in [obstruent+sonorant] clusters and prothesis in [sibilant+stop] clusters include Central Pahari, Hindi, Sinhalese and Bengali. These typological findings are summarized in Table 6.2 (Fleischhacker 2001: 33):[29]

	Epenthesis pattern	*Language(s)*
A	**Anaptyxis** in all clusters	Korean, Japanese and Punjabi
B	**Prothesis** in all clusters	Iraqi Arabic, Central Siberian Yupik
C	**Anaptyxis** in [obst+son] / **Prothesis** in [sib+son] and [sib+stop]	Farsi
D	**Anaptyxis** in [obst+son] and [sib+liquid] / **Prothesis** in [sib+nasal] and [sib+stop]	Kazakh
E	**Anaptyxis** in [obst+son] and [sib+son] / **Prothesis** in [sib+stop]	Egyptian Arabic
F	**Prothesis** in [obst+son] / **Anaptyxis** in [sib+stop]	NOT ATTESTED
G	**No** vowel epenthesis	English

Table 6.2 Typology of vowel epenthesis patterns in onset clusters

Note that the absence of F, the only unattested pattern in Table 6.2, confirms the claim stated earlier about the unlikelihood of a pattern which associates prothesis with more marked clusters, and anaptyxis with less marked clusters. Also unattested, based on Fleischhacker's (2001) typological study, is the insertion of prothetic vowels in [sibilant+sonorant] but anaptyctic vowels in [sibilant+stop] clusters. The fact that

[29] This is a simplified summary. The reader is referred to Fleischhacker (2001) for a comprehensive and more in-depth discussion.

these two patterns do not appear in any known language is crucial. Not only does it support the perceptually-based analysis of anaptyxis/prothesis asymmetries in this dissertation, but it further explains why such aberrant patterns were extremely rare in SA learners of English.

6.6. Summary

This chapter has formalized the perceptual markedness hierarchy developed for the perception of onset clusters by SA learners of English. The hierarchy of difficulty was transformed into relative perceptibility or distinctiveness scales – derivable from the L2 learner's P-map knowledge – and projected onto a family of context-sensitive faithfulness OT constraints. It was argued that the interaction of markedness *COMPLEX^{Ons} and ONSET on the one hand and faithfulness DEP-ə/A+B and DEP-ə/A_B on the other provided a principled account of the different perceptual grammars of SA learners. With few exceptions reported, learners' perception of vowel epenthesis and its location proceeded as a function of demoting the markedness constraints *COMPLEX^{Ons} and ONSET relative to the strata of vowel-blocking faithfulness constraints. In addition, it was shown how the notion of constraint demotion could offer some insight into L2 learners' ability to approach the native L2 grammar. The chapter ended with a discussion of the implications the analysis bears for the typology of vowel epenthesis patterns in onset consonant clusters.

CHAPTER 7

CONCLUSION

7.1. Evaluating the Research Questions

This dissertation has looked at non-native ability to perceive L1-illegal syllable-initial consonant clusters. Saudi Arabic has an active constraint on the occurrence of syllable-initial clusters, yet the performance of its speakers did not entirely conform to the native language phonology, nor did it fully approximate the target language norms. Instead, SA speakers in this research perceived and produced English onset clusters with a varying degree of accuracy, often alternating between the native and non-native form of the cluster. Further, their ability to perceive and produce clusters natively varied according to cluster type. While SA learners were able to perceive certain onset clusters accurately by correctly identifying vowel epenthesis errors, their identification was less successful with other types. When producing onset clusters, they incurred more vowel epenthesis errors in some clusters, but not others. Location of the epenthesized vowel seemed to be cluster-dependent as well. SA learners perceived and produced anaptyxis in some clusters, but prothesis in others.

In what follows, I summarize the main objectives of the dissertation by reiterating the research questions set forth in Chapter 1, and briefly discussing how well each was addressed by the results.

The research questions in this dissertation were as follows:

1. How will SA learners of English, whose first language is averse to CC-sequences, perceive English onset consonant clusters? Will their perception be native-like, non-native-like, or variably both?

The first question concerns cluster variation. The findings indicate that SA listeners' perception of English onset clusters ranged from native to non-native in all clusters. In Experiment 1, listeners had trouble identifying vowel epenthesis 49.86% of the time in [obstruent+glide], 36.6% in [obstruent+liquid], 25.94% in [sibilant+sonorant] and only 11.87% in [sibilant+stop] clusters. This means that listeners' ability to perceive onset clusters within each type varied from native-like with vowel epenthesis being correctly identified to non-native-like with incorrect vowel identification.

These findings reflect a crucial role of the native language phonology in L2 perception. SA listeners' insensitivity to the presence or absence of vowel epenthesis in onset clusters is very likely the result of a structural ban on initial clusters in their native language (see §4.2.4.1). Similar findings are yielded by other studies, which demonstrate a strong influence of the native-language phonotactic system and syllabic constraints on the L2 perceptual process (Chang et al. 2007, Dehaene-Lambertz et al. 2000, Dupoux et al. 1999, Dupoux et al. 2001, Kabak and Idsardi 2007, Moreton 2002, Polivanov 1931).

2. Will SA listeners' perception of vowel epenthesis vary from one cluster to another? If so, can this relative perceptibility of vowel epenthesis be systematically attributed to cluster type?

The second research question investigates variation across clusters, which was found to be present in SA learners in this study. In Experiment 1, listeners' inability to identify vowel epenthesis varied by cluster type. Vowel epenthesis was hardest to identify in [obstruent+glide] clusters, with a high error rate of almost 50% (49.86%). Next were [obstruent+liquid] clusters, in which listeners identified vowel epenthesis only 36.6% of the time. Difficulty subsided in [sibilant+sonorant] clusters (25.94%). The lowest error rate belonged, however, to [sibilant+stop] clusters, in which SA listeners were able to correctly identify instances of vowel epenthesis almost 90% of the time. Only in 11.87% of the stimuli did they have trouble indicating whether vowel epenthesis was present or not. This variation across different types of clusters was significant statistically, suggesting that perceptibility of vowel epenthesis is cluster-determined, as shown in the following scale of perceptual difficulty:

(1) Perceptual hierarchy of difficulty among onset clusters

less difficult ⟷ *more difficult*

[sibilant+stop] >> [sibilant+sonorant] >> [obstruent+liquid] >> [obstruent+glide]

Taking vowel perceptibility to correspond to perceptual difficulty or markedness (see §4.2.4.2), sensitivity to vowel epenthesis increased as perceptual difficulty in onset clusters decreased.

Although shown to have existed in production (e.g. Broselow and Finer 1991, Davidson 2006, Kwon 2005, Major 1996), such variation in learners' ability to perceive non-native onset clusters represents a new finding in the second language acquisition field (although see Berent et al. 2007). It implies that L2 perception, parallel to L2 production, is subject to a host of regulatory factors that trigger the differential success observed by the L2 learners. Here, chief among these factors is the role of phonetics along the dimensions of segmental quality, voicing, gestural overlap, and C_1 to C_2 transitional properties in explaining the acoustic prominence of vowel epenthesis under different consonantal contexts (§4.3). The conclusion that SA speakers of English are sensitive to markedness relations built on phonetic considerations agrees with the Berent et al. study in which the perception of L1-illicit onset clusters by English speakers was modulated by the sonority profile of the cluster.

3. How sensitive are SA listeners to the presence/absence of the inserted vowel in onset clusters? Does the degree of sensitivity vary according to vowel length, cluster type, or both?

The third question is addressed by Experiment 2, which measures aural sensitivity to vowel epenthesis by manipulating vowel length. The results of Experiment 2 reveal a significant impact of vowel duration on listeners' ability to perceive anaptyxis in various onset clusters. SA listeners were able to systematically better discriminate stimuli the longer the duration of the epenthetic vowel. In the shortest 16 ms condition, discrimination suffered as both beginning and advanced listeners were unable to detect the epenthesized vowel. In the longest 80 ms duration, discriminability was 87.71% in beginning and 96.66% in advanced learners. Thus, consistent with the Japanese and French subjects of Dupoux et al. (1999), for example, SA listeners' sensitivity to vowel epenthesis was directly related to the duration of the inserted vowel.

Research question 3 goes further to raise the possibility of vowel length being reacted to differently depending on the cluster type. Although partial evidence was found for this claim, the results of Experiment 2 indeed suggest that listeners' alertness to subtle variations in epenthetic vowel length was more acute in some clusters than others. For example, under the 32 ms, 48 ms and 64 ms conditions, beginning listeners were able to detect vowel epenthesis fairly well in [sibilant+stop] onsets compared to other types of clusters. These listeners also heard 64 ms durations better in [sibilant+sonorant] clusters than [obstruent+liquid] clusters. Advanced listeners also demonstrated greater sensitivity toward epenthesis in [sibilant+stop] clusters even when vowel duration was only 32 ms. When duration increased to 48 ms, listeners became attentive to vowel epenthesis in [sibilant+C] onsets only, but showed auditory indifference to its presence in [obstruent+C] onsets. By the time duration reached 64 ms, vowel epenthesis was clearly audible in all clusters. To sum up, the auditory sensitivity of SA listeners varied along the dimensions of vowel length as well as cluster type.

4. Will SA listeners demonstrate a bias toward anaptyxis vs. prothesis, and vice versa, in their perception of vowel epenthesis location? If so, will this asymmetry in their performance be cluster-dependent?

The fourth research question pertains to the location of vowel epenthesis in a cluster. Experiment 3 probes SA listeners for cluster-based differences between anaptyxis and prothesis. Despite research establishing cluster-dependent differences in vowel location (e.g. Fleischhacker 2001, 2005), different treatment of anaptyxis vs. prothesis by SA listeners in this dissertation obtained only in few cases, generally lacking strong support. While auditory bias toward anaptyxis was displayed by beginner listeners in [sibilant+sonorant] clusters only, advanced listeners expressed bias toward anaptyxis in [obstruent+glide] clusters and prothesis in [sibilant+sonorant] clusters. Overall, in Experiment 3 there was mild tendency in support of the hypotheses. Unlike Fleischhacker, who demonstrated a clear cluster-dependant place asymmetry in the perception of vowel epenthesis, such failed to be strongly supported here, possibly because Fleischhacker's subjects were native speakers of English while the subjects here were non-native listeners.

5. Will perception of onset clusters be better in higher proficiency learners of English, compared to lower proficiency learners?

The fifth research question deals with the effect of linguistic proficiency on L2 perception. The findings of this dissertation are mixed in this regard. While Experiment 1 found no significant differences due to proficiency, Experiments 2 and

4 both resulted in better performance by advanced learners. In Experiment 2, the ability to discriminate stimuli correctly along a continuum of epenthetic vowel duration was better observed in advanced learners. In Experiment 4, advanced SA speakers had significantly fewer overall errors than beginning speakers (27.5% vs. 45% error rate), although differences in vowel epenthesis error rates were insignificant between the two groups.

6. How closely do L2 learners' perception and production of consonant clusters correspond?

The sixth question addresses the relationship between perception and production. Similar findings between Experiments 1 and 4 point to a perception-production link, at least in the acquisition of onset clusters. Broadly speaking, SA learners consistently experienced more difficulty in both experiments with [obstruent+glide] and [obstruent+liquid] than with [sibilant+sonorant] and [sibilant+stop] clusters; vowel identifiability was worse in [obstruent+C] than [sibilant+C] clusters; and production of [obstruent+C] clusters, in turn, was fraught with vowel epenthesis errors, compared to [sibilant+C] clusters.

The fact that the same vowel epenthesis pattern surfaced in both experiments despite task-related differences in eliciting responses suggests a relationship, albeit a modest one, between SA learners' perception and production of English onset clusters. Such conclusion is corroborated by research that maintains a relationship between L2 perception and production in general (see §3.7).

7. Can variable degrees of difficulty in the perception of English onset clusters by SA listeners be accounted for in formal phonological theory?

The last research question asks whether the varying success SA listeners show in their perception of onset clusters can be analyzed formally. Chapter 6 maintains that based on the theoretical framework of Optimality Theory (Prince and Smolensky 1993) and the principles of the P-map (Steriade 2009), it is possible to describe SA learners' perceptual grammars in a formal way by projecting distinctiveness scales onto a family of context-sensitive faithfulness OT constraints. In particular, it was argued that SA learners' zero-epenthesis patterns in the perception of onset clusters are captured by the ordering of the markedness constraint *COMPLEX^{Ons} relative to a ranked set of the faithfulness constraint DEP-V/A+B. With regard to anaptyxis-prothesis asymmetries, the ranking of ONSET relative to DEP-V/A_B generates multiple grammars that include the individual SA learner patterns in this dissertation.

7.2. Limitations and Further Research

This dissertation has mainly examined the perceptual success SA learners of English achieve in acquiring different English onset consonant clusters. One contribution this dissertation makes in the field of second language acquisition is that it re-emphasizes the role of perception in L2 speech processing and phonological theory. The dissertation also dismisses L1 transfer as the sole factor for the relative difficulty in the perception of vowel epenthesis in onset clusters; instead, other factors that are intrinsic to the cluster itself such as the phonetic properties of the segments are involved. The processing of vowel epenthesis is subject to a hierarchy of phonetic properties which helps explain why vowel epenthesis is more successfully identified in some clusters but not others.

The current study is limited in scope to evaluating the impact vowel epenthesis has on L2 listeners' ability to perceive clusters accurately. Further research should examine the effects of other repair strategies such as consonant deletion or phonemic substitution on learners' perception of onset, medial, or coda clusters. In addition, the research here investigated L2 perception of onset clusters only in SA learners of

English. It would be interesting to see if the findings provided by the learners in this dissertation extrapolate to other subject populations with different native language backgrounds. Findings where similarities in relative perceptibility of cluster-dependant vowel epenthesis indeed exist across different subject groups would suggest universality of the proposed scale of difficulty.

Other possible causes for the variable difficulty which L2 learners exhibit in their perception of consonant clusters also merit attention. This dissertation has only considered, besides the native language, the role of phonetics (vowel prominence) and linguistic proficiency in L2 perceptual ability. Future research can examine other phonetically related factors in L2 cluster perception such as speech rate and amount of gestural overlap in the native input.

APPENDICES

APPENDIX A

Oral questionnaire answered by participants

1. Name and age?
2. What part of Saudi Arabia do you come from? In what city-or cities-have you lived?
3. Is Arabic your native language?
4. Are your parents both native speakers of Arabic?
5. What Arabic dialect within Saudi Arabia is primarily spoken by you and your parents?
6. What level of education have you completed so far?
7. When did you first start learning English, and for how long?
8. Were you exposed to English at an early age (*e.g. had an English speaking caretaker; attended a private school where English is the primary language of instruction*)?
9. Have you lived in an English speaking country before coming to the US? If so, for how long?
10. How long have you been staying in the US?
11. Have you ever taken any specialized pronunciation courses or received any training on speaking skills?
12. What level of the English program are you placed in?/What year of college are you in?
13. How would you rate your general English ability (*excellent-very good-good-poor*)?
14. Do you speak other languages besides Arabic and English?
15. Do you have or suffer from any hearing or speaking difficulties?

APPENDIX B

List of English onset cluster words used in the experiments

[stop+liquid]

cluster	***attested***	***test word (IPA)***
br-	breach/broom	bris, brus
bl-	bleak/bloom	blis, blus
tr-	treat/troop	tris, trus
dr-	dream/drool	drit, drut
kr-	creep/croon	krit, krut
kl-	clean/clued	klit, klut
gl-	gleam/gloom	glis, glus
gr-	green/groove	gris, grus

[non-sibilant+liquid]

cluster	***attested***	***test word (IPA)***
fr-	freeze/fruit	fris, frus
fl-	fleet/fluke	flis, flus
θr-	threes/through	thris, thrus

[sibilant+stop]

cluster	***attested***	***test word (IPA)***
st-	steer/stool	stis, stus
sk-	skeet/school	skis, skus

[sibilant+nasal]

cluster	***attested***	***test word (IPA)***
sn-	sneeze/snooze	snit, snut
sm-	smear/smooch	smit, smut

[sibilant+liquid]

cluster	***attested***	***test word (IPA)***
sl-	sleeve/sloop	slis, slus
ʃr-	shriek/ shrewd	ʃrit, ʃrut

[stop+glide]

cluster	***attested***	***test word (IPA)***
bj-	beauty	bjut, bjut
kj-	cute	kjus, kjus
mj-	muse/mute	mjas, mjus
tw-	tweak/twang	twis, twas
dw-	dweeb/dwarf	dwit, dwat
kw-	queen/quart	kwit, kwat

[fricative+glide]

cluster	***attested***	***test word (IPA)***
hj-	huge/Hughes	hjut, hjut
fj-	fuse/fume	fjut, fjut
sw-	sweet/swoop	swis, swus

APPENDIX C

Wordlist produced by the native speaker of English (actual list randomized)

CCVC	CVCVC	VCCVC	CCVC	CVCVC	VCCVC
bris	bəris	əbris	brus	bərus	əbrus
gris	gəris	əgris	grus	gərus	əgrus
tris	təris	ətris	trus	tərus	ətrus
drit	dərit	ədrit	drut	dərut	ədrut
krit	kərit	əkrit	krut	kərut	əkrut
blis	bəlis	əblis	blus	bəlus	əblus
klit	kəlit	əklit	klut	kəlut	əklut
glis	gəlis	əglis	glus	gəlus	əglus
fris	fəris	əfris	frus	fərus	əfrus
flis	fəlis	əflis	flus	fəlus	əflus
θris	θəris	əθris	θrus	θərus	əθrus
stis	sətis	əstis	stus	sətus	əstus
skis	səkis	əskis	skus	səkus	əskus
snit	sənit	əsnit	snut	sənut	əsnut
smit	səmit	əsmit	smut	səmut	əsmut
slis	səlis	əslis	slus	səlus	əslus
ʃrit	ʃərit	əʃrit	ʃrut	ʃərut	əʃrut
twis	təwis	ətwis	twas	təwas	ətwas
dwit	dəwit	ədwit	dwat	dəwat	ədwat
kwit	kəwit	əkwit	kwat	kəwat	əkwat
swis	səwis	əswis	swus	səwus	əswus
mjas	məjas	əmjas	mjus	məjus	əmjus
bjut	bəjut	əbjut	bjut	bəjut	əbjut
kjus	kəjus	əkjus	kjus	kəjus	əkjus
hjut	həjut	əhjut	hjut	həjut	əhjut
fjut	fəjut	əfjut	fjut	fəjut	əfjut

APPENDIX D

Number of responses for the 156 test words in Experiment 1

Stimulus		*Response*	
	Prothesis (A)	*Anaptyxis (B)*	*None (C)*
bris	9	11	20
əbris	36	0	4
bəris	0	27	13
brus	12	8	20
əbrus	25	0	15
bərus	1	33	6
blis	11	3	26
əblis	33	2	5
bəlis	0	32	8
blus	6	9	25
əblus	29	1	10
bəlus	0	31	9
tris	8	10	22
ətris	30	1	9
təris	0	32	8
trus	12	11	17
ətrus	28	4	8
tərus	0	32	8
drit	10	7	23
ədrit	28	0	12
dərit	1	28	11
drut	10	13	17
ədrut	24	1	15
dərut	0	32	8
krit	16	21	3
əkrit	27	1	12
kərit	0	19	21
krut	14	7	19
əkrut	24	1	15
kərut	1	37	2
klit	9	17	14
əklit	34	3	3
kəlit	0	26	14
klut	13	7	20
əklut	29	1	10
kəlut	2	27	11
glis	11	29	0
əglis	27	1	12
gəlis	0	23	17
glus	14	15	11
əglis	28	1	11

gəlus	0	29	11
gris	13	10	17
əgris	27	0	13
gəris	1	31	8
grus	17	11	12
əgrus	28	0	12
gərus	0	27	13
fris	11	9	20
əfris	31	0	9
fəris	0	28	12
frus	8	14	18
əfrus	29	1	10
fərus	0	26	14
flis	2	18	20
əflis	30	2	8
fəlis	0	20	20
flus	6	12	22
əflus	33	1	6
fəlus	0	33	7
thris	14	12	14
əthris	25	0	15
thəris	1	26	13
thrus	10	14	16
əthrus	31	3	6
thərus	0	32	8
stis	6	0	34
əstis	33	1	6
sətis	3	34	3
stus	3	1	36
əstus	34	0	6
sətus	0	40	0
skis	4	2	34
əskis	39	0	1
səkis	0	38	2
skus	7	6	27
əskus	33	1	6
səkus	0	36	4
snit	7	9	24
əsnit	35	0	5
sənit	0	32	8
snut	10	5	25
əsnut	30	1	9
sənut	0	36	4
smit	15	4	21
əsmit	32	0	8
səmit	1	38	1
smut	14	9	17
əsmut	30	4	6

səmut	1	27	12
slis	4	12	24
əslis	36	2	2
səlis	0	32	8
slus	6	17	17
əslus	32	0	8
səlus	1	23	16
shrit	0	6	34
əshrit	29	3	8
shərit	1	37	2
shrut	6	14	20
əshrut	32	2	6
shərut	0	32	8
bjut	10	23	7
əbjut	29	3	8
bəjut	0	15	25
bjut	13	21	6
əbjut	30	1	9
bəjut	2	14	24
kjus	10	17	13
əkjus	32	0	8
kəjus	0	21	19
kjus	9	21	10
əkjus	29	2	9
kəjus	0	26	14
mjas	12	28	0
əmjas	27	2	11
məjas	0	15	25
mjus	9	30	1
əmjus	27	3	10
məjus	0	8	32
twis	4	15	21
ətwis	34	0	6
təwis	1	20	19
twas	8	25	7
ətwas	29	2	9
təwas	0	19	21
dwit	5	28	7
ədwit	36	0	4
dəwit	0	11	29
dwat	8	21	11
ədwat	34	1	5
dəwat	0	26	14
kwit	3	29	8
əkwit	31	0	9
kəwit	1	11	28
kwat	7	29	4
əkwat	27	2	11

kəwat	0	21	19
hjut	5	18	17
əhjut	35	2	3
həjut	0	23	17
hjut	7	23	10
əhjut	35	0	5
həjut	1	27	12
fjut	7	33	0
əfjut	32	4	4
fəjut	0	12	28
fjut	4	27	9
əfjut	39	0	1
fəjut	0	10	30
swis	10	21	9
əswis	33	1	6
səwis	0	23	17
swus	8	17	15
əswus	36	0	4
səwus	3	20	17

APPENDIX E

Criteria for the selection of onset clusters in Experiment 2

Selected		**Possible**	**Rationale**
[kl-]	vs.	[kr-]	/r/ C_2
[kl-]	vs.	[bl-] or [gl-]	Voiced C_1
[fl-]	vs.	[fr-]	/r/ C_2
[fl-]	vs.	[θr-]	Frequency (293 vs. 78).
[sl-]	vs.	[ʃr-]	/r/ C_2 and frequency (179 vs. 31)
[sn-]	vs.	[sm-]	Frequency (110 vs. 74)
[kw-]	vs.	[tw-] or [dw-]	Frequency (227 vs. 53 and 9). Also C_1 is consistent with [kl-]
[sw-]	vs.	[hj-] or [fj-]	Frequency (127 vs. 41 and 56). C_1 is consistent with [sk-], [sn-], and [sl-]. C_2 is consistent with [kw-]

*Frequency counts based on (Kjellmer 1998: 86).
*The justification for singling out these clusters among other onset clusters was a bit complicated as these clusters differ on several dimensions. Generally, the clusters were selected based on the desire to keep frequency and the voicing status of the cluster as much consistent as possible, as well as on preference for /l/ over /r/ in C_2. Voiceless, as opposed to voiced, clusters were preferred because (1) C_2 is always voiceless in some cluster types (e.g. [sibilant+stop]), (2) a voiced cluster could potentially form a confound in detecting vowel epenthesis: in a fully voiced environment, it might be hard to detect the epenthetic vowel compared to a voiceless environment. C_2 /r/ is avoided because it is realized quite differently in both languages: /r/ is realized as the retroflex approximant [ɹ] by many American speakers of English (Ladefoged 2001), but it is an alveolar (or sometimes dental) trill [r] for the majority of Arabic speakers (Al-Ani 1970). /l/, on the other hand, is realized as an alveolar lateral in both English and Arabic.

APPENDIX F

List of fillers used in the production experiment[1]

mind	drop
lord	voice
land	risk
force	faith
sweet	court
trust	rock
seek	thick
grass	heart
brother	count
night	ground
fair	length
fact	deep
house	cattle
dinner	single
machine	deal

[1] The list is taken from the 'Most Frequent 2200 English Words' retrieved on January, 9, 2009 from: http://www.auburn.edu/~nunnath/engl6240/kucera67.html, which is based on the Kucera and Francis's (1967) Computational Analysis of Present-day American English.

REFERENCES

REFERENCES

Abrahamsson, Niclas. 1999. Vowel epenthesis of /sC(C)/ onsets in Spanish/Swedish Interphonology: A longitudinal case study. *Language Learning*, 49, 473-508.

Abrahamsson, Niclas. 2003. Development and recoverability of L2 codas: A longitudinal study of Chinese-Swedish interphonology. *Studies in Second Language Acquisition*, 25, 313-49.

Abramson, A.S. and Leigh Lisker. 1970. Discriminability along the voicing continuum: Cross-language tests. *Proceedings of the Sixth International Congress of Phonetic Sciences*, 569-573. Academia, Prague.

Al-Ani, Salman. 1970. *Arabic phonology: An acoustical and physiological investigation*. The Hague, Netherlands: Mouton.

Ali, Azra, Mohamed Lahrouchi, and Michael Ingleby. 2008. Vowel epenthesis, acoustics and phonology patterns in Moroccan Arabic. *Interspeech*, 1178-1181.

Al-Shahrani, Saeid. 1988. *The phonology of the spoken Arabic of the Shahran tribe in southern Saudi Arabia*. Ph.D dissertation, University of Michigan, Ann Arbor.

Altenberg, Evelyn. 2005. The judgment, perception, and production of consonant clusters in a second language. *International Review of Applied Linguistics in Language Teaching*, 43, 53-80.

Anderson, Janet. 1983. Syllable simplification in the speech of second language learners. *The Interlanguage Studies Bulletin*, 3, 4-36.

Anderson, Janet. 1987. The markedness differential hypothesis and syllable structur difficulty. *Interlanguage phonology: The acquisition of a second language sound system*, ed. by Georgette Ioup and Steven Weinberger, 279-91. Cambridge, MA: Newbury House.

Archibald, John. 1998. *Second language phonology*. Philadelphia: John Benjamins.

Archibald, John. 2003. Learning to parse second language consonant clusters. *The Canadian Journal of Linguistics*, 48, 149-77.

Audacity Team. 2008. Audacity. [Computer program]. Retrieved January 12, 2008, from: http://audacity.sourceforge.net/

Bakalla, Muhammad. 1979. *The morphological and phonological components of the Arabic verb: Meccan Arabic*. London: Longman Group.

Barlow, Jessica. 1997. *A constraint-based account of syllable onsets: Evidence from developing systems*. Ph.D dissertation, Indiana University.

Barlow, Jessica. 2001. The structure of /s/-sequences: Evidence from a disordered system. *Journal of Child Language*, 28, 291-324.

Barry, William. 1989. Perception and production of English vowels by German learners: instrumental-phonetic support in language teaching. *Phonetica*, 46, 155-168.

Beddor, Patrice and Terry Gottfried. 1995. Methodological issues in cross-language speech perception research with adults. *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*, ed. by Winifred Strange, 207-32. Timonium, MD: York Press Inc.

Benson, Bronwen. 1988. Universal preference for the open syllables as an independent process in interlanguage phonology. *Language Learning*, 38, 221-35.

Berent, Iris, Donca Steriade, Tracy Lennertz, and Vered Vaknin. 2007. What we know about what we have never heard: evidence from perceptual illusions. *Cognition*, 104, 591-630.

Best, C. T. 1993. Emergence of language-specific constraints in perception of nonnative speech: A window on early phonological development. *Developmental Neurocognition: Speech and Face Processing in the First Year*, ed. by B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, and J. Morton. Kluwer Academic, Dordrecht, the Netherlands.

Best, C. T. 1994. The emergence of native-language phonological influences in infants: a perceptual assimilation model. *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, ed. by H. C. Nusbaum. MIT, Cambridge, MA.

Best, Catherine. 1995. A direct realist view on cross-language speech perception. *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*, ed. by Winifred Strange, 171-204. Timonium, MD: York Press Inc.

Best, Catherine, Gerald McRoberts, and Nomathemba Sithole. 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, 345-360.

Best, Catherine, Gerald McRoberts, and Elizabeth Goodell. 2001. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, 109, 775-794.

Best, Catherine and Winifred Strange. 1992. Effects of phonological and phonetic factors on cross language perception on approximants. *Journal of Phonetics*, 20, 305-330.

Birdsong, David. 1992. Ultimate attainment in second language acquisition. *Language*, 68, 706-755.

Blevins, Juliette. 1995. The syllable in phonological theory. *A handbook of phonological theory*, ed. by John Goldsmith, 206-44. Oxford: Blackwell.

Boersma, Paul and David Weenink. 2005. Praat: doing phonetics by computer. [Computer program]. Retrieved September 12, 2008, from: http://www.praat.org/

Bongaerts, Theo. 1999. Ultimate attainment in foreign language pronunciation: The case of very advanced late foreign language learners. *Second language acquisition and the Critical Period Hypothesis*, ed. by David Birdsong, 133-159. Mahwah, NJ: Erlbaum.

Borden, Gloria, Adele Gerber, and Gary Milsark. 1983. Production and perception of the /r/-/l/ contrast in Korean adults learning English. *Language Learning*, 33, 499-526.

Borden, Gloria, Katherine Harris, and Lawrence Raphael. 2003. *Speech science primer: Physiology, acoustics and perception of speech*. Lippincott Williams and Wilkins.

Bradlow, Ann, David Pisoni, Reiko Akahane-Yamada, and Yoh'ichi Tohkura. 1997. Training Japanese listeners to identify English /r/ and /l/ IV: Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, 101, 2299-2310.

Brasington, Ron. 1981. Epenthesis and deletion in loan phonology. Work in progress 3. *Reading University Phonetics laboratory*, 97-103.

Brière, Eugene. 1966. An investigation of phonological interferences. *Language*, 42, 768-95.

Broselow, Ellen. 1984. An investigation of transfer in second language phonology. *International Review of Applied Linguistics*, 22, 253-269.

Broselow, Ellen. 1987. Non-obvious transfer: on predicting epenthesis errors. *Interlanguage phonology: The acquisition of a second language sound system*, ed. by Georgette Ioup and Steven Weinberger, 292-304. Cambridge, MA: Newbury House.

Broselow, Ellen.1988. Prosodic phonology and the acquisition of a second language. *Linguistic Theory in Second Language Acquisition*, ed. by Suzanne Flynn and Wayne O'Neil, 295-308. Dordrecht: Kluwer Academic Publishers.

Broselow, Ellen. 1992a. *The structure of fricative-stop onsets*. Ms., State University of New York at Stony Brook.

Broselow, Ellen. 1993. Transfer and universals in second language epenthesis. *Language Transfer in language learning*, ed. by Susan Gass and Larry Selinker, 71-86. Amsterdam: John Benjamins.

Broselow, Ellen. 2004. Unmarked structures and emergent rankings in second language phonology. *International Journal of Bilingualism*, 8, 51-65.

Broselow, Ellen. Su-I Chen, and Chilin Wang. 1998. The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition*, 20, 261-280.

Broselow, Ellen and Daniel Finer. 1991. Parameter setting in second language phonology and syntax. *Second Language Research*, 7, 35-59.

Browman, Catherine and Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology*, 6, 201-51.

Brown, Cynthia. 1997. *Acquisition of segmental structure: consequences for speech perception and second language acquisition*. Ph.D dissertation, McGill University, Montreal, Quebec.

Brown, Roger and Donald Hildum. 1956. Expectancy and the perception of syllables. *Language*, 32, 411-419.

Byrd, Dani. 1996. Influences on articulatory timing in consonant sequences. *Journal of Phonetics*, 24, 209-44.

Cairns, Charles and Mark Feinstein. 1982. Markedness and the theory of syllable structure. *Linguistic Inquiry*, 13, 19-225.

Carlisle, Robert. 1997. The modification of onsets in a markedness relationship: Testing the Interlanguage Structural Conformity Hypothesis. *Language Learning*, 47, 327-361.

Carlisle, Robert. 1998. The acquisition of onsets in a markedness relationship: A longitudinal study. *Studies in Second Language Acquisition*, 20, 245-60.

Chan, Alice. 2006. Strategies used by Cantonese speakers in pronouncing English initial consonant clusters: Insights into the interlanguage phonology of Cantonese ESL learners in Hong Kong. *International Review of Applied Linguistics in Language Teaching*, 44, 4, 331-355.

Chang, Yueh-chin, Jiaqing Hong, and Pierre Hallé. 2007. English cluster perception by Taiwanese Mandarin speakers. *The Proceedings of the International Conference for Phonetic Sciences*, 16, 797-800.

Chen, Szu-wei. 2003. Acquisition of English onset clusters by Chinese learners in Taiwan. *Postgraduate Conference 2003*, University of Edinburgh.

Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Clements, George. 1990. The role of the sonority cycle in core syllabification. *Papers in laboratory phonology I*, ed. by John Kingston and Marry Beckman, 283-333. Cambridge: Cambridge University Press.

Clements, George and Samuel Keyser. 1983. *CV phonology: A generative theory of the syllable*. Cambridge, MA: MIT Press.

Coté, Marie-Hélène. 2000. *Consonant cluster phonotactics: A perceptual approach*. Ph.D dissertation, MIT.

Davidson, Lisa. 1997. *An Optimality Theoretic approach to second language acquisition*. Senior Honors Thesis, Brown University.

Davidson, Lisa. 2001. Hidden rankings in the final state of the English grammar. *RuLing Papers II*, ed. by Graham Horwood and Se-Kyung Kim, 21-48. New Brunswick: Rutgers University.

Davidson, Lisa. 2003. *The atoms of phonological representation: Gestures, coordination, and perceptual features in consonant cluster phonotactics*. Ph.D dissertation, Johns Hopkins University.

Davidson, Lisa. 2006. Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics*, 34, 104-37.

Davidson, Lisa. 2007. The relationship between the perception of non-native phonotactics and loanword adaptation. *Phonology*, 24, 261-286.

Davidson, Lisa, Peter Jusczyk, and Paul Smolensky. 2004. The initial and final states: Theoretical implications and experimental explorations of richness of the base. *Constraints in phonological acquisition*, ed. by René Kager, Wim Zonneveld, and Joe Pater, 791-839. Cambridge: Cambridge University Press.

Davis, Stuart. 1990. An argument for the underspecification of [coronal] in English. *Linguistic Inquiry*, 21, 301-06.

Davis, Stuart. 1998. Syllable Contact in Optimality Theory. *Journal of Korean Linguistics*, 23, 181-211.

Davis, Stuart and Michael Hammond. 1995. On the status of on-glides in American English. *Phonology*, 12, 159-182.

Dehaene-Lambertz, Ghislaine, Emmanuel Dupoux, and Ariel Gout. 2000. Electrophysiological correlates of phonological processing: A cross-linguistic study. *Journal of Cognitive Neuroscience*, 12, 635-647.

DeKeyser, Robert. 2000. The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22, 499-534.

Demuth, Katherine. 1995. Markedness and the development of prosodic structure. *Proceedings of the North East Linguistic Society 25*, ed. by Jill Beckman, 13-25. Amherst: Graduate Linguistic Student Association.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christopher Pallier, and Jacques Mehler. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1568-1578.

Dupoux, Emmanuel, Christopher Pallier, Kazuhiko Kakehi, and Jacques Mehler. 2001. New evidence for pre-lexical phonological processing in word recognition. *Language and Cognitive Processes*, 16, 491-505.

Dupoux, Emmanuel, Christopher Pallier, Nuria Sebastian, and Jacques Mehler. 1997. A distressing 'deafness' in French? *Journal of Memory and Language*, 36, 406-421.

Duyck, Wouter, Kevin Diependaele, Denis Drieghe, and Marc Brysbaert. 2004. The size of the cross-lingual masked phonological effect does not depend on second language proficiency. *Experimental Psychology*, 51, 1-9.

Eckman, Fred. 1987. Markedness and the contrastive analysis hypothesis. *Interlanguage phonology: The acquisition of a second language sound system*, ed. by Georgette Ioup and Steven Weinberger, 55-69. Cambridge, MA: Newbury House.

Eckman, Fred. 1991. The structural conformity hypothesis and the acquisition of consonant clusters in the interlanguage of ESL learners. *Studies in Second Language Acquisition*, 13, 23-41.

Eckman, Fred. 2004. Universals, innateness and explanation in second language acquisition. *Studies in Language*, 28, 682-703.

Eckman, Fred and Gregory Iverson. 1993. Sonority and markedness among onset clusters in the interlanguage of ESL learners. *Second Language Research*, 9, 234-52.

Eckman, Fred, Edith Moravcsik, and Jessica Wirth. 1989. Implicational universals and interrogative structures in the interlanguage of ESL learners. *Language Learning*, 39, 173-205.

Fikkert, Paula. 1994. *On the acquisition of prosodic structure*. Ph.D dissertation, Holland Academic Graphics.

Flege, James. 1987. The production of new and similar phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, 47-65.

Flege, James. 1995. Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*, ed. by Winifred Strange, 233-277. Timonium, MD: York Press Inc.

Flege, James and Wieke Eefting. 1987a. Production and perception of English stops by Spanish speakers. *Journal of Phonetics*, 15, 67-83.

Flege, James and Kathryn Fletcher. 1992. Talker and listener effects of perceived foreign accent. *Journal of Acoustical Society of America*, 19, 370-389.

Flege, James and Serena Liu. 2001. The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, 23, 527-552.

Flege, James, Murray Munro, and Ian MacKay. 1995. Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97, 3125-3134.

Fleischhacker, Heidi. 2001. Cluster-dependent epenthesis asymmetries. *UCLA working papers in linguistics*, 7, papers in phonology 5.

Fleischhacker, Heidi. 2005. *Similarity in phonology: Evidence from reduplication and loan adaptation*. Ph.D dissertation, UCLA.

Flemming, Edward. 1995. *Auditory representations in phonology*. Ph.D. dissertation, UCLA.

Fowler, Carol. 1989. Real objects of speech perception: A commentary on Diehl and Kluender. *Ecological Psychology*, 1, 145-160.

Freitas, João and Maria Rodrigues. 2003. On the nature of sC-clusters in European Portuguese. *Journal of Portuguese Linguistics*, 2, 55-85.

Frisch, Stefan, Nathan Large, and David Pisoni. 2000. Perception of wordlikeness: Effects of segment probability and length on the processing of non-words. *Journal of Memory and Language*, 42, 481-496.

Gafos, Adamantios. 2002. A grammar of gestural coordination. *Natural Language and Linguistic Theory*, 20, 269-337.

Gass, Susan and Larry Selinker. 2001. *Second language acquisition: An introductory course.* London: Lawrence Earlbaum.

Giegerich, Heinz. 1992. *English Phonology: An introduction*. Cambridge: Cambridge University Press.

Gierut, Judith. 1999. Syllable onsets: Clusters and adjuncts in acquisition. *Journal of Speech, Language, and Hearing Research*, 42, 708-26.

Goad, Heather and Yvan Rose. 2004. Input elaboration, head faithfulness and evidence for representation in the acquisition of left-edge clusters in West Germanic. In *Fixing Priorities: Constraints in Phonological Acquisition*, ed. by René Kager, Joe Pater and Wim Zonneveld, 109-157. Cambridge: Cambridge University Press.

Goldsmith, John. 1990. *Autosegmental and metrical phonology*. Oxford: Blackwell.

Goldsmith, John. 1995. *The handbook of phonological theory*. Cambridge, Mass: Blackwell.

Goto, Hiromu. 1971. Auditory perception by normal Japanese adults of the sounds /l/ and /r/, *Neuropsychologia*, 9, 317-323.

Gouskova, Maria. 2001. *Falling sonority onsets, loanwords, and Syllable Contact*. Available on Rutgers Optimality Archive (ROA-491): http://roa.rutgers.edu

Grassegger, Hans. 1991. Perception and production of Italian plosives by Austrian Learners. *Proceedings of the XIIth International congress of phonetic sciences*, 5, 290-293.

Greenberg, Joseph. 1976. *Language universals*. The Hague: Mouton.

Greenberg, Joseph. 1978. Some generalization concerning initial and final consonant clusters. *Universals of human language: Phonology*, ed. by Joseph Greenberg, 243-79. Stanford, CA: Stanford University Press.

Greenberg, Joseph and James Jenkins. 1964. Studies in the psychological correlates of the sound system of American English. *Word*, 20, 157-77.

Grimshaw, Jane. 1997. Projection, heads, and optimality. *Linguistic Inquiry*, 28, 373-422.

Guion, Susan, James Flege, Reiko Akahane-Yamada, and Jessica Pruitt. 2000. An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America*, 107, 2711-2724.

Gundel, Jeanette, Kathleen Houlihan, and Gerald Sanders. 1986. Markedness and distribution in phonology and syntax. *Markedness*, ed. by Fred Eckman, Edith Moravcsik and Jessica Wirth, 107-138. New York: Plenum Press.

Halle, Morris and Jean-Roger Vergnaud. 1980. Three dimensional phonology. *Journal of Linguistic Research*, 1, 83-105.

Hallé, Pierre and Catherine Best. 2007. Dental-to-velar perceptual assimilation: A cross-linguistic study of dental stop+/l/ clusters. *Journal of Acoustical Society of America*, *Vol. 121*, 5, 2899-2914.

Hallé, Pierre, Juan Segui, Uli Frauenfelder, and Christine Meunier. 1998. Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance*, 24, 592-608.

Hancin-Bhatt, Barbara and Rakesh Bhatt. 1997. Optimal L2 syllables: Interactions of transfer and developmental effects. *Studies in Second Language Acquisition*, 19, 331-78.

Hansen, Jette. 2004. Developmental sequences in the acquisition of English L2 syllable codas: A preliminary study. *Studies in Second Language Acquisition*, 26, 85-124.

Hardison, Debra. 2003. Acquisition of second-language speech: Effects of visual cues, context and talker variability. *Applied Psycholinguistics*, 24, 495-522.

Hodne, Barbara. 1985. Yet another look at interlanguage phonology: The modification of English syllable structure by native speakers of Polish. *Language Learning*, 35, 405-417.

Holes, Clive. 2004. *Modern Arabic: Structures, functions, and varieties*. Washington, D.C: Georgetown University Press.

Hudson, Grover. 1995. Consonant release and the syllable. *Linguistics*, 33, 655-72.

Imai, Terumi. 2004. *Vowel devoicing in Tokyo Japanese: A variationist approach*. Ph.D dissertation, Michigan State University.

Ingham, Bruce. 1971. Some characteristics of Meccan speech. *Bulletin of the School of Oriental and African Studies*, *Vol. 34*, No. 2, 273-297. University of London.

Ingham, Bruce. 1982. *Northeast Arabian Dialects*. London and Boston: Kegan Paul International.

Ingham, Bruce. 1994. *Najdi Arabic: central Arabian*. Amsterdam: John Benjamins.

Jakobson, Roman. 1962. *Selected writings 1: phonological studies*. The Hague: Mouton.

Johnson, Thomas and Winifred Strange. 1982. Perceptual constancy of vowels in rapid speech. *Journal of the Acoustical Society of America*, 72, 1761-770.

Johnstone, Thomas. 1967b. *Eastern Arabian dialect studies.* London Oriental Series. London: Oxford University Press.

Jusczyk, Peter, Paul Luce, and Jan Charles-Luce. 1994. Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, 33, 630-645.

Kabak, Bariş and William Idsardi. 2007. Perceptual distortions in the adaptation of English consonant clusters: Syllable structure or consonantal contact constraints? *Language and Speech*, 50, 23-52.

Kager, René. 1999. *Optimality theory*. Cambridge, UK: Cambridge University Press.

Kajikawa, Sachiyo, Laurel Fais, Ryoko Mugitani, Janet Werker, and Shigeaki Amano. 2006. Cross-language sensitivity to phonotactic patterns in infants. *The Journal of the Acoustical Society of America*, 120, 2278-2284.

Kaye, Jonathan and Jean Lowenstamm. 1981. Syllable structure and markedness theory. *Theory of markedness in generative grammar*, ed. by Adriana Belletti, Luciana Brandi, Luigi Rizzi, 287-315. Pisa, Italy: Scuola Normale Superiore.

Kaye, Jonathan, Jean Lowenstamm, and Jean-Roger Vergnaud. 1990. Constituent structure and government in phonology. *Phonology*, 7, 193-231.

Keer, Edward and Eric Baković. 1997. Have FAITH in syntax. *Proceedings of the West Coast Conference on Formal Liguistics 16*. Available on Rutgers Optimality Archive (ROA-200): http://roa.rutgers.edu

Kenstowicz, Michael. 1994. *Phonology in generative grammar*. Oxford, UK: Blackwell.

Kenstowicz, Michael. 2003. The role of perception in loanword phonology. *Studies in African Linguistics*, 32, 95-112.

Kim, Hyouk-Keun. 2000. *The interlanguage phonology of Korean learners of English: A computational implementation of optimality theoretic constraints*. Ph.D dissertation, Georgetown University.

Kirk, Cecilia and Katherine Demuth. 2005. Asymmetries in the acquisition of word-initial and word-final consonant clusters. *Journal of Child Language*, 32, 709-734.

Kjellmer, Göran. 1998. Why is 'nrather dlifficult' rather difficult? On changes in English initial consonant clusters. *Explorations in Corpus Linguistics*, ed. by Antoinette Renouf, 83-97. Atlanta, GA: Amsterdam.

Kreitman, Rina. 2006. Cluster buster: A typology of onset clusters. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, *Vol. 42*, 163-179.

Kucera, Henry and W. Nelson Francis. 1967. *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Kühnert, Barbara, Philip Hoole, and Christine Mooshammer. 2006. Gestural overlap and C-center in selected French consonant clusters. *Proceedings of the 7th International Seminar on Speech Production*, 327-334. Ubatuba, Brazil.

Kuhl, Patricia. 1980. Perceptual constancy for speech sound categories in early infancy. *Child phonology: Vol. 2. Perception*, ed. by Grace Yeni-Komshian, James Kavanagh, and Charles Ferguson, 41-66. San Diego, CA: Academic Press.

Kuhl, Patricia. 1991. Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, 50, 93-107.

Kuhl, Patricia. 1993a. Early linguistic experience and phonetic perception: Implications for theories of developmental speech perception. *Journal of Phonetics*, 21, 125-139.

Kuhl, Patricia. 2004. Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*, 5, 831-843.

Kuhl, Patricia and Paul Iverson. 1995. Linguistic experience and the perceptual magnet effect. *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*, ed. by Winifred Strange, 121-154. Timonium, MD: York Press Inc.

Kuhl, Patricia, Erica Stevens, Akiko Hayashi, Toshisada Deguchi, Shigeru Kiritani, and Paul Iverson. 2006. Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9, 2, 13-21.

Kwon, Bo-Young. 2005. The patterns of vowel insertion in IL phonology: The P-map account. *Studies in Phonetics, Phonology and Morphology*, 11, 21-49.

Kwon, Bo-Young. 2006. *Korean speakers' production of English consonant clusters: Articulatory and perceptual accounts*. Ph.D dissertation, Michigan State University.

Ladefoged, Peter. 2001. *A course in phonetics*. Fort Worth: Harcourt College Publishers.

Ladefoged, Peter. 2005. *Vowels and consonants: an introduction to the sounds of language*. Malden, MA: Blackwell Pub.

Lado, Robert. 1957. *Linguistics across cultures, applied linguistics for language teachers*. Ann Arbor, MI: University of Michigan Press.

Larson-Hall, Jenifer. 2004. Predicting perceptual success with segments: A test of Japanese speakers of Russian. *Second Language Research*, 20, 35-80.

Lenneberg, Eric. 1967. *Biological foundations of language*. New York: Wiley.

Levelt, Clara, Niels Schiller, and Willem Levelt. 2000. The acquisition of syllable types. *Language Acquisition*, 8, 237-64.

Levin, Juliette. 1985. *A Metrical Theory of Syllabicity*. Ph.D dissertation, MIT.

Liberman, Alvin, Franklin Cooper, Donald Shankweiler, and Michael Studdert-Kennedy. 1967. Perception of the speech code. *Psychological Review*, 74, 431-461.

Liberman, Alvin and Ignatius Mattingly. 1985. The motor theory of speech perception revised. *Cognition*, 21, 1-36.

Lin, Yuh-Huey. 2001. Syllable simplification strategies: A stylistic perspective. *Language Learning*, 51, 681-718.

Lindblom, Björn. 1983. Economy of speech gestures. *The production of speech*, ed. by Peter MacNeilage, 217-246. New York: Springer-Verlag.

Lindblom, Björn and Ian Maddieson. 1988. Phonetic universals in consonant systems. *Language, speech and mind*, ed. by Larry Hyman, and Charles Li, 62-78. London: Routledge.

Lisker, Leigh and Arthur Abramson. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384-422.

Logan, John, Scott Lively, and David Pisoni. 1991. Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America*, 89, 874-886.

Major, Roy. 1996. Markedness in second language acquisition of consonant clusters. *Second language acquisition and linguistic variation*, ed. by Robert Bayley and Dennis Preston, 75-96. Amsterdam: John Benjamins.

Major, Roy. 1999. Chronological and stylistic aspects of second language acquisition of consonant clusters. *Language Learning*, 49, 123-50.

Major, Roy and Eunyi Kim. 1996. The similarity differential rate hypothesis. *Language Learning*, 46, 465-496.

Malécot, Andrei. 1958. The role of releases in the identification of released final stops. *Language*, 34, 370-80.

Massaro, Dominic and Michael Cohen. 1983. Phonological context in speech perception. *Perception and Psychophysics*, 34, 338-348.

McCarthy, John and Alan Prince. 1995. *Faithfulness and reduplicative identity*. Available on Rutgers Optimality Archive (ROA-60): http://roa.rutgers.edu

McGurk, Harry and John MacDonald. 1976. Hearing lips and seeing voices. *Nature*, 264, 746-748.

Morelli, Frida. 1999. *Markedness relations and implicational universals in the typology of onset obstruent clusters*. Available on Rutgers Optimality Archive (ROA-420): http://roa.rutgers.edu

Morelli, Frida. 2003. The relative harmony of /s+stop/ onsets: Obstruent clusters and the sonority sequencing principle. *The syllable in optimality theory*, ed. by Caroline Fery and Ruben van de Vijver, 356-71. New York: Cambridge University Press.

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

Moyer, Alene. 1999. Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition*, 21, 81-108.

Mugitani, Ryoko, Laurel Fais, Sachiyo Kajikawa, Janet Werker, and Shigeaki Amano. 2007. Age-related changes in sensitivity to native phonotactics in Japanese infants. *The Journal of the Acoustical Society of America*, 122, 1332-1335.

Murray, Robert and Theo Vennemann. 1983. Sound change and syllable structure in Germanic phonology. *Language*, 59, 514-528.

Omar, Margaret. 1975. *Saudi Arabia urban Hejaz dialect*. Washington, DC: Foreign Service Institute.

Ohala, John. 1990. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. *Papers from the Regional Meetings, Chicago Linguistic Society*, 2, 319-38.

Oller, John. 1972. Contrastive analysis, difficulty, and predictability. *Foreign Language Annals*, 6, 95-105.

Osburne, Andrea. 1996. Final cluster reduction in English L2 speech: A case study of a Vietnamese speaker. *Applied Linguistics*, 17, 164-81.

O'Shaughnessy, Douglas. 2000. *Speech communications: Human and machine*. Piscataway, NJ: IEEE Press.

Paradis, Carole. 1996. The inadequacy of filters and faithfulness in loanword adaptation. *Current trends in phonology: Models and methods*, ed. by Jacques Durand and Bernard Laks, 509-534. Salford: University of Salford Publications.

Paradis, Carole and Jean-François Prunet. 1991. *The special status of coronals: Internal and external evidence*. San Diego, California: Academic Press.

Pater, Joe and Jessica Barlow. 2003. Constraint conflict in cluster reduction. *Journal of Child Language*, 30, 487-526.

Pitt, Mark. 1998. Phonological processes and the perception of phonotactically illegal consonant clusters. *Perception and Psychophysics*, 60, 941-951.

Polivanov, Evgenij. 1931. La perception des sons d'une Langue´etrang`ere. *Travaux du Cercle Linguistique de Prague*, 4, 79-96.

Prince, Alan and Paul Smolensky. 1993. *Optimality theory: Constraint interaction in generative grammar*. Ms., Rutgers University and University of Colorado at Boulder.

Prince, Alan and Bruce Tesar. 1999. *Learning phonotactic distributions*. Technical Report TR-54. Available on Rutgers Optimality Archive (ROA-353): http://roa.rutgers.edu

Rice, Keren. 2007. Markedness in phonology. *The Cambridge handbook of phonology*, ed. by Paul de Lacy, ch.4. Cambridge University Press.

Riney, Timothy and James Flege. 1998. Changes over time in global foreign accent and liquid identifiability and accuracy. *Studies in Second Language Acquisition*, 20, 213-242.

Rochet, Bernard. 1995. Perception and production of L2 speech sounds by adults. *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*, ed. by Winifred Strange, 379-410. Timonium, MD: York Press Inc.

Sato, Charlene. 1984. Phonological processes in second language acquisition: Another look at interlanguage syllable structure. *Language Learning*, 34, 43-57.

Schütz, Albert. 1978. English loanwords in Fijian. *Fijian Language Studies: Borrowing and Pidginization*, ed. by Albert Schütz, 1-50. Bulletin of the Fiji Museum No. 4. Suva: Fiji Museum.

Selkirk, Elisabeth. 1982. The syllable. *The structure of phonological representation*, ed. by Harry Van der Hulst and Norval Smith, 337-84. Dordrecht: Foris Publications.

Selkirk, Elisabeth. 1984. On the major class features and syllable theory. *Language and sound structure: Studies in phonology presented to Morris Halle by his teacher and students*, ed. by Mark Aronoff and Richard Oerhle. Cambridge, MA: MIT Press.

Sheldon, Amy and Winifred Strange. 1982. The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, 3, 243-261.

Sievers, Eduard. 1881. *Grundzüge der Phonetik*. Leipzig: Breitkopf and Hartel.

Silverman, Daniel. 1992. Multiple scansions in loanword phonology: Evidence from Cantonese. *Phonology*, 9, 289-328.

Singh, Rajendra. 1985. Prosodic adaptation in interphonology. *Lingua*, 67, 269-282.

Smolensky, Paul. 1996. On the comprehension/production dilemma in child language. *Linguistic Inquiry*, 27, 720-731.

Smorodinsky, Iris. 2002. *Schwas with and without active control*. Ph.D. dissertation, Yale University.

Spencer, Andrew. 1996. *Phonology: Theory and Description*. Oxford: Blackwell Publishers Ltd.

Steriade, Donca. 1982. *Greek prosodies and the nature of syllabification*. Ph.D dissertation, MIT.

Steriade, Donca. 1999a. Alternatives to syllable-based accounts of consonantal phonotactics. *Proceedings of the 1998 Linguistics and Phonetics Conference*, ed. by Osamu Fujimura, Brian Joseph, and Bohumil Palek, 205-45. Prague: Karolinum. Available from: http://www.linguistics.ucla.edu/people/steriade/papers/Alternatives_to_Syllables.pdf

Steriade, Donca. 2001a. Directional asymmetries in place assimilation: A perceptual account. *The role of speech perception in phonology*, ed. by Elizabeth Hume and Keith Johnson. San Diego: Academic Press.

Steriade, Donca. 2009. The phonology of perceptibility effects: the P- map and its consequences for constraint organization. *The nature of the word: Studies in honor of Paul Kiparsky*, ed. by Kristin Hanson and Sharon Inkelas, 151-179. Cambridge, MA: MIT Press.

Stevens, Kenneth and Sheila Blumstein. 1978. Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, 64, 1358-1368.

Strange, Winifred. 1995. Cross-language studies of speech perception: A historical review. *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*, ed. by Winifred Strange, 2-45. Timonium, MD: York Press Inc.

Strange, Winifred and Patricia Broen. 1981. The relationship between perception and production of /w/, /r/, and /l/ by three-year-old children. *Journal of Experimental Child Psychology*, 31, 81-102.

Tarone, Elaine. 1987. Some influences on the syllable structure of interlanguage phonology. *Interlanguage phonology: The acquisition of a second language sound system*, ed. by Georgette Ioup and Steven Weinberger, 232-47. Cambridge, MA: Newbury House.

Tench, Paul. 2003. Non-native speakers' misperceptions of English vowels and consonants: Evidence from Korean adults in UK. *IRAL*, 41, 145-173.

Tesar, Bruce and Paul Smolensky. 1998. Learnability in optimality theory. *Linguistic Inquiry*, 29, 229-68.

Thompson, Irene. 1991. Foreign accents revisited: The English pronunciation of Russian Immigrants. *Language Learning*, 41, 177-204.

Towell, Richard and Roger Hawkins. 1994. *Approaches to second language acquisition*. Clevedon, UK: Multilingual Matters.

Trubetzkoy, Nikolai. 1939. *Principles of Phonology*. Berkeley, CA: University of California Press.

Vennemann, Theo. 1988. *Preference laws for syllable structure and the explanation of sound change: With special reference to German, Germanic, Italian, and Latin*. Berlin: Mouton de Gruyter.

Vitevich, Michael and Paul Luce. 2005. Increases in phonotactic probability facilitate spoken non-word repetition. *Journal of Memory and Language*, 52, 193-204.

Walden, Brian and Allen Montgomery. 1975. Dimensions of consonant perception in normal and hearing impaired listeners. *Journal of Speech and Hearing Research*, 18, 445-55.

Wardhaugh, Ronald. 1970. The contrastive analysis hypothesis. *TESOL Quarterly*, 4, 123-130.

Watson, Janet. 2002. *The phonology and morphology of Arabic*. New York: Oxford University Press.

Weinberger, Steven. 1987. The influence of linguistic context on syllable simplification. *Interlanguage phonology: The acquisition of a second language sound system*, ed. by Georgette Ioup and Steven Weinberger, 401-17. Cambridge, MA: Newbury House.

Werker, Janet and John Logan. 1985. Cross-language evidence for three factors in speech perception. *Perception and Psychophysics*, 37, 35-44.

Werker, Janet and Richard Tees. 1984. Developmental change across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology*, 37, 278-286.

White, Lydia. 1987. Markedness and second language acquisition: The question of transfer. *Studies in Second Language Acquisition*, 9, 261-286.

Wingstedt, Maria and Richard Schulman. 1988. Listeners' judgments of simplifications of consonant clusters. *Linguistics*, 26, 105-123.

Winitz, Harris, Marlin Scheib, and James Reeds. 1972. Identification of stops and vowels for the burst portion of /p, t, k/ isolated from conversational speech. *Journal of the Acoustical Society of America*, 51, 1309-1317.

Wright, Richard. 2004. A review of perceptual cues and cue robustness. *Phonetically-based phonology*, ed. by Bruce Hayes and Donca Steriade, 34-57. New York: Cambridge University Press.

Yamada-Akahane, Reiko, Yoh'ichi Tohkura, Ann Bradlow, and David Pisoni. 1996. Does training in speech perception modify speech production? *Proceedings of 4th International Conference on Spoken Language Processing*, 606-609.

Yeon, Sang-Hee. 2003. Perception of English palatal codas by Korean speakers of English, *Journal of the Acoustical Society of America*, 4, 23-30.

Yoo, Hyebae. 2004. A longitudinal study of consonant cluster acquisition. Retrieved April, 15, 2007, from: http://society.kisti.re.kr/~pmc/work/vol_10_3/08Hyebae%20Yoo.doc

Zsiga, Elizabeth. 2003. Articulatory timing in a second language: Evidence from Russian and English. *Studies in Second Language Acquisition*, 25, 399-432.

Zuraw, Kie. 2007. The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog infixation. *Language*, 83, 277-316.