4-12-2017

Acquisition of English /S/-Clusters by Brazilian Portuguese Speakers

Carolyn June Baker
Florida International University

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

ACQUISITION OF ENGLISH /S/-CLUSTERS BY BRAZILIAN PORTUGUESE SPEAKERS

A master's project submitted in partial fulfillment of the

requirements for the degree

MASTER OF ART

in

LINGUISTICS

by

Carolyn June Baker

2017

FIU LINGUISTICS PROGRAM

MA PROJECT

FINAL SUBMISSION

To: Director, Linguistics Program

College of Arts, Sciences and Education

This MA Project, written by Carolyn June Baker, and entitled Acquisition of English /s/-clusters by Brazilian Portuguese Speakers, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this MA Project and recommend that it be approved.

____

Virginia C. Mueller Gathercole

____

Feryal Yavaş

____

Mehmet Yavaş, Major Professor

Date of Defense: April 12, 2017

The MA Project of Carolyn June Baker is approved.

____

Prof. Virginia C. Mueller Gathercole

Director, Linguistics

College of Arts, Sciences, and Education

Florida International University, 2017

DEDICATION

This Master’s project is dedicated to my father, my rock. Thank you for your unwavering support in every endeavor I have ever undertaken. Your continuous love has never failed to motivate and inspire me.

I could have not done this without you.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor Dr. Mehmet Yavaş for his invaluable wisdom in the realm of Phonology and Phonetics. His guidance was truly appreciated every step of the way

I would also like to thank my committee members, Dr. Virginia C. Mueller for her insightful comments, editing expertise, and commitment of time to this project and Dr. Feryal Yavaş for her unwavering support in this project and throughout the duration of my Master's.

I am in deep gratitude to all the participants who volunteered their time to make this study possible, muito obrigada!

Finally, a very special thank you to Marcelo Richard Hilário, for his endless help and support throughout this project and for reminding me when I needed it most that 'ciência é mexer na sujeira".

ABSTRACT OF THE PROJECT

ACQUISITION OF ENGLISH /S/-CLUSTERS BY BRAZILIAN PORTUGUESE SPEAKERS

by

Carolyn June Baker

Florida International University, 2017

Miami, Florida

Professor Mehmet Yavaş, Major Professor

This research examines the error patterns of two member /s/ + consonant clusters (sC) in word initial position by native speakers of Brazilian Portuguese learning English. Previous research has found that not all sC clusters are modified at the same rate and studies involving native Brazilian Portuguese speakers have produced conflicting findings (Cardoso, 2008; Major, 1996; Rebello & Baptista, 2006; Rauber, 2006). The current study aims to clarify which markedness relationships influence the frequency of prothesis in sC clusters and test the opposing predictions made by two phonological principles, Sonority Sequencing and Obligatory Contour Principle for [continuant] (OCP [cont]), to determine which most accurately accounts for the productions observed. It will also address the possible contributions of native language transfer and effect of input frequency.

Participants (n=32) completed two tasks, a sentence reading and a picture-based task, and their productions were recorded for acoustic analysis. Error rates were examined using traditional transcription methods and duration of prothetic vowels were measured. Analysis showed that frequency of prothesis contradicted predictions of sonority; /s/+sonorant clusters were modified significantly more often than /s/+obstruent clusters. However, high rates of voicing on /s/ due to native language transfer may have mediated the relative influence of sonority. Results based on vowel duration indicated prothetic vowels tended to get longer when clusters became more marked. Vowel duration was longest on /sl/ and may indicate markedness effects of the OCP [cont].

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

It is widely known that learning to produce foreign language sounds that do not exist in the native language can be difficult for the foreign language learner. However, even when two languages share the same sounds, the ways in which they are combined can present complications to the learner. The rules governing the possible phoneme sequences in a language are known as phonotactics and learning a foreign language necessarily involves learning the unique rule set of that language.

Difficulties arise for Brazilian Portuguese speakers learning English words beginning with the combination of /s/ consonant clusters (sC clusters) such as '**st**art' and '**sm**all'. This research seeks to examine learners' performance with sC clusters, with the wider aim of determining the order of acquisition and of establishing a hierarchy of sC difficulty. The findings here will be used to test the opposing predictions made by two phonological principles, Sonority Sequencing and Obligatory Contour Principle, to determine which most accurately accounts for the productions observed.

Several studies have observed the effect of sonority distance on the acquisition of complex clusters (Barlow & Dinnsen, 1998; Broselow & Finer, 1991; Carlisle, 2006; Eckman & Iverson 1993; Pater & Barlow, 2003). Syllables consisting of initial consonant clusters with a large sonority distance are regarded as preferred, simpler, or less marked and are predicted to be acquired before those more marked (Carlisle 2006; Clements, 1990). However, although sonority has accounted for many patterns in onset cluster production, there are times when data are not consistent in showing the role of sonority; clusters with large sonority distance are sometimes acquired before those with small sonority distance (Abrahamson 1999; Enochson, 2014; Major, 1996; Rebello & Baptista, 2006; Yavas &Someillan, 2005).

These exceptions have led to the suggestion that alternative accounts may better represent the patterns found in data. The current study will examine the patterning of two member initial consonant clusters beginning in /s/ in the interphonology of native Brazilian Portuguese speakers acquiring English, a population in which previous research has produced contradictory results. The aim is to address possible influences, such as the role of continuance, that have been less investigated with respect to productions by these speakers. The study will analyze data using traditional transcription methods to examine error rates and will also assess a more detailed form of phonetic evidence, the duration of prothetic vowels, as an indication of the potential influence of markedness.

II. LITERATURE REVIEW

Markedness and Interlanguage Phonology

The acquisition of a second language (hereafter L2) phonology is an evolving system whereby phonological patterns are successively restructured throughout the process of L2 acquisition (Carlisle, 1991). In learning a foreign language phonology, there is a transitory phase, or an interphonology, characterized by variability in performance. Variability occurs when at least two or more surface variants appear, usually alternating between target and non-target-like forms in the production of L2 speakers. Research often seeks to determine which conditioning factors are most influential in accounting for these variants. An abundance of studies have approached these questions by investigating the relationship between markedness constraints and acquisition of a second language phonology (Abrahamsoon 1999; Carlisle 1988, 1997, 2006; Eckman 1991; Eckman & Iverson 1993; Major 1996; Rebello & Baptista, 2006). This research grew in large part in response to Eckman's (1977) Markedness Differential Hypothesis (MDH) which states that, "The areas of difficulty that a language learner will have can be predicted on the basis of a systematic comparison of the grammars of the native language, the target language and the markedness relations stated in universal grammar" (321). The areas of the target language that differ from the native language and are more marked than that of the native language will be more difficult to acquire. The degree of difficulty caused by differences between the native language and target language is predicted to correspond directly to the relative degree of markedness. Often this prediction depends on implicational universals. An example of an implicational universal for cluster length states that if a language contains an onset cluster of length N, then it also contains onset clusters of length N-1. In other words, if a language allows three segments to cluster in onsets, it will necessarily allow a two-segment cluster. A study by Anderson (1978) found evidence of this markedness relationship by investigating the acquisition of onset and coda clusters in English for learners from three different NL backgrounds, Egyptian Arabic, Mandarin Chinese, and Amoy Chinese. It was found that all native language groups made significantly more modifications as cluster length increased, thereby supporting predictions made by implicational universals.

Although the MDH was successful in predicting second language behavior for numerous structures, an important question remained: what happens when a speaker of a language that does not allow consonant clusters tries to acquire L2 consonant clusters? To explore questions such as this, the MDH was extended to include instances when the target language contains structures that are not found in the native language. To investigate markedness relationships that arise entirely within the target language, Carlisle developed the Intralingual Markedness Hypothesis (IMH) (Carlisle, 1988; 17) which states,

> If structures in the target language differ from those in the native language, and if those structures in the target language are in a markedness relationship, then the more marked structures will be more difficult to acquire than will the less marked structures.

Along similar lines to the IMH, Eckman (1991) developed the Structural Conformity Hypothesis (SCH) to explain patterns in L2 structures that do not constitute an area of difference between the NL and TL. The basic premise of the hypothesis states “The universal generalizations that hold from primary languages hold also for interlanguages” (p. 24). The SCH appeals to a set of underlying markedness principles and eliminates the requirement for a direct comparison between the NL vs TL structures. Evidence supporting the SCH would consist of an interlanguage pattern that is neither NL-like nor TL-like, but rather obeys the kinds of universal patterns found across the world’s languages. An active area of research investigating the influence of markedness principles in interlanguages has come from cases where the TL allowed both a greater number and more marked consonant clusters than did the NL (Carlisle, 1997, 1998; Eckman, 1991; Eckman & Iverson, 1994).

Attempts to explain why some consonant clusters were more difficult to acquire than others sought to establish markedness relationships to predict which structures would be more difficult for learners to acquire. Marked segments are those that are less likely to occur cross-linguistically and exist in an implicational relationship such that if a language has the more marked form, then it will also have its unmarked counterpart (Anderson, 1987). Results from studies have consistently found that the margins considered more marked are modified significantly more frequently than less marked margins in second language productions. Some of these markedness considerations include (1) cluster length (Abrahamsson, 1999; Anderson, 1997; Carlisle 1997, 1998, 2002; Greenberg 1965; Kiparsky 1979, Vennemann 1988; Weinberger, 1987), (2) phonological environments in which clusters occur (Abrahamsson, 1999; Carlisle, 1991a/b, 1992, 1994, 1997; Rauber, 2006), and (3) sonority relations between cluster members (Broselow & Finer 1991; Eckman & Iverson 1993; Barlow & Dinnsen 1998; Pater & Barlow 2003, Carlisle 2006).

Initial /s/+consonant clusters

A particular group of complex clusters that have been widely studied are those in English whose first member (C1) is /s/ followed by a second consonant (C2) (hereby sC clusters). These clusters violate several principles of syllable well-formedness for English and therefore present a fruitful testing ground for many syllable structure universals due to their exceptional behavior. For example, there is a generalization that prohibits many homorganic clusters (e.g. /pw/ and /bw/) but the clusters /st, sl, sn/ are still allowed. Another generalization in English disallows many 'obstruent + nasal' clusters (/pn, kn/), but /sn, sm/ are allowed. In addition, while other double onset clusters follow the Sonority Sequencing Principle (SSP), which dictates a rise in sonority from the outer member of the cluster to the peak of the syllable and a fall from the peak to the end, some of the sC clusters violate this principle. For example, in the /s/ + obstruent clusters (e.g. *speak* [spik], *stop* [stap], and *ski* [ski]), the sonority level drops, instead of rises, from C1 to C2.

The current study examines the acquisition of English sC clusters by native speakers of Brazilian Portuguese (BP), a language in which all complex onsets abide by the SSP. It gauges the extent to which learners observe sonority distance in their L2 initial cluster production.

Onset structure in Portuguese and English

Word-initial complex onsets in BP are restricted to obstruent plus liquid combinations. Double onsets are formed whereby C1 is a stop or /f, v/ and C2 is a tap or a lateral. More specifically, we have the following combinations: /pr, pl, br, bl, tr, dr, kr, kl, gr, gl, fr, fl, vr, tl/

(Ferreira Neto 2001, Ribas 2004). Therefore, BP prohibits sC clusters to occur in word-initial and word-internal positions. In the case of word-medial position, sC clusters syllabify heterosyllabically (e.g. de/s+t/oante → de/s.t/oante 'discordant'). When occurring word-initially, the sequence is preceded by an epenthetic [i] via prothesis resulting in syllabification into two separate syllables (#CCV→VC.CV as in *escola* [skɔla] → [is.kɔla], 'school').

The syllable structure conditions of English, in contrast, permit a greater number (33 as opposed to 14 in BP) of well-formed two-member onsets. In one set of onsets, the first member consists of an obstruent and the second an oral consonantal sonorant (/pl, pr, pj, bl, br, bj, tr, tw, dr, dw, kl, kr, kw, kj, gl, gr, gw, vj, fl, fr, fj, θr, θw, sl, sm, sn, sw, ʃr, hj, mj/). The second set consists of /s/ followed by a non-continuant: /sp/, /st/, /sk/ (Carlisle, 1991b). When Brazilian speakers encounter foreign clusters in English, their tendency is to insert an /i/ (phonetically realized as [ɪ]) to resolve the incoming unacceptable consonant cluster (Rebello & Baptista, 2006). For example, the illicit cluster in *school* [skul] would be produced as [ɪs.kul] via prothesis. In other words, the aim is to repair an input that does not meet the native language's phonotactic requirements.

When acquiring English sC clusters, frequency of prothesis is variable in production by Brazilian speakers (Major, 1999). That is, each cluster type is not modified with the same frequency. The major research question in the field has attempted to answer which markedness relationships best capture this differential treatment. The following sections explain the predictions made by the principles of sonority and continuance and how these predictions may be reflected in second language production data.

Markedness Effects on sC Clusters

OCP [continuant]

In addition to sonority, it has been suggested that the feature of continuance may play a role in the frequency of modifications of sC-initial clusters. As a general cross-linguistic tendency, Greenberg (1965) noted that languages prefer obstruent clusters that differ in their specification for the feature [continuant] over sequences that agree in this specification. Evidence for the OCP has also been found in a typological survey of markedness by Morelli (2003), whose findings indicate that out of four types of clusters (fricative-stop, stop-fricative, stop-stop, and fricative-fricative), fricative-stop clusters are the least marked typologically, even though they exhibit a more serious sonority violation.

The general tendency disfavoring identical or near identical sequences is captured by the Obligatory Contour Principle (hereafter OCP) (Yip, 1988). Clusters that violate the OCP for continuance (hereby OCP [cont]) are considered more marked because there is no change in the value of [continuant] from C1 to C2. Those that obey OCP [cont] have a change in the value of [continuant] from C1 to C2 and are therefore considered less marked. Morelli (2003) accounted for her findings by proposing a set of markedness constraints that prohibit elements in a cluster from having the same value for [continuant]. According to markedness considerations based on OCP [cont], those clusters that abide will be produced correctly more often than those that violate. In the case of sC clusters, the more marked clusters in terms of continuancy (/sl/ and /sw/) should be modified more frequently than the unmarked (/s/+stop and /s/ + nasal clusters).

In the domain of L1 acquisition, Yavaş & Somellian (2005) and Yavaş (2010) have found that children's reduction patterns may be explained by the continuancy of the second member (C2) of the consonant cluster. In these studies, children were more successful in producing /s/+approximate clusters, than /s/+stop/nasal clusters leading the authors to suggest a binary split in renditions between /s/ + [-continuant] vs /s/ + [+continuant]. In regards to L2 production, Enochson (2014) has observed a similar effect of continuance on the production of English onset clusters by native speakers of Mandarin Chinese, Cantonese, and Japanese. In an investigation of both sC clusters and CC clusters, sonority distance did not correlate with correct production in the sC clusters; instead she found an influence of OCP [cont] in the production of sC clusters. Specifically, sC clusters that obey the OCP [cont] were produced correctly more often than sC clusters in violation. She concluded that production patterns suggested a binary grouping between sC clusters that obey OCP[cont] and clusters that violate it.

The two universal tendencies, markedness based on sonority distance and markedness based on OCP [continuant] make opposite predictions about sC production. A summary of the predictions is illustrated below (where ">" indicates "acquired before")

(3) Developmental order of sC clusters: Two hypothetical learning paths

a. Markedness based on sonority: s+ approximant > s+ nasal > s+ obstruent

b. Markedness based on OCP [cont]: s+ obstruent/ s+ nasal > s+ approximant

sC Sequences in SLA: Previous Findings

For ease of exposition, the majority of the relevant studies on the L2 acquisition of sC initial clusters are compiled in a diagrammatic fashion, adapted from Cardoso (2008), illustrating authors, number of participants in each study, the L1 and L2 involved, and the order of acquisition of the sC sequences. I will comment on these, in turn, in relation to the L1.

Sonority

The function of sonority as a universal determinant in preferred syllable structure across languages has been well established in the literature. Cross-linguistically, the most preferred syllable displays a continuous rise in sonority from the periphery of the onset through the nucleus, constituting the sonority peak, followed by a continuous fall in sonority through the most peripheral member of the coda (Clements, 1990). The Sonority Sequencing Principle (SSP) captures this strong cross-linguistic generalization in syllable structure preferences (Clements, 1990). Despite its wide use as a model in explaining syllable related phenomena, the exact correlates and number of segmental classes identified in the scale differ among researchers. The most basic sequence is a five-point sonority scale proposed by Clements (1990) which indicates a rise in sonority from the obstruents through the vowels:

(1) Obstruents < nasals < liquids < glides < vowels

The current study follows Yavaş & Someillan (2005) by referencing the ten-point sonority scale put forth by Hogg and McCully (1987). Yavaş & Someillan (2005) argue that a finer cut needs to be made in the original sonority scale to allow for a more precise examination into the sonority relations among sC clusters, specifically to separate the stops and fricatives. By combining stops and fricatives into the single category ‘obstruent’, the five-point scale results in classifying /s/ +stop clusters as sonority plateaus rather than sonority reversals. The sonority scale, with the relative sonority indexes are displayed in Table 1.

Table 1: **Hogg and McCully sonority scale**

Sounds	Sonority Index (S.I.)
Low vowels	10
Mid vowels	9
High vowels	8
Flaps	7
Laterals	6
Nasals	5
Voiced fricatives	4
Voiceless fricatives	3

Voiced stops	2
Voiceless stops	1

According to the scale shown in table 1, the most expected relative sequencing of the onset of a syllable is shown as in (2) (Yavaş et al., 2008):

Onset Nucleus

(1) Voiceless stop>voiced stop>voiceless fricative>voiced fricative>nasal>lateral>flap>**V**

In other words, segments with a higher sonority tend to occur closer to the nucleus and less sonorous segments tend to occur further away.

As previously mentioned, research has established that sonority differences between cluster members translate into markedness relationships. For example, Greenberg (1965) found that in languages containing complex onsets, it is common for clusters to consist of ‘obstruent + liquid’, such as *blue* [blu] and *free* [fri], whereas reversing the sequence of sounds in these onsets, such as [lbu] and [rfi] results in clusters that are extremely rare and highly marked in terms of sonority. However, there are exceptions to generalizations based on sonority. In English, violations of SSP occur in initial /s/ + stop clusters (e.g /sp/, /st/, /sk/) which consist of sonority reversals while other sC clusters, such as /s/ + sonorant (/sm/, /sn/, /sl/, /sw/, abide by the SSP. Research in L2 acquisition has found that clusters in violation of SSP are considered to be more marked and therefore prone to modification more frequently than unmarked clusters. Much of this research has investigated native Spanish speakers learning English and has found that markedness based on sonority relations between members of a cluster plays a major role in influencing correct productions (Carlisle 1988, 1991a, 1991b, 2006; Escartin, 2005). The majority of the empirical studies on sC prothesis involving Spanish speakers of English have been conducted by Carlisle

(1988, 1991, 1999, 2006) who has consistently found that rates of prothesis were significantly lower for clusters with larger sonority distances (e.g., /sl/, /sn/, and /sm/). However, data involving Brazilian Portuguese (a typologically similar language to Spanish) speakers of English, has not been as consistent. While some authors (Cardoso, 2008) support a prediction of difficulty based sonority, other results have been statistically insignificant (Rauber, 2006; Rebello & Baptista, 2006) and even contradictory to sonority (Major, 1996).

Table 2. sC cluster in L2 Acquisition

Study	**Participants (n)**	**L1/L2**	**sC order of acquisition**
Abrahamsson (1999)	1 adult	Spanish/Swedish	sn, sm > sl, sp, sk, st
Carlisle (1988)	14 adults	Spanish/English	sl > sn > sm
Carlisle (1991b)	11 adults	Spanish/English	sl > st
Carlisle (2006)	17 adults	Spanish/English	sl > sn > st
Escartin (2005)	23 adults	Spanish/English	sm, sn > sl, sp, sk, st
Major (1996)	4 adults	BP/English	st, sp, sk > sl ,ʃr
Rebello & Baptista (2006)	6 adults	BP/English	sp, st, sk > sl, sm, sn
Rauber (2006)	10 adults	BP/English	sm, sn, sl, sw > sp, st, sk**
	9 adults	Spanish/English	sm, sn, sl, sw > sp, st, sk
Enochson (2014)	8 adults	Chinese, Cantonese, Japanese/English	s+stop, s +nasal > sl, sw
Cardoso (2008)	10 adults	BP/English	sl > sn > st

BP= Brazilian Portuguese

**= Results were not statistically significant

'>' = Acquired before

Studies involving L1 Spanish

As can be seen from table (2), numerous studies involving Spanish as the L1 have shown acquisition is predicted by sonority relations. In studies by Carlisle (1988, 1991ab, 1992, 1994, 1997 2006) participants were given reading tasks consisting of topically unrelated sentences in order to investigate markedness effects based on sonority relations. Many of these studies also aimed to determine if modification increased according to the sounds preceding the target clusters. That is, they considered phonological context as a constraint for determining frequency of prothesis. Overall, the findings showed that regardless of the degree of markedness, word-initial clusters were modified significantly less frequently after word-final vowels than after word-final consonants (Abrahamson, 1999; Carlisle, 1991a, 1992, 1994, 1997). The resulting conclusion from these studies is that modification is influenced by sonority relations but depends heavily on the preceding phonological environment.

Studies involving L1 Brazilian Portuguese

One of the earliest studies investigating acquisition of consonant clusters by Brazilian learners of English was a longitudinal study by Major (1996) which spanned approximately twelve weeks. In his study, four participants read three types of speech materials, each designed to elicit a different register of formality: word list, text, and conversation. The target sC clusters in question consisted of /sp, sk, st, sl, ʃr/, as well as several non-sC clusters. In the first two tasks, participants began by listening to a native speaker of American English reading the word list and text, after which they read each respective text five times followed by a 5 to-15-minute conversation. They repeated this three times at intervals of four weeks. The /s/+obstruent clusters were acquired before /s/+sonorant, contradicting predictions by the SSP. Major suggested that the nonconformity of the data to markedness based on sonority may be due to the adjunct status of /s/ in onset clusters (lies outside the syllable) and possibly due to positive transfer from a process in Brazilian Portuguese whereby in running speech the /i/ occurring before voiceless sC clusters is often devoiced and then deleted. His main conclusion was that NL transfer most likely influenced the results and must be taken into consideration.

A second study involving BP speakers and also producing contradictory results with respect to the influence of sonority in acquiring sC clusters was conducted by Rebello and Baptista (2006). This study examined production of initial sC clusters by six Brazilian learners of English to determine which universal tendencies would have the greatest influence on L2 syllable production. They considered as markedness constraints cluster length, sonority relations within the cluster, and preceding phonological environment. They included six total participants, two from three proficiency levels: lower intermediate, intermediate, and upper intermediate. Participants completed a reading task consisting of topically unrelated sentences (n=312). Each of the 12 clusters /sp, st, sk, sm, sn, sl, sw, spr, str, skr, spl, skw/ were placed in different preceding phonological contexts: 21 consonants, 4 vowels, and sentence initial. The only cluster excluded from analysis was /sw/ because this cluster yielded only two occurrences of prothesis out of 154 valid tokens. The authors attributed the correct productions to the occurrence of [sw] sequences phonetically in BP.

Overall, their results showed that /s/+obstruent clusters were produced with fewer errors than /s/ + sonorant, even though the former are in violation of SSP. Similar to Major (1996), the authors suggest the results may be due to NL transfer effects. Specifically, they attribute the error rates from /s/+sonorant clusters to an anticipatory assimilation in voicing of sibilants occurring in BP. For example, due to this transfer effect, the /s/ in *smoke* would undergo voicing assimilation due to the voicing on the nasal [m] in C2 to be produced as [zmok]. The resulting voiced cluster is considered more marked than a voiceless cluster according to Greenberg's (1965) implicational universals of consonant clusters, and therefore prone to more modification via prothesis (e.g. [izmok]). The second transfer process the authors speculate may have contributed to the results is the vowel deletion in /s/+obstruent clusters that occurs in quickly spoken native speech, as mentioned by Major (1996). The findings from this study demonstrate that the SSP was shown to have limited influence on L2 production patterns. The authors concluded that native language transfer, specifically due to voicing and positive transfer from vowel deletion, may have intervened in determining the relative influence of markedness relationships in the interphonology.

In a related study, Rauber (2006) compared the influence of cluster length, the sonority sequencing principle, and environment on the production of English initial sC clusters by Argentine Spanish and Brazilian Portuguese speakers. She used the same instrument for 10 BP and 9 Spanish speakers to determine whether the NL has an influence on TL production and to clarify the contradictory findings in Rebello & Baptista (2006). Overall, more frequent prothesis was found for /s/ + obstruent clusters than for /s/ + sonorant clusters, but only for Spanish speakers. In reference to Rebello & Baptista (2006), similar conclusions were drawn by Rauber (2006) in that native language transfer (from BP) was found to mediate the type and degree of influence of markedness.

An additional component that has been previously examined is the frequency with which certain linguistic structures occur in a language. Cardoso (2008) examined the development of English sC homorganic sequences (st, sn, sl) in the interlanguage of 10 BP speakers learning English in a classroom environment (low-intermediate and advanced) to investigate the effects of markedness and input frequency in acquisition. He chose homorganic clusters to ensure that sonority was the only markedness feature upon which the three target clusters differed. He compiled frequency data from an oral corpus consisting of student-directed speech of an English teacher over a two-month period to determine the relative frequencies of cluster types in the input. The prediction was made that if L2 learners are sensitive to the frequency of these clusters in the input, they should acquire the most frequent form /st/ before /sl/ and /sn/ (where '>' means acquired first). However, if they are more sensitive to markedness based on sonority, they should acquire /sl/ > /sn/ > /st/.

The findings were that sonority based markedness, and not frequency, most accurately predicted the developmental learning path of sC clusters. In other words, the least frequent clusters (/sl/ and /sn/), which were also the least marked in terms of sonority, were acquired before the most marked but more frequent cluster (/st/). Similar results have been found in an examination of the acquisition of English sC clusters by Turkish speakers (Yavaş & Altan, 2016) who found that both lexical and phonotactic frequencies did not seem suggestive in any way, of the patterns in the data.

Lastly, a study by Carlisle (2016) involved BP speakers to examine the effect of environment in determining the frequency of prothesis. The first hypothesis under investigation was that native speakers of Portuguese will use prothesis before English sC initial clusters significantly more frequently when the preceding context contains word-final consonants than after word-final vowels. This pattern has been found with native Spanish speakers learning Swedish (Abrahamsson, 1999) and English (Carlisle, 1991a, 1991b, 1992, 1997 & 2006), but never with Portuguese speakers (Rauber, 2006; Rebello & Baptista, 2006). To test this, Carlisle (2016) used a data gathering instrument consisting of 336 sentences each containing one of six target clusters: /st/, /sp/, /sk/, /str/, /spr/ and /skr/. These clusters occurred twice after 28 word final segments consisting of 14 obstruents, 3 nasals, 2 liquids, 3 diphthongs, and 6 vowels. He found that correct productions patterned in accordance with previous studies with native Spanish speakers, in that a greater percentage of target clusters were produced correctly after vowels than after consonants.

In addition to examining the distinction between word-final vowels and consonants, he further refined his study to include an examination of markedness relationships within four word-final demisyllables that each ended in a segmental class that differed in sonority: vowel + obstruent (VO), vowel + nasal (VN), vowel + liquid (VL) and vowel + glide (VG). The results showed the greater the sonority of the last segment in the two-member word-final demisyllable, the lower the frequency of prosthesis. That is, he found the highest rates of prothesis when the preceding environment contained an obstruent (Carlisle, 2016).

The Current Study

Given the conflicting findings with respect to Brazilian Portuguese speakers, the current study aims to clarify the contribution(s) of markedness and native language transfer in determining the variable production of English sC clusters. The primary markedness constraints considered will be sonority relations between cluster members and markedness with respect to the OCP [cont], as well as consideration of any mediating NL transfer effects, namely voicing assimilation and vowel deletion. A key component of this research will include an acoustic investigation of the prothetic vowel to provide a more nuanced investigation into production patterns. In addition to reporting frequency of prothesis, the study will include measurements of duration of the prothetic vowel.

Two previous studies using vowel duration have found the duration of the prothetic vowel to be an indicator of markedness (Hussain, 2014; Yavaş & Altan, 2016). Hussain found that for Saudi Arabian L2 English learners, the duration of the prothetic vowel increased as the markedness of the cluster increased. In Yavaş and Alton’s (2016) study on the acquisition of

English sC clusters by Turkish speakers, long epenthetic vowel durations were linked to lower success rates of cluster productions. Since none of the previous studies involving native Brazilian Portuguese speakers have included this type of analysis, an aim within the present study is to determine if phonological inconsistencies can be resolved using phonetic evidence.

Research Questions

The research questions are the following:

1. Does frequency of prothesis reveal differential treatment of two member /s/-clusters with respect to varying C2?

2 Is there binary categorization based on clusters considered ‘SSP-violating’ vs. ‘SSP-following’ or ‘OCP[cont]- violating’ vs. ‘OCP[cont]- abiding’?

3. Will duration of the prothetic vowel be an indicator of markedness?

4. Does lexical frequency or phonotactic probability play a role in frequency of prothesis?

5. Is voicing of the preceding environment, /t/ or /d/, an influential factor for prothesis?

III. METHODOLOGY

Participants

Participants included 32 native Brazilian Portuguese speakers currently residing in Belo Horizonte, Brazil. Out of the 32 participants, 14 were excluded on the basis of 100% correct productions (i.e. no rates of prothesis). Of the 18 remaining, the ages ranged from 18-38 years of age (M=23.4) and included 8 females and 9 males. Fourteen participants had learned English through private English courses and four reported self-study as the method of learning. All reported no knowledge of other languages. Proficiency was assessed through self-reported questionnaires. The complete questionnaire is given in appendix D. Participants were divided into two groups based on self-reported auditory proficiency (listening and speaking abilities). The maximum score for total auditory proficiency was 12 points (6 for each ability). The first group included those who reported an auditory proficiency score of under 6 points (n=8) and the second a score of 6 points and higher (n=10)

Materials

The materials used for the data collection in this study consisted of two types of sentence reading tasks: a formal sentence reading task and a picture based sentence reading task. The formal task involved reading two lists containing 40 sentences each. Of the 40 sentences, 30 sentences contained a word with one of the target sC initial clusters (/sp/, /st/, /sk/, /sm/, /sn/, and /sl/) and 10 filler sentences. The cluster /sw/ was excluded based on previous findings showing less than 1% error rate for this cluster, most likely due to the appearance of this sequences phonetically in Portuguese (e.g. *suiça* [swisɑ] ‘swiss’) (Rauber, 2006; Rebello & Baptista, 2006).

Preceding environment is not one of the central research questions in the current study, therefore, the preceding environment in stimuli will be strictly controlled to only two contexts, word final [t] and [d]. Given previous findings by Carlisle (2016), the phonological environments in the current study were chosen with the aim of inducing more prothesis. Also, since it was not a primary goal of the present study to investigate environment, the limited number will serve to provide more tokens for direct comparison. By restricting preceding environment to alveolar stops, there will be a greater number of items to analyze and will also present an opportunity to determine if voicing of the preceding environment influences rates of prothesis. In the formal reading task, 5 words for each of the six target clusters appeared twice after the environments /t/ and /d/, for a total of 10 sentences for each cluster. The word lists are given in Appendix A. These sentences were then divided amongst the two blocks (labeled A and B and counterbalanced). In other words, if the target word *speak* appeared in block A with a final [d] in the preceding word, it would appear in block B after a word ending in [t]. A complete set of stimuli is given in Appendix B.

In the picture based task, four words of each cluster type were chosen from the formal reading task based on picturability for a total of 24 target sentences and 16 fillers. Two words for each cluster type were preceded by /t/ and the other two by /d/. These target words were presented in picture form above the target sentence and their place within the sentence was indicated by an underscore "______". A complete sentence list is given in Appendix C.

Procedure

The participants were individually tested in a quiet office space at Federal University of Minas Gerais, in Belo Horizonte, Brazil. Each session began with a brief presentation of directions and a general goal of the study, with no revelation of the precise focus of the investigation. Directions were presented on the first PowerPoint slide and read verbally in Portuguese and English by the investigator:

> Portuguese directions:
>
> "Vou pedir para você ler as frases e vou gravá-lo lendo-as. Serão 3 blocos contendo 40 frases cada um. Em um desses blocos, as frases tem uma palavra faltando. Você verá, ao invés, uma foto mostrando o objeto cujo nome deve ser pronunciado ao ler a frase. Caso você não saiba o nome daquilo que a foto mostra, por favor me pergunte e eu irei mostrá-lo."
>
> English directions:
>
> "I will ask you to read the phrases and I will record you reading them. There will be three blocks each containing 40 phrases. In one of these blocks, the phrases have a word

missing. Instead, you will see a photo showing an object whose name you should say aloud while completing the phrase. If you don't know the name of the object in the photo, please ask me and I will play you a recording of the word."

Order of tasks were counterbalanced so that a participant either received the reading task or the picture based task first. The order of the two blocks of the formal reading task were further counterbalanced resulting in a total of four possible orders of presentation:

Block A- Block B-Picture	Picture- Block A- Block B
Block B- Block A-Picture	Picture-Block B- Block A

All stimuli were presented via PowerPoint slides on a laptop computer. For the formal reading tasks, participants were instructed to read the sentences at a normal/natural rate. After each sentence, the researcher moved onto the next slide by pressing the space bar. In between the formal reading task and the picture based task a written language background questionnaire was given as well as a short break.

In the picture task, participants were instructed to read the sentence by filling in the blank with the word that the picture represented. If participants chose a semantically similar word (i.e. *shop* instead of *store*), they were asked by the researcher to provide another word if possible. If a participant didn't know a word, the researcher played a pre-recorded audio clip of the target. This occurred on average five times for each participant. In this case, the participant would reread the sentence using the word they had just heard. When the sentence was completed, the researcher progressed to the next slide. All audio were recorded in Praat as well as WavePad Masters Edition v 6.33.

IV. RESULTS

The following presents the results, first, for the occurrence of prothesis, and then for the length of prothetic vowels when these occurred.

Frequency of Prothesis:

The results are reported first separately for the two tasks, since the absolute numbers of trials per condition differed in the two. Then the two tasks are compared for just those items that were common to the two tasks.

In total, the tasks produced 1,512 clusters. 49 clusters out of the total had to be removed either due to other simplification strategies (/s/ deletion n=7), change of C2 (n=15), substitution with words other than the target (i.e. 'shop' instead of 'store', n=7) and change of preceding environment (n=20). These exclusions left 1,463 clusters for analysis. In total, there were 514 instances of prothesis. Frequencies and rates of prothesis are reported in Table 3 for each cluster type.

Table 3 Frequencies and rates of error (prothesis) for each cluster type from all tasks

	SL	**SM**	**SN**	**Total /s/+sonorant**	**SP**	**ST**	**SK**	**Total /s/+obstruent**
N	249	235	250	734	244	241	244	729
Errors (prothesis)	104	101	125	330	67	56	61	184
Percentage Prothesis	42%	43%	50%	45%	27%	23%	25%	25%

Sentence Reading Task:

An analysis of variance was carried out to investigate the effects of: (1) cluster type and (2) preceding environment (/t,d/) in the sentence reading task. Cluster type (/sp/, /sk/, /st/, /sm/, /sn/, and /sl/), preceding environment (/t/ and /d/), and auditory proficiency (under 6 points, and 6 points and over) were entered as independent variables into the analysis. The dependent measure, inaccuracy, was determined by the presence of a prothetic vowel. That is, if participants modified a target word through insertion of a prothetic vowel, their response was counted as incorrect. The results of the analysis demonstrated significant effects for cluster type F (5, 12) = 13.31, p <.001. The cluster type with the most modification was /sn/ (M=2.5), followed by /sm/ (M=1.88) and then /sl/ (M=1.74). The prothesis rate for /sn/ was significantly different than all other clusters (/sl/, p=.002; /sm/, p=.009; /sk/, p=.000; /sp/, p=.000; /st/, p=.000). There was no significant difference between /sm/ and /sl/ (p=.460). The fourth most frequently modified cluster was /sk/ (M=1.18), which was significantly different from all sonorant clusters (/sl/, p=.032; /sm/, p=.003; and /sn/, p=.000). There was no significant difference between /st/ (M=1.113) and /sp/ (M=1.02) (p=.196).

Overall, the sonorant clusters were produced less successfully than the obstruents. While the obstruents all patterned together, there was some separation between /sn/ and the other sonorants. The following pattern emerges out of these significances, starting from greatest rates of prothesis to least we have (where '>' indicates higher rates of prothesis):

/sn/ > /sm/,/sl/ > /sk/,/st/,/sp/

As for effects of auditory proficiency, although the rates of prothesis were higher for the lower proficiency group (M=1.88, SD=.365) than that of the more proficient group (M=1.26, SD=.327), this difference was not statistically significant.

Picture task:

In the picture task, an analysis of variance was carried out to investigate the effects of: (1) cluster type and (2) preceding environment (/t,d/). Cluster type (/sp/, /sk/, /st/, /sm/, /sn/, and /sl/), preceding environment (/t/ and /d/), and auditory proficiency (under 6 points, and 6 points and over) were entered as independent variables into the analysis. The analysis showed a statistically significant effect of cluster type on rates of prothesis F(5, 12)=7.60, p <.001. In this task, there was a general trend for prothesis to be highest in /sl/ (M=1.21), followed by /sn/ (M=1.08), then /sm/ (M=.981). There were no significant differences in modification rates among /s/+sonorant clusters (p>.05). Of the /s/+obstruent clusters, the trend was for /sp/ to be modified most frequently (M=.856), followed by /st/ (M=.519) and lastly, /sk/ (M=.575). Pairwise comparisons showed that /sp/ was statistically different from /sl/ (p=.003), /sm/ (p=.007), and /sn/ (p=.002). Differences in rates of prothesis were not statistically different among /s/+obstruent clusters (p>.05).

From most modification to least, the following pattern emerges:

/sl/, /sn/, /sm/> /sp/ /sk/, /st/

The same general trends found in the sentence reading task for preceding environment and auditory proficiency hold in the picture task. Specifically, there was a trend for higher rates of prothesis to occur in the preceding context of [d] (M=.906, SD=.135) than of [t] (M=.837, SD=.098), however this did not reach statistical significant (p=.410). There was also a trend for higher rates of prothesis to occur in the lower proficiency group (M=1.01, SD=.165) than in the higher proficiency group (M=.733). Again, these results did not reach statistical significance (p=.222).

Picture Vs. Sentence reading task:

Lastly, an analysis of variance was carried out for the target words that appeared in both the picture and sentence reading tasks in the same context (preceding environment matched) to directly compare the effects of task type. Cluster type, preceding environment, auditory proficiency, and task type (picture task vs. sentence reading task) were entered as variables, and the dependent measure was occurrence of prothesis. The results showed a main effect of task type, $F(1,16)= 5.54$, $p=.032$, and of cluster type, $F(5,12)= 13.07$, $p < .001$. The effect of task type showed that participants produced more prothesis in the picture task, with a mean rate of prothesis of .872, than in the sentence reading task, mean of .626 (*p*=.032). This suggests that the picture task may have been more difficult, possibly due to different processing demands, which led to higher rates of prothesis.

As for the main effect of cluster type, the trend for the cluster to be most frequently modified was /sn/, followed by /sl/ and then /sm/, with means of 1.10, .93, and .78, respectively. The modification of /sn/ was significantly greater than that of /sm/ (*p*=.002), /sk/ (*p*=.000), /sp/ (*p*=.000) and /st/ (*p*=.000). The only /s/+sonorant clusters to be significantly different from each other were /sm/ and /sn/ (p=.002). Out of the obstruent clusters, the trend was for /sp/ to be modified most frequently, followed by /sk/ and then /st/, with means of .672, .572, and .434, respectively. However, differences among /s/+obstruent clusters are not significant (p >.05). From these pairwise comparisons, the following pattern emerges:

/sn/, /sl/ >/sm/ > /sp/, /sk/, /st/

There were also significant interactions of Task Type X Cluster Type, $F(5,12)= 3.11$, $p=.013$ and Cluster Type X Preceding Environment $F(5,12) = 3.15$, $p =.012$. Results are shown in figures 1 and 2, respectively. The Task Type X Cluster interaction shows that participants performed better on all clusters when they appeared in the sentence reading task, except for /sn/ (M=1.112, *p*=.011) which was produced with more error. The cluster /sn/ also appeared as the oddity in the Cluster Type X Preceding Environment interaction appearing with significantly greater prothesis after [d] (M=1.344) than after [t] (M=.850).

Figure 1 Task Type X Cluster Type Interaction

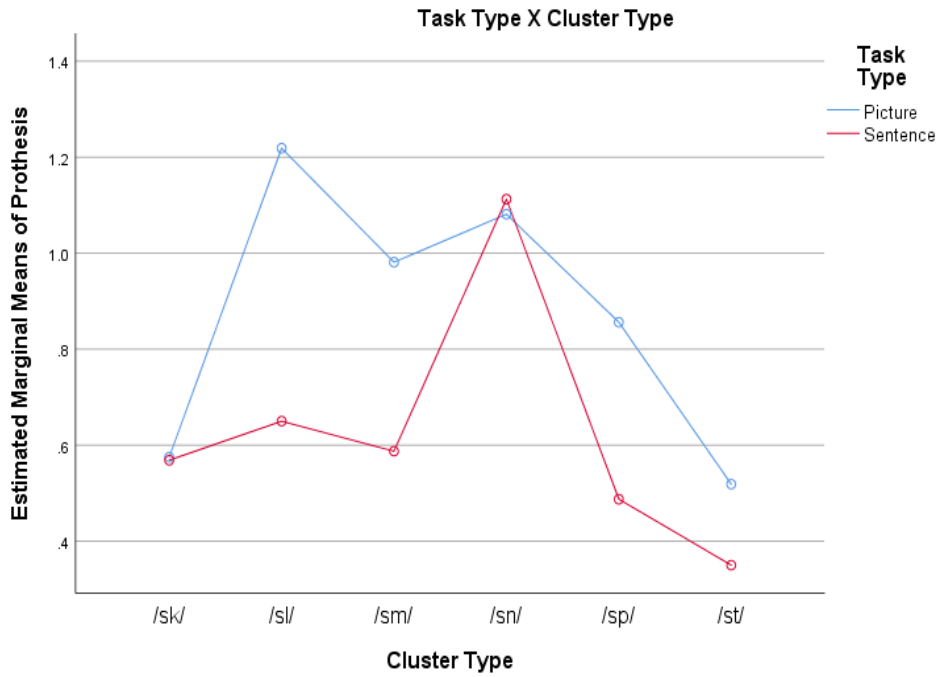
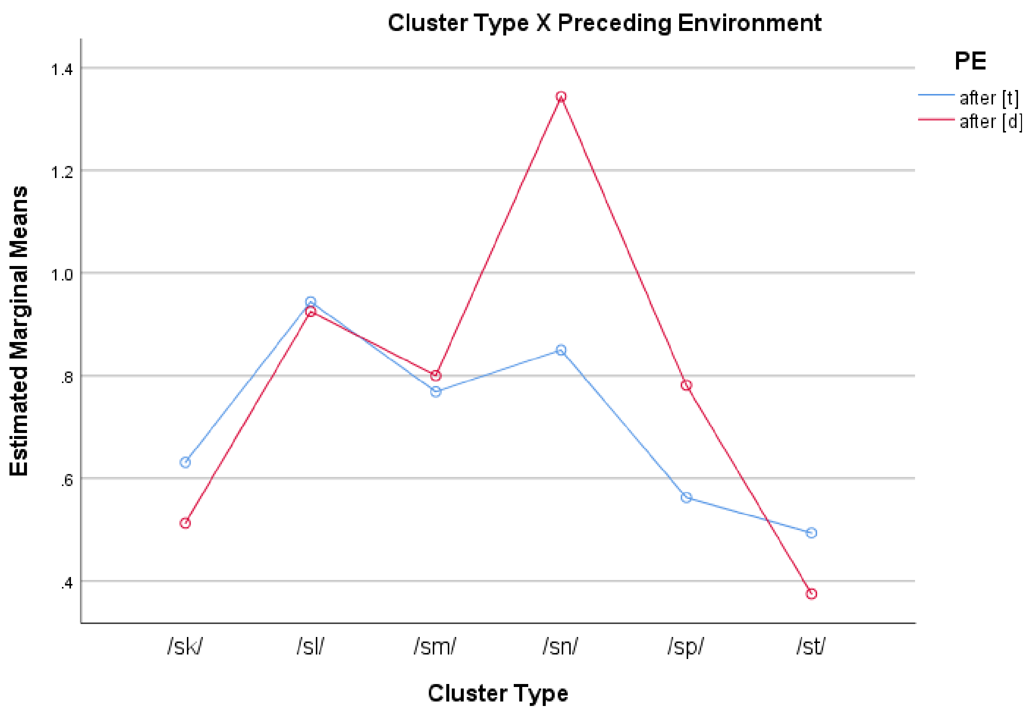


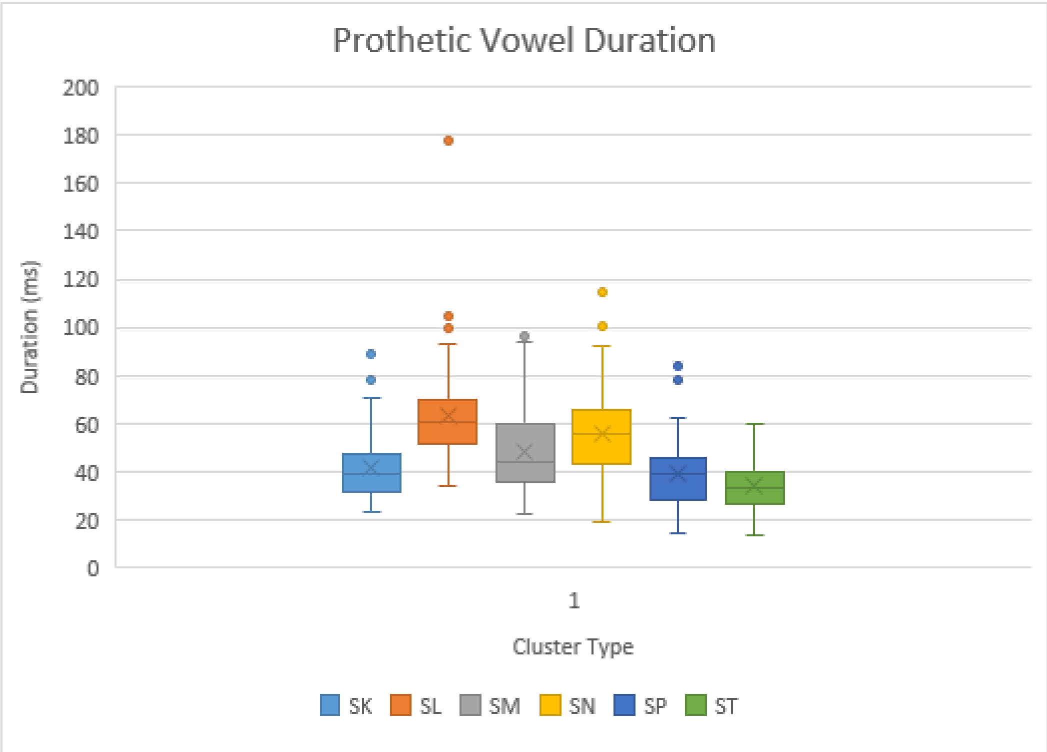
Figure 2 Cluster Type X Preceding Environment Interaction



Duration of Prothetic Vowel:

Prothetic vowels were analyzed to further examine the markedness of cluster types. Duration of the prothetic vowel was considered here to be an indicator of increased effort in production (i.e longer vowel indicates increased difficulty). Vowel lengths are shown by cluster type in Graph 3.

Figure 3 Prothetic vowel duration among cluster types.



An analysis of variance was conducted to compare mean durations of prothetic vowels by each cluster type across tasks. Cluster type and auditory proficiency were entered as independent variables. A main effect of cluster type on vowel durations was found $F(5, 12)=13.74$, $p<.001$. The cluster with the longest vowel duration was /sl/(M=64.25) and was significantly different from all other clusters ($p<.001$). Clusters /sn/ (M=54.94) and /sm/ (M=47.82), were the next longest and were not significantly different from each other (*p*=.196). The next longest clusters were /sk/ (M=42.93), /sp/ (M=38.21), and lastly /st/ (M=33.80). Vowel duration in /sk/ was significantly different from /sl/ (*p*=.001), /sn/ (*p*=.023), and /st/ (*p*=.027). /st/ was statistically different from /sn/ (*p*=.000), /sm/ (*p*=.002), /sl/ (*p*=.000) and /sk/(*p*=.027). Following from pairwise comparisons in vowel duration differences, the following pattern emerges in order from greatest duration to least.

/sl/ > /sn/, /sm/ > /sk//sp/, /st/

Voicing and Prothesis Rates:

As previous research found an influence of voicing assimilation on the /s/ in sC clusters on prothesis, the relationship between voicing and prothesis was also investigated (Major 1996; Rebello & Baptista, 2006). Table 4 shows the rates of voiced and voiceless sibilants and their respective rates of prothesis.

Table 4 Rates of voiced and voiceless sibilant clusters and prothesis

Sibilant:	Voiced	Voiceless	Total
N clusters	343	1083	1426
Prothesis	231	283	514
Percent Modified	67%	26%	36%

As the table shows, prothesis is more frequent when /s/ becomes voiced, resulting in a voiced cluster. Prothesis occurred 67% of the time in voiced clusters. The only cluster type to undergo voicing was the sonorants, which were voiced 47% of the time. Table 5 shows prothesis rates for each individual cluster type.

Table 5 Errors in voiceless and voiced sibilant clusters with sonorant C2

	SL	**SM**	**SN**	**Total**
Voiceless	25	34	40	99
Voiced	79	67	85	231

Percent Voiced	76%	66%	67%	

As the table above shows, the /sl/ cluster was voiced most frequently (76%), followed by /sn/ (67%) and lastly /sm/ (66%).

Frequency:

Cluster frequency was analyzed to determine if frequency plays a role in mediating occurrence of prothesis by cluster type. Three different measures of frequency were collected to examine the possible influences of phonotactic frequency, cluster type frequency, and frequency of target words by cluster type.

The first investigation into frequency examined phonotactic probability of different /s/ clusters for English from Vitevitch and Luce (2004). The following table presents phonotactic probabilities by cluster type in order from most probable to least. A side by side comparison of mean rates of prothesis is given in table 6, beginning with most successful renditions to least. If phonotactic probabilities predict prothesis rates, we should see a match between most probable cluster type and most successfully produced cluster (lowest means).

Table 6 Comparisons of phonotactic probabilities of cluster types and results from sentence and picture tasks

Cluster	Phonotactic Probability
/st/	0.0177
/sp/	0.0091
/sk/	0.0041
/sl/	0.0041
/sm/	0.0017
/sn/	0.0015

Sentence Reading		**Picture Task**	
Cluster	Mean	Cluster	Mean
/sp/	1.019	/st/	0.519
/st/	1.113	/sk/	0.575
/sk/	1.188	/sp/	0.856
/sl/	1.738	/sm/	0.981
/sm/	1.875	/sn/	1.081
/sn/	2.469	/sl/	1.219

As can be seen from table 6, the order of phonotactic probability closely matches error rates from both the sentence and picture tasks. The most probable cluster sequences (/s/+obstruents) are also those with the least errors in renditions, while the /s/+sonorant clusters are the least probable and produced with more error. Except for a switch between /sp/ and /sk/ clusters, the ordering of phonotactic probabilities closely mirrors those of cluster renditions. This may suggest that phonotactic probability could be influential in correct renditions by participants.

To account for cluster type frequency, two lexical data bases were consulted and compared with the renditions. The first was the CELEX lexical data base (Baayen et. al). In the 7,217 entries from all word-initial CC clusters in non-compound words, there was a total of 1,856 /sC/ clusters. Table 7 shows frequency data from the CELEX corpus in comparison with error rates in the sentence and reading tasks.

Table 7 Frequency data from CELEX Corpus compared with results from sentence and picture tasks

CELEX Corpus	
/st/	559
/sp/	328
/sk/	324
/sl/	239
/sn/	146
/sm/	92

Sentence Reading		**Picture Task**	
Cluster	Mean	Cluster	Mean
/sp/	1.019	/st/	0.519
/st/	1.113	/sk/	0.575
/sk/	1.188	/sp/	0.856
/sl/	1.738	/sm/	0.981
/sm/	1.875	/sn/	1.081
/sn/	2.469	/sl/	1.219

Again, we see similar concordance between error and frequency rates. The only discrepancy is the reverse ranking of /sn/ and /sm/ clusters in cluster frequencies. Overall, the CELEX database captures the distinctions between the /s/+stop clusters, which are both more frequent and more successfully rendered than the less frequent /s/+sonorant clusters.

The next frequency measure consisted of the frequency of each cluster type based on the stimuli words in the current study. These frequency measures were taken from the SUBTLEXus Corpus. The SUBTLEXus database is based on 50 million words whose frequency measures are derived from television and movie subtitles from the U.S. Since 15 out of 18 participants reported instances of learning English through American television series, the database is thought to provide an ecologically valid corpus to draw frequency samples from to better approximate everyday language exposure (Brysbaert & New, 2009). To determine cluster type frequency, the complete word list used in the experiment was entered into the online SUBTL database and average frequency measures were obtained for each cluster as shown in table 8.

Table 8 Frequency data from SUBTLXus database

Words Used for Sentence Reading			Words used in Picture Task		
Cluster	FREQcount	SUBTLwf	Cluster	FREQcount	SUBTLwf
/st/	7466	146.39	/st/	11166.25	218.945
/sk/	3488.33	68.39	/sk/	5005.25	98.1425
/sl/	2871.67	56.31	/sm/	4227	82.8825
/sm/	2831.33	55.52	/sl/	4174.25	81.8475
/sp/	2346	46	/sp/	3454.5	67.7375
/sn/	720.67	14.13	/sn/	742.75	14.5625

*SUBTLWF is a standard measure of word frequency independent of corpus size, based on frequency per million words.
*FREQcount is the number of times the word appears in the corpus.

Lexical frequencies from the SUBTL database do not seem particularly suggestive of the differences seen in the data, although there are points of comparison. For example, the /sn/ cluster is least frequent and produced least successfully in the sentence reading task and second least successfully in the picture task. Points of divergence are found with the /sp/ cluster, which is the second most infrequent cluster, however it is produced most successfully in the sentence reading task. Also, the /sl/ cluster words are fourth most frequent in the picture task, however they are the least successfully rendered. This divergence may be the result of the type of frequency measures utilized. It may be that an analysis of individual lexical items could be more valuable than averaging frequencies by cluster type. For example, two of the target words used for /sn/ clusters ‘sneeze’ (#32) and ‘snore’ (#30) were two of the most infrequent words out of the total 36 target words included according SUBTL database. The infrequency of these particular words may have contributed to higher error rates for that cluster. The complete frequency list for the stimuli is given in appendix E.

V. DISCUSSION

The results for frequency of prothesis revealed higher success rates in productions of /s/+obstruent clusters (75%) over /s/+sonorants (55%). In other words, as C2 becomes more sonorous, prothesis occurs more frequently. In relation to the first research question, “does frequency of prothesis reveal differential treatment of two member /s/-clusters with respect to varying C2?”, the results clearly show differential treatment of cluster types. The findings from the sentence reading and picture tasks indicated that /s/+sonorant clusters are modified significantly more often than /s/+obstruent. Between the two tasks, there were minor differences with respect to the subgroupings of clusters. For example, in the sentence reading task, /sn/ emerged as the most frequently modified cluster, while in the picture task, the trend was for /sl/ to be the most modified. Sub-groupings of clusters showed that in the sentence task, /sn/ was significantly different from all other clusters, followed by a grouping of /sm/ and /sl/. The patterning of sonorant clusters in the picture task revealed a binary grouping of /s/+sonorants versus /s/+obstruents. Results showed the obstruent clusters, /sk/, /st/, and /sp/ patterning together in each task.

The second research question asked if there would be a binary categorization based on clusters considered ‘SSP-violating’ vs. ‘SSP-following’ or ‘OCP[cont]- violating’ vs. ‘OCP[cont]- abiding’. The general impression given from these data appears to contradict the markedness predictions based on sonority. It was found that clusters with a sonorant C2, which are most well-formed in terms of sonority, were produced least successfully. In contrast, the /s/+obstruent clusters, those in violation of the SSP, were produced most successfully. However, as the results indicated, the most frequently modified clusters were also those that underwent sibilant voicing assimilation most often; 67% of prothesis occurred in voiced clusters, whereas only 26% was found in voiceless clusters. While this is not quite as strong as rates found in Rebello and Baptista's (2006) study, which was 79% for voiced and 39% for voiceless, this could be due to the lower overall frequency of prothesis in this study (36%) compared to theirs (62%).

With regards to voicing assimilation, the findings from this study are in concordance with those from Rebello and Baptista's (2006) and Major's (1996), which found that the role of native language transfer may be responsible for predicting success in L2 performance. The influence of voicing may be related to Greenberg's implicational universals regarding consonant clusters. In Greenberg's (1965) investigation into typological universals of syllable structure across languages, there was a general preference for voiceless clusters over voiced clusters. For example, there was an implicational relationship favoring voiceless obstruent + nasal over voiced obstruent + nasal. His data were insufficient to make a similar generalization about voicing in obstruent + liquid clusters. As Carlisle (2001) has argued, findings from L2 production research have revealed that the implicational relationships found in Greenberg's (1965) study can be used to determine which syllable structure universals may have an influence on interlanguage phonology. From the implicational relationships found in Greenberg's (1965) investigation, the voiced clusters (resulting from voicing assimilation produced by BP speakers) would be considered more marked and therefore prone to a greater rate of prothesis.

In addition, there are other perspectives which may also account for the patterns found in the data. Based on work by Goad (2011), who espouses a coda analysis of /s/ in sC clusters, cross-linguistic preferences on the sonority profile of the C2 in languages that permit sC clusters reveals a pattern strikingly similar to the data presented here. She found that as the sonority of C2 increases, the well-formedness of the cluster decreases. In a typological comparison of Spanish, French, Greek, Dutch, English, and Russian, Goad (2011) shows that no languages that permit sC clusters forbid /s/+stop. The typological evidence suggests the following scale of sC cluster well-formedness: /s/ + stop > /s/ + nasal > /s/ + lateral > /s/ + rhotic (where '>' means is more harmonic/well-formed than). The pattern of successful productions of /s/+stop over /s/+sonorants observed by Brazilian Portuguese speakers are supportive of this typological prediction.

Returning to the second question regarding the predictions based on OCP, the data do not present a conclusive binary grouping between 'OCP [cont]- pviolating' vs. 'OCP [cont]- abiding'. The only definitive grouping that occurs is between /s/+sonorant vs. /s/+obstruent clusters. While frequency of prothesis may not be indicative at this point, there is potential to determine markedness by looking at the duration of the prothetic vowel. Specifically, the longest vowel durations in /sl/ clusters suggest continuance may affect duration. /sl/ is the only cluster in violation of the OCP for [continuance] because it doesn't demonstrate a change of continuance from C1 to C2 ([+cont][+cont]). Under the perspective of the OCP [cont], by disobeying the preference for a change in the feature of continuant, /sl/ is an ill-formed cluster and should undergo modification more frequently.

In answering research question three, "will duration of the prothetic vowel help to rank clusters based on their relative degrees of markedness?" the results suggest that prothetic vowel duration can be taken as an indicator of increased difficulty and would place /sl/ as the most marked cluster. In accordance with previous research, investigations into the phonetic characteristics of error patterns such as prothetic vowel durations prove to be a fruitful course to pursue in determining markedness (Hussain, 2014; Yavaş & Altan, 2016).

The fourth research question asked if frequency may explain participants' error patterns. Frequency measures for cluster types and phonotactic probabilities were obtained to examine this question. If cluster frequency plays a role in frequency of prothesis, we should expect to see greater rates of prothesis in the infrequent clusters. From the frequency comparisons, phonotactic probabilities, most clearly mirrored the observed patterns in the data. Although a comparison is established, future work would be required to determine if prothesis results from frequency considerations alone. This would likely include a larger quantity of target words from each cluster type ranging from highly frequent to infrequent to observe which lexical items are accompanied by modification. Future investigations at the individual lexical level should also be conducted to explore frequency effects more accurately.

In answering the last research question, "Is voicing of the preceding environment, /t/ or /d/, an influential factor for prothesis?" there was no effect of preceding environment found in the data and therefore was not determined to be an influential factor for prothesis. The only instance when preceding environment became significant was the interaction effect between Cluster Type X Preceding Environment in the direct comparison between the picture and sentence tasks. In this case, /sn/ was produced incorrectly more often when the preceding word ended in [d] than in [t]. These results are not in accordance with previous work by Carlisle (2016) who found that as the markedness of the environment preceding the sC cluster decreases, the frequency of correct production increases. He concluded that the higher the sonority value of the final segment, the more preferred the final demisyllable. Between [t] and [d] then, the optimal environment then is [d] because it has slightly higher sonority per the Hogg and McCully sonority scale employed by the current study. While there is not a definitive answer as to why [d] yielded higher error rates, a conjecture would be that the voicing of [d] may have strengthened the voicing assimilation of the sibilant in sonorant clusters.

VI. CONCLUSION

Summary

Acquisition of initial English sC clusters by Brazilian Portuguese learners was examined to determine the influence of markedness based on sonority and the Obligatory Contour Principle for [continuance]. These phonological principles make opposing predictions for the developmental order of acquisition of sC clusters. Two hypothetical learning paths are predicted based on preferences of syllable well-formedness by each markedness constraint:

a. Markedness based on sonority: s+ approximant > s+ nasal > s+ obstruent

b. Markedness based on OCP[cont]: s+ obstruent/ s+ nasal > s+ approximant

Data were examined in terms of error rates in the form of frequency of modification, via prothesis, to determine which phonological principle most accurately predicted order of acquisition. Error patterns overwhelmingly showed the /s/ + sonorant clusters were modified more often than /s/+obstruents. Contradictory to predictions made by sonority, the data showed higher error rates for the clusters abiding by sonority. However, it is important to consider that more than half of the modified clusters also underwent sibilant voicing assimilation due to native language transfer. Since a voiced cluster is considered more marked (Greenberg, 1965), error rates were likely mediated by effects of voicing.

While it cannot be concluded definitively to what extent sC cluster productions are caused by markedness effects from sonority or OCP[cont] based on prothesis rates alone, using a finer grained acoustic analysis, through durational measurements of prothetic vowels, provided insight into cluster markedness. Although /sn/ had the highest frequency of prothesis, the longest prothetic vowels were found before the /sl/ cluster. When taken as an indicator of production difficulty, the prothetic vowel duration of /sl/ renditions indicate that this cluster is the most marked. /sl/ is the only cluster in violation of the OCP[cont], which may have contributed to markedness as indicated by vowel length. Although prothetic vowel durations and prothesis rates do not align perfectly, given more data points of prothesis, the two are likely to enter alignment.

In addition to the two markedness principles based on sonority and continuance, this study took into consideration other factors such as cluster frequency and environment preceding target sC clusters. Rates of prothesis closely followed the trajectories laid out by phonotactic probabilities of clusters and the cross-linguistic implicational hierarchy of sC clusters developed by Goad (2011). Since these trajectories overlap in their orders of acquisition (/s/+sonorant > /s/ + obstruent), these results warrant further investigation to tease apart the factors at play in influencing L2 production of initial sC clusters.

Limitations

This study sought to examine the production of English initial sC clusters by native Brazilian Portuguese speakers. The method employed was two forms of a self-paced reading task: one in which participants read sentences containing target words in controlled environments and a second in which target words were elicited by pictures. Future studies would benefit from the inclusion of different elicitation methods that do not rely exclusively on structured and highly controlled reading and picture elicitation tasks, such as a more conversational style elicitation. Also beneficial may be a delayed repetition task, where participants listen to sentences and then repeat them. Lastly, productions of nonce words may be included to tease apart possible confounding effects of lexical frequency.

Finally, the number of participants in this study was eighteen. Fourteen of the original 32 participants were excluded due to a lack of error in productions. The successful renditions of these participants may have been due to their advanced level of proficiency in English. Future studies may include a higher number of low proficiency participants to ensure sufficient error rates and provide a higher power when performing statistical analyses.

References

Abrahamsson, N. (1999). Vowel epenthesis of /sC(C)/ onsets in Spanish/Swedish interphonology: A longitudinal case study. *Language Learning* 49 (3), 473-508.

Anderson, J. (1987). The markedness differential hypothesis and syllable structure difficulty. In G. Ioup & S. Weinberger (Eds.), *Interlanguage Phonology: The acquisition of a second language sound system* (pp. 279-291). Cambridge, MA: Newbury House.

Baayen, R., Piepenbrock, R. & Gulikers, L. CELEX2 LDC96L14. Web Download. Philadelphia: Linguistic Data Consortium, 1995.

Barlow, J. & Dinnsen, D. (1998). Asymmetrical Cluster Development in a Disordered System. *Language Acquisition* 7: 1-49

Broselow, E. & Finer, D. (1991). Parameter setting in second language phonology and syntax. *Second Language Research* 7: 35-39.

Brysbaert, M. & New, B. (2009). Moving beyond Kucera and Francis: A Critical Evaluation of Current Word Frequency Norms and the Introduction of a New and Improved Word Frequency Measure for American English. Behavior Research Methods, 41 (4), 977-990.

Cardoso, W. (2008). The Development of #sC Onset Clusters in Interlanguage: Markedness vs. Frequency Effects. In R. Slabakova, J. Rothman, P. kempchinsky, & E. Gavruseva (Eds.), *Proceedings of the Generative Approaches to Second Language Acquisition (GASLA 9)* (pp. 15-29). Somerville, MA: Cascadilla Proceedings Project

Cardoso, Waclir, & Liakin. (2009). When input frequency patterns fail to drive learning: Evidence from Brazilian Portuguese English. In B. Baptisa, A. Rauber, and M. Watkins (eds.), *Recent Research in Second Language in Second Language Phonetics/Phonology: Perception and Production*, pp.174-202. Newcastle Upon Tyne: Cambridge Scholars.

Carlisle, R. (1991a). The influence of syllable structure universals on the variability of interlanguage phonology. *Applied Linguistics, 12 (1),* 76-95.

Carlisle, R. (1991b). The influence of environment on vowel epenthesis in Spanish/English interphonology. In A. D. Volpe (Ed.), *The Seventeenth LACUS Forum 1990* (pp. 134-145). Lake Bluff, IL: Lingusitic Association of Canada and the United States

Carlisle, R. (1992). Environment and markedness as interacting constraints on vowel epenthesis. In J. Leather, & A. Jarnes (Eds.), *New sounds* 92 (pp. 64-75). Amsterdam: University of Amsterdam Press.

Carlisle, R. (1994). Markedness and environment as internal constraints on the variability of interlanguage phonology. In M. Yavas (Ed.), *First and Second Language Phonology* (pp. 223-249). San Diego: Singular Publishing Company.

Carlisle, R. (1997). The modification of onsets in a markedness relationship: Testing the interlanguage structural conformity hypothesis. *Language Learning*, *47*, 327-361.

Carlisle, R. (1998). The acquisition of onsets in a markedness relationship: A longitudinal study. *Studies in Second Language Acquisition*, 20, 245-260.

Carlisle, R. (2002). Syllable structure universals and second language acquisition. *International Journal of English Studies* 1(1), 1-19.

Carlisle, R. (2006). The sonority cycle and the acquisition of complex onsets. In B. Baptista & M. Watkins (Eds.), *English with a Latin Beat: Studies in Portuguese/Spanish–English Interphonology* (pp. 105-137). Amsterdam/ Philadelphia: John Benjamin Publishing.

Carlisle, R. (2016). Environmental Markedness in Portuguese-English Contact. In M. Yavaş, M. Kehoe, & W. Cardoso (Eds.), *Romance-Germanic Bilingual Phonology* (ppg. 157-180). United Kingdom: Equinox eBooks Publishing.

Clements, G. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. Beckmans (Eds.), *Papers in Laboratory Phonology I: Between the grammar and physics of speech* (pp. 283-333). Cambridge: CUP.

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

Eckman, F. (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, F. & Iverson, G. (1993). Sonority and markedness among the onset clusters in the interlanguage of ESL learners. *Second Language Research,* 9, 234-252.

Enochson, Kelly. (2014). The effect of continuance on the L2 production of onset clusters. In Chia-Ying Chu, et al. (Eds.) *Selected proceedings of the 5th Conference on Generative Approaches to Language Acquisition North America (GALANA 2012)* Somerville, MA: Cascadilla Proceedings

Escartin, C. (2005). *The acquisition of English sC onset clusters by Mexican Spanish EFL learners.* Unpublished master's thesis, Concordia University, Montréal.

Ferreira Neto, W. (2001). *Introdução à Fonologia da Língua Portuguesa.* São Paulo: Hedra.

Greenberg, J. (1965). Some generalizations concerning initial and final consonant clusters. *Linguistics*, 18, 5-34.

Goad, H. (2011). The representation of sC clusters. In M. vanOostendorp, C. Ewen, E. Hume & K. Rice (Eds.) *The Blackwell Companion to Phonology*, 898-923. Oxford: Wiley-Blackwell.

Hogg, R. & McCully, C.B. (1987). *Metrical Phonology: a coursebook.* New York: Cambridge University Press.

Hussain, A. (2014) "Acoustic Investigation of Production of Clusters by Saudi Second Language Learners of English". *FIU Electronic Theses and Dissertations*. 1235. http://digitalcommons.fiu.edu/etd/1235

Major, R. C. (1996). Markedness in second language acquisition of consonant clusters. In r. Bayley & D. R. Preston (Eds.), *Second language Acquisition and Linguistic Variation* (ppg. 71- 96). Amersterdam: John Benjamins

Morelli, F. (2003). The relative harmony of /s+stop/ onsets: Obstruent clusters and the sonority sequencing principle. In C. Fery & R. van de Vijver (Eds.), *The Syllable in Optimality Theory* (pp. 356-371). Cambridge: CUP.

Rauber, A. (2006). Production of English initial /s/-clusters by speakers of Brazilian Portuguese and Argentine Spanish. In B. Baptista & M. Watkins (Eds.), *English with a Latin Beat:*

Studies in Portuguese/Spanish–English Interphonology (pp. 154-167). Amsterdam/ Philadelphia: John Benjamin Publishing.

Rebello, J. & Baptista, B. (2006). The influence of voicing on the production of initial /s/-clusters by Brazilian learners. In B. Baptista & M. Watkins (Eds.), *English with a Latin Beat: Studies in Portuguese/Spanish–English Interphonology* (pp. 139-153). Amsterdam/ Philadelphia: John Benjamin Publishing.

Ribas, V. (2004). Sobre a aquisição de onset complex. In Regina Lamprecht (Eds.), *Introdução à Estudos de Fonologia do Português Brasileiro*, 151-166. São Paulo: ARTMED

Vennemann, T. (1988). *Preference Laws for Syllable Structure and the Explanation of Sound Change.* Berlin: Mouton de Gruyter

Weinberger, S. (1987). The influence of linguistic context on syllable simplification. In G. Loup & S. Weinberger (Eds.), *Interlanguage Phonology. The acquisition of a second language sound system* (pp. 401-417). Rowley, MA: Newbury House

Yavaş, M. (2010). Acquisition of /s/-clusters in Spanish-English bilingual children with phonological disorders. *Clinical Linguistics & Phonetics*. 24 (3): 188-198.

Yavaş, M. (2011). Patterns of cluster reduction in the acquisition of #sC onsets: Are bilinguals different from monolinguals? *Journal of Multilingual Communication Disorders.* 25 (11-12): 981-988.

Yavaş, M. (2013). What explains the reductions in /s/-clusters: Sonority or [continuant]?, *Clinical Linguistics & Phonetics*, 27:6-7, 394-403.

Yavaş, M. and Altan, A. (2016). Vowel epenthesis in the acquisition of English /s/-clusters by Turkish speakers. In M. Guven, D. Akar, B. Öztürk & M.Kelepir (Eds), *Exploring the Turkish Linguistic Landscape: Essays in honor of Eser Erguvanli-Taylan.* (pp.3-20)

Yavaş, M. & Someillan, M. (2005). Patterns of Acquisition of /s/-clusters in Spanish-English Bilinguals. *Journal of Multilingual Communication Disorders.* 3(1): 50-55.

Yavaş, M., Ben-David, A., Gerrits, E., Kristoffersen, K., Simonsen, H. (2008). Sonority and cross-linguistic acquisition of initial s-clusters. *Clinical Linguistics & Phonetics.* 22(6): 421-441.

Yip, M. (1988). The Obligatory Contour Principle and Phonological Rules: A Loss of Identity. Linguistic Inquiry. 19: 65-100.

APPENDIX

APPENDIX A- Complete word list containing target sC clusters

/sp/	/st/	/sk/	/sm/	/sn/	/sl/
speak	star	sky	smoke	snow	sleep
spider	stamp	skirt	small	snake	sleeve
spoon	stop	school	smile	sneeze	slow
space	store	scarf	smell	snore	slide
spook	stump	sketch	smog	snack	slug
spade	stork	scout	smock	snap	slot

APPENDIX B- Sentence stimuli for formal reading task

Preceding Env.	Block A	Preceding Env.	Block B
	/Sp/		**/Sp/**
t	They will not speak about it.	d	My friend speaks Russian
t	He saw a giant spider in the car.	d	She said spiders are dangerous.
d	They put a fork and spoon on the table.	t	Please put spoons on the table.
d	I hope you can find space in the closet.	t	They want space in the room.
t	They don't spook easily.	d	A red spade was in the deck.
	/St/		**/St/**
t	He drew a white star on the paper.	d	The teacher gave a gold star for effort.
t	He put stamps on the letter.	d	He collected stamps as a hobby.
d	The old stop sign was replaced.	t	They went to that store yesterday.
d	This is a good store for shopping.	t	He didn't stop at the light.
d	He said storks don't exist	t	He wrote stump on the paper.
	/Sk/		**/Sk/**
t	There is a light sky, not a dark one.	d	He said sky, not cloud.
t	The girls bought skirts at the mall.	d	Her red skirt is new.
d	Harvard is a hard school to get into.	t	They want school to end.
d	She found her old scarf today.	t	They put scarves on in the winter.
t	They don't sketch anymore.	d	He is a good scout that learns quick./scope
	/Sm/		**/Sm/**

t	You can not smoke in this room.	d	There was fire and smoke that day.
t	They want small dogs.	d	He said small cars are better.
d	He said smile for the camera.	t	She didn't smile for the picture.
d	There is a good smell in the kitchen.	t	There was a sweet smell from the cookies.
d	The factory produced record smog that year.	t	She wore a paint smock over her clothes.
	/Sn/		**/Sn/**
t	There was a light snow fall this year.	d	The cold snow froze their hands.
t	The quiet snake surprised us.	d	She saw a red snake today.
d	My dad snores at night.	t	They don't snore anymore.
d	His friend sneezed loudly.	t	Matt sneezes when he is sick.
t	The mom brought snacks for her kids.	d	I heard a loud snap outside.
	/Sl/		**/Sl/**
t	The problem is not sleep, it's time.	d	You should sleep if you are tired.
t	He prefers short sleeves, not long ones.	d	The shirt had sleeves with holes.
d	He said slow and steady wins the race.	t	They write slow, not fast.
d	The old slide at the park broke.	t	There is a great slide outside.
t	The right slot was full.	d	He found slugs in the grass.

APPENDIX C- Sentence stimuli for picture based reading task

Preceding Env.	**Cluster Type**
	/Sp/
t	He doesn't speak French but he is learning.
t	They hate spiders and bugs.
d	The polished spoon is clean.
d	He loved space as a child.
	/St/

t	There is a bright star in the picture.
t	He bought stamps at the post office.
d	He said stop, not go.
d	They saw an old store on the corner.
	/Sk/
t	The bright sky is blue.
t	There is a white skirt in the picture.
d	The kids go to a good school to learn.
d	She made a good scarf for her friend.
	/Sm/
t	The volcano made white smoke for days.
t	This apple is not small, its big.
d	He said smile and be happy.
d	There is a bad smell from the trash.
	/Sn/
t	The white snow covered the road.
t	There is a giant snake in the picture.
d	The husband snores and the wife hates it.
d	The boy was sick and sneezed all day.
	/Sl/
t	The cat sleeps on the table.
t	The shirt has a short sleeve and a collar.
d	The turtle walks quiet and slow all day.
d	The red slide is at the playground.

APPENDIX D- Complete language background questionnaire

Questionário histórico-linguístico

1. Idade: ____________ 2. Sexo: Masculino / Feminino

3. Local de nascimento (cidade, estado): ________________________

4. Nível de escolaridade (marque o mais alto grau obtido ou o mais alto nível escolar frequentado, mesmo que ainda incompleto):

___ ensino fundamental

___ ensino médio

___ graduação

___ pós-graduação

___ mestrado

___ doutorado

Outro (especifique):_____________________

5. Se você fala outras línguas, por favor liste as línguas em ordem de proficiência (a mais proficiente primeiro, não incluindo português)

Línguas

6. Quando você começou a estudar inglês (excluindo os cursos no ensino fundamental e médio)?

Se você só estudou no ensino fundamental e médio, marque um ‘X’.

Idade ________

7. Como você aprendeu inglês (ou está aprendendo)? Marque tantas opções quantas forem necessárias.

- Curso de línguas
  - Qual curso?_________ Nivel? ________ Quando foi? ________
- Em casa
- Escola
- Sozinho
- Outro _____________

8. Caso você já tenha realizado algum teste de proficiência em sua língua não-nativa: (por exemplo, TOEFL) por favor, indique o nome do teste, a língua avaliada e a pontuação obtida (se você não se lembra exatamente da pontuação, escreva um número aproximado. Se você só se lembra do percentual, escreva-o na coluna do escore real).

Teste	Língua	Pontuação real	Pontuação aproximada

9. Você já passou mais do que **três meses** em um país de lingua inglesa? Se sim, informe o número de meses que você passou.

Meses:_________

10. Circule em uma escala de 1 a 6, seu nível de proficiência nas línguas que sabe (não incluindo o português).

1=muito baixo 2 = baixo 3 = razoável 4=bom 5=muito bom 6=proficiente

Lingua:_________

Leitura:	1	2	3	4	5	6
Escrita:	1	2	3	4	5	6

Compreensão auditiva: 1 2 3 4 5 6

Fala: 1 2 3 4 5 6

Lingua:__________

Leitura: 1 2 3 4 5 6

Escrita: 1 2 3 4 5 6

Compreensão auditiva: 1 2 3 4 5 6

Fala: 1 2 3 4 5 6

10. Você fala uma outra lingua que não é português na sua casa? ____________

Com amigos? ____________

No trabalho/escola? ____________

11. Estime a porcentagem do tempo que você usa cada língua diariamente (o total deve ser 100%):

	% do tempo
Língua 1	
Língua 2	
Língua 3	
Língua 4	

13. Caso haja alguma outra informação que você julgue importante sobre o aprendizado ou o uso das suas línguas, por favor, escreva abaixo:

APPENDIX E- Complete language background questionnaire

Word	FREQcount	SUBTLwf
stop	36071	707.27
school	16989	333.12
sleep	11625	227.94
speak	9546	187.18
small	6373	124.96
smell	4240	83.14
store	4178	81.92
star	4149	81.35
slow	3877	76.02
space	3369	66.06
smoke	3337	65.43
smile	2958	58
sky	2285	44.8
snow	1599	31.35
snake	1140	22.35
slide	909	17.82
snap	887	17.39
scout	657	12.88
spider	515	10.1
skirt	508	9.96
snack	466	9.14
spoon	388	7.61
sleeve	286	5.61
slot	280	5.49
stamps	267	5.24
slug	253	4.96
sketch	252	4.94
scarf	239	4.69
sneeze	150	2.94
spook	140	2.75
stump	125	2.45
spade	118	2.31
snore	82	1.61
smog	55	1.08
smock	25	0.49
storks	6	0.12