We are pleased to present the ninth volume of the George Mason Working Papers in Linguistics. The papers in this volume are works in progress submitted by George Mason University students and reviewed by the editors of GMU Working Papers in Linguistics.

Special thanks are due to Steven Weinberger and Tel Monks for their assistance with this volume and for their valuable guidance and support during the writing of the works herein.

We would also like to thank all students who submitted to this volume for their general interest in the field of linguistics and their willingness to share what they have learned with others. Finally, we are grateful to the College of Humanities and Social Sciences for the funding of this project.

Editors

LaBrea Bass
Claudia Calissie
Katie Shoaf

Chris Alderman

Glossomimesis

The Swedish Chef is a gibberish-spouting hand puppet from Jim Henson's *The Muppet Show* who speaks a nonsense language. As his name suggests, he is supposed to be Swedish, and Jim Henson intended his "language" to resemble the perceived lilting, sing-song quality of Swedish. Ask a Swede, however, and he or she will claim it sounds like Norwegian, but that does not stop many from asking Swedes to try and "translate" some of the Chef's gibberish (Stahl, 2012). Yet that would be impossible for the most fluent Swede or Norwegian, since most of the wacky cook's discourse are improvised neologisms that conform to what the actor believes is how the Swedish language sounds.

Language imitation is a pervasive element of human speech. All humans with the capacity for speech do it out of spite, for the sake of laughter, or to demonstrate their perception of a language's phonetic character. As a testament to this pervasiveness, the word "barbarian," though accepted as onomatopoeic, was coined by the Ancient Greeks to mimic the phonetic nature of other languages, which they also believed resembled the bleating of sheep (Bredin, 1996). Thus, "barbarian" is actually a conflation of language mimicry and onomatopoeia.

English has worldwide cultural dominance, and many people, yet not everyone in non-anglophone countries, can speak it. With the ubiquity and popularity of English media, the English language itself is often the subject of imitation. One variety the French call *yaourter* literally means "to yogurt" or "to attempt to speak or sing in a foreign language that they don't know very well...often...mishear[ing] and misinterpret[ing] the word or lyrics and substitut[ing] them with familiar words." A further description is "an imitation of a very nasal language, kind of like a baby crying...mostly imitating the 'cowboy' accent" (Liberman, 2009). I find the "nasal" description odd, given French's phonemic possession of nasals and English's phonemic lack thereof.

Yaourter does not satisfy the definition, however, since it refers to something more akin to malapropism. English has words that describe certain facets of nonsensical foreign language imitation, such as *glossolalia* (which I will treat later on) and even *onomatopoeia*, but not a portmanteau referring to the whole spectrum of language imitation. I wish to posit a term for it: *glossomimesis*. The word consists of two compounded parts--*glossa*, which is Greek for "language," and *mimesis*, Greek for "imitation" (and *-mimetic* is the adjective derived therefrom) (DeMoss, 2001; *Oxford Greek Dictionary*).

Usage

In English, glossomimesis often figures into some improvisational comedy. It is the theme of some skits titled *foreign film dub*, which were often figured into the American television show *Whose Line is it Anyway?* (12Medbe, 2008). British comedian Catherine Tate has a hilarious, and rather illuminating, skit called *Catherine Tate Translator*, whereby she "translates" a coworker's words during a meeting into six "languages"--French, Swahili, Castilian Spanish, Italian, and Chinese--all of which are stereotyped imitations (golocalise, 2009).

Purpose

That leads to the premise of this study--what phonemes do speakers use to comprise their glossomimetic speech? In the case of the Swedish Chef, not only do individual segments contribute to the accuracy of the utterances, but the prosody as well, since in the case of Swedish and Norwegian, prosody constitutes the primary differences between the two languages. Every language has its own phonetic peculiarities, and practitioners of glossomimesis will focus on the phonetic features that are the most salient to them. For that reason, one could not expect for two glossomimetic impressions of a particular language to sound completely alike or possess the same segment inventory.

The scope of this essay will be limited to the segmental phonology of glossomimetic Chinese, since it is my own observation that most people can mimic Chinese in some form with very little exposure to it. Though prosody in Chinese is of paramount communicative importance to Chinese since tonal distinctions between words serve to differentiate lexica and it is important to glossomimetic Chinese production (Comrie, 1990), I will exclude it for the purposes of this essay and narrow my focus to segmental and phonotactic consistency.

There are numerous dialects of "Chinese," and most share phonological features. Chinese is a collective term for a number of dialects spoken in China and is actually part of a larger East Asian sprachbund. Membership in this grouping makes most of the mainland Southeast Asian languages (such as Vietnamese, Lao, and Thai, among others) sound quite similar since all of the languages possess simple syllable structure, are tonal, and most words are monosyllabic (Comrie, 1990). Therefore, a speaker could base his or her glossomimetic "Chinese" on any number of Southeast Asian languages and attribute it to Chinese proper. I will narrow the focus to one dialect, Mandarin, also known as Standard Chinese (Duanmu, 2000).

I would like to make a brief note on nomenclature used in this paper. I will use *Chinese* when referring to the speakers' generalized concept of the target language. I will use the phrase *glossomimetic Chinese* when referring to the speakers' imitations. And I will use *Mandarin* to refer to the spoken language itself.

Pre-existing Literature

Little treatment of glossomimesis exists in scholarly literature. Samarin mentions it as a subset of glossolalia, from which glossomimesis differs by having a ludic function, mentioning Charlie Chaplin's impression of Hitler (Samarin, quoted in Mueller, 1981). Motley mentions how some glossolalia sounds like Spanish or Russian, though the utterances had no target impersonation.

Despite the paucity of academic literature regarding glossomimesis, it is the subject of various internet blogs and discussions. One such blog is the University of Pennsylvania's *Languagelog*, where I came across *yaourter* detailed above. Most Internet sites that discuss glossomimesis try and assign it a name, but usually only succeed in describing it to some degree. The discussions will either begin with or move into how speakers of one language perceive other languages.

Comparison of Glossomimesis and Glossolalia

I rely on phonological descriptions of glossolalia to inform my analysis because glossomimetic behavior is very similar to glossolalia, except glossomimesis has a deliberate, targeted phonology comprising the characteristics of the target language folded into the glossomimetic performance, whereas the phones of any given glossolalic utterance are seemingly random and generated from a subset of the speaker's native inventory. Glossomimesis and glossolalia also differ in their usage. Glossolalia is a religious and trance-induced linguistic act (Goodman, 1969), and though the literature does not mention a target audience, one could assume a glossolalist's only "audience" would be his or her God and those around him or her. The audience for glossomimesis are individuals from whom the speaker is trying to draw a reaction.

Another major difference between glossolalia and glossomimesis is the use of a speaker’s pre-existing familiarity or exposure to other languages. Glossolalia does not require any such knowledge. As stated before, a speaker must identify salient features of a given language, particularly those that are absent from his or her own language, and make productive mimetic use of that to achieve his or her intent. For instance, an utterance containing [x] or extensive nasalization - the former absent entirely from English and the latter a marginal allophonic result - would figure prominently into glossomimetic representations of Arabic and French, respectively. An utterance such as [bukalakabadawada] would not come across as Arabic, French, or even Mandarin. Thus, without prior exposure to the target language, a glossomimesist would have a difficult time producing interpretable glossomimetic speech. As I discuss further on in the sections describing individual utterances, more general glossomimetic renditions of Chinese (that is, renditions where there is little phonological resemblance to spoken Chinese varieties), as well as more nuanced representations, betray the speakers’ knowledge or exposure to the language.

Mandarin Phonology

In order to evaluate the phonological accuracy of the glossomimetic Chinese data to the spoken target language, a brief overview of Chinese phonology and phonotactics is required. My primary reference is Standard Chinese, also known as Mandarin (Duanmu, 2000), but I also include quick notes on other languages of Southeast Asia, including Thai, Vietnamese, and Burmese but only where their mention is germane to the discussion. I do this because of the close similarities between the languages and the fact that not many people outside of the Southeast Asian speaking area can distinguish them. I do not include Japanese, Korean, or any Austronesian languages such as languages of the Philippines or Indonesia.

Mandarin Chinese possesses the following consonant phonemes (Campbell & King, 2011; Duanmu, 2000):

stops: p pʰ t tʰ k kʰ
affricates: t͡s t͡sʰ t͡ɕ t͡ɕʰ t͡ʂ t͡ʂʰ
fricatives: f s ɕ ʂ ʐ x
nasals: m n ŋ
lateral: l
approximates: j ɻ w ɥ

Mandarin also possesses the following vowels (Campbell & King 2011[1]):

i|y ɨ u
ɛ ə o
a

The presence of the retroflex consonants gives Mandarin a rhotic quality akin to some varieties of English.

Very strict syllable structure characterizes Mandarin phonotactics. No dialect of Chinese allows consonant clusters (Comrie, 1990), and only vowels, diphthongs, triphthongs, and /n/ and /ŋ/ are permitted in coda position. The nasals are further restricted as to the vowels and diphthongs with which they may occur (Campbell & King, 2011). Thus, the maximum Mandarin syllable would be CVVV or CVVN. Other dialects, such as Cantonese, permit other consonant codas, as do Thai and Vietnamese (Comrie, 1990).

Another attribute of Mandarin, and the whole Southeast Asian linguistic group discussed above, with the sole exception of Khmer, is the propensity for single syllable words with distinctive tones. Mandarin has four tones and is the simplest of all the Chinese varieties (Cantonese has nine, and the Chaozhou dialect has eight basic tones, plus many more contour tones) (Comrie, 1990).

Data

I derived my glossomimetic data from open source media exclusively from *youtube.com.* On the website, I searched "imitation Chinese" and transcribed the speech of ten different speakers (which I will refer to as S1-S10) which were all recorded on video. All of the speakers' linguistic backgrounds, and all but S1's (Catherine Tate, who is British) nationality are unknown, since native language may have some effect. The context of some of the samples is unknown, though they appear to be reacting to elicitation by the person recording. S9 is trying to teach his audience "how to copy an Asian," and he provides imitations of not only Mandarin, but also other Southeast Asian languages.

[1] Duanmu claims there are only five vowels, [i u y ə a], with [o, e, ɤ] being allophones of the middle vowel (2000, p 45).

Rosie O'Donnell is imagining what a Chinese spokesperson would have to say about an incident on her television show *The View*.

Below, I transcribe the glossomimetic Chinese of S1-S9. As stated above, I exclude prosodic notation. However, as a convention, I will parse the samples into "words" when they appear to belong to a stress group to give a cursory indication of prosody and to break up the sound stream. In the sample transcriptions below, I briefly describe prime characteristics of each.

Here, S1, Catherin Tate, is part of a television comedy sketch acting as a translator: [niɔːŋː t͡ʃɔː aː i no ma jaː t͡ʃin to waː t͡ʃi kɔ wa jɑ mɔːŋ i toŋ bɑ wɑ jɔː]. All of these words contain phonemes that occur in English. /t͡ʃ/ does not occur in Mandarin. All of the words are phonotactically congruent with English and Chinese syllable structures; however the long vowels are not present phonemically in Mandarin or English.

S2's words were "...this is what Mandarin sounds like to me.": [ʂɚ wɑ ʂɪŋ hɚ hɚ hɑ ʂɪŋ hiɻ kwɑːŋ ɪɻ ʃɪŋ iɻ wɑ ʂɪ dɑ hɚŋ go hiɻ wɑ ɑɻ hɪŋ wɑ ʂɪ ɑ̰ ɑ̰ o jaʊ mɪŋ ʂɚ ɖɛ ʂɪɛ ɪn bi ɛi wɑ ʈʂɪɻ gɪŋ ɪɻ ɖʐɑɔ wɔɑ tʰɪŋ hɚ hɑ hɚ tʰɪŋ pʰɑŋ mɪŋ hɔː wɑ ʈʂɪɻ]. Once again, the linguistic background of the individual speakers is unknown, but I would surmise this man had either studied Mandarin or has had considerable exposure to it. The rhoticity of the entire statement was likely intended to imitate a feature of some Mandarin subdialects called *erhua*, which I will discuss below (Zhang, 2005).

S3 is from a short film called *Babble*, in which people are recorded speaking various glossomimetic languages: [dɑs duː ti ʃãː pũ/puŋ tãː jɑɔ ti ʃə fə tʌŋ tɔ jʌː ɑːː ʃi tau θi si tum ʃistə θãũ θãũ di ʒau]. First and foremost, the /s/ coda in the first word is immediately apparent, which violates Mandarin in that Mandarin does not possess a /s/ and, if it did, it would not allow it to occupy coda position. Another feature of this sample is the interdental fricative, present in Burmese, but not Mandarin (Campbell & King, 2011).

S4 is from the same film as above: [bɔŋ t͡ʃɔŋ pãuːŋ daŋːː kikaɹɔŋ]. All words have a velar nasal, which is present in every Southeast Asian language. Note, however, the final word is comprised of three syllables.

S5 is from the same film as above: [ɑkɑtɑkɑtɑːtɑːtɑtukud͡ʒo ʁu t͡ʃau]. This particular sample was difficult to parse, and the first word looks to be an example of generalized glossolalia. Note the voiced uvular fricative, which is not present in Mandarin.

S6 is from the same film as above: [hɪntəntənfʊnt͡ʃajoː]. This sample was very brief but continuous, thus making any comments on syllable structure tenuous. However, individual syllables are simple and [t͡ʃ] is lacking in Mandarin. The reduced vowels could be the result of the speaker's native language habits.

S7 is from the same film as above: [mɪ nɔ wɑɲa naː]...the remainder was unintelligible. [ɲ] does not occur phonemically in Mandarin.

S8 is from the same film as above: [nalala dalala haŋ hadza zoa wɔːŋə nadə daː]. It is impossible to tell if the [d] and [z] are in contact as a result of adjacent syllables, or [dz] is an onset. Either way, Mandarin does not possess [z], and [dz] only would only occur in an unstressed syllable, according to Duanmu (2000).

S9 is from a short film called "How to Speak a Fake Asian Language," in which the speaker mimics various Southeast Asian languages; the transcribed Mandarin section follows: [niː wɔɻ ʃɔɻ hɔɻ wɚ tɔɻ ʃɔɻ wɔwɔwɔɻʃɔ wɔɻ]. This is another rhotic version whereby almost all the words contain [ɻ]. However, the [ɔɻ] rhyme does not occur in Mandarin (Duanmu, 2011).

S10 is Rosie O'Donnell's imagined Chinese newscast of Danny Devito's drunken appearance on *The View* (DannyRebus, 2013): [t͡ʃɪŋ dɔŋ hu ga t͡ʃi d͡ʒɔŋ tɔŋ t͡ʃɪŋ t͡ʃɔŋ (Danny Devito) t͡ʃɪŋ t͡ʃɔŋ t͡ʃɔŋ t͡ʃɔŋ t͡ʃɔŋ (drunk *The View*) t͡ʃɪŋ t͡ʃɔŋ]. Rosie relies heavily on the word [t͡ʃɔŋ] and the phrase [t͡ʃɪŋ t͡ʃɔŋ]. Almost all the words end with the velar nasal.

Analysis

I will discuss briefly the segment inventory of the samples above and compare it with Mandarin, then I will move on to describe the samples' phonotactics and compare that as well with actual Mandarin. When describing the sounds, I will refer to them as *phones*, following a convention used by Michael Motley (Motley, 1981) to refer to the individual sounds comprising glossolalia. This I do since glossomimesis, like glossolalia, is a non-communicative pseudo-language and individual sounds do not carry the same functional load as a phoneme. They do bear a particular functional load, however, and I will discuss that later.

Consonant Inventory

Below is the condensed phone inventory of glossomimetic Chinese based on ten samples:

stops: p b p^h t d t^h ɖ k g
affricates: t͡ʃ t͡ʂ ɖʐ
fricatives: f θ s z ʒ ʂ ʁ h
nasals: m n ɲ ŋ
lateral: l
approximates: j ɻ w

Some of the phones are English, and some are Mandarin. /ɲ/ and /ʁ/ occur in neither. Note that [ʁ], [f], [ʒ], [ɲ], and [z] occur only once throughout all of the samples. The full retroflex series occurs only in S2, and the interdentals only occur in S3. Otherwise, the most widespread and repeated consonant sounds are [t͡ʃ], [w], [h], and [ŋ]. S2 and S9 used [ɻ] heavily. As stated above in the commentaries following the individual samples, [t͡ʃ] does not occur in Mandarin (though it does occur in many other Southeast Asian languages), nor does [h]. In Mandarin, there is an affricate series consisting of three lenis affricates--[t͡s], [t͡ɕ], and [ʈ͡ʂ]--and three aspirated--[t͡sʰ], [t͡ɕʰ], and [ʈ͡ʂʰ]--each of these occurs exclusively as an onset and is of relatively high frequency (Duanmu, 2011), which minimal exposure would immediately demonstrate. Thus, some of the segments that find their way into glossomimetic speech are those which possess the highest frequency in the target language. Assuming all of the speakers in the samples were either English speakers or had knowledge of English, these affricates would have been interpreted as [t͡ʃ], explaining its frequency as the nearest familiar sound available.

The most common coda in the glossomimetic samples is [ŋ], and it occurs in the majority of the samples. This is likely due to similar factors as the affricate--its sheer frequency in Mandarin, which is second only to [n] as coda (Duanmu, 2011). However, English speakers are familiar with the Chinese borrowings or adaptations yin yang, kung fu, ping pong, Hong Kong, Beijing, and names such as Mao Zedong and Yao Ming. All of these examples share the velar nasal and are often the first and only exposure to Chinese many people have.

A perceptive individual or somebody who has studied Mandarin or phonology may notice the rhotic nature of northern Mandarin. This is particularly evident in Beijing and other dialects. As stated in the commentaries, this phenomenon is known as *erhua* (Zhang, 2005). The retroflex approximant occurs as a coda in very few words, but this frequency increases when the influence of *erhua* is introduced. The repetitive usage of the retroflex approximant in two of the samples above illustrates segment choice by what could be termed as a "differential trait," or some segment that is or seems utterly alien to the mimicking speaker, even if that segment is present in the speaker's own phonology and only differs in its frequency or distribution. For example, although the alveolar flap is present as an allophone in English, thus severely restricting its distribution to the point that most English speakers are unaware of its existence in their language, when a word begins with [ɾ], it is almost instantaneously regarded as non-English. As another example, in the same *youtube* video where Catherine Tate "translates" English into different languages, she also mimics Castilian Spanish, which possesses an interdental fricative (spelled <c> or <z>) (Campbell & King, 2011). The extent of her imitation of Castilian Spanish consisted of repeated production exclusively of the syllable [εθ] (golocalise, 2009). English, of course, has this consonant, and it is one of the language's distinctive features, but either its distribution or frequency was peculiar enough to Catherine Tate that is became the staple of her imitation.

Vowel Inventory

Below is the condensed vowel inventory glossomimetic Chinese based on ten samples, all of which could be English vowels:

i ɪ ʊ/u
o
ε ə ʌ/ɔ
a ɑ

The vowel inventory is a near reflection of English with the tense-lax high vowel distinctions, and in that regard the vowel list is fairly unremarkable. The samples universally lack the mid-high front vowel. The following vowels occurred only once: [ʊ], [ʌ], and [ɛ]. I adhered to the standard convention of transcribing the high lax front vowel with the velar nasal, which was the most frequent position of the high lax front vowel, but it did occur in a word final position in S2 and S7. S2 was also the only sample containing [ɛ]. The vowel occurring in the most samples was [a], which is consistent with its dispersal across normal human languages and Mandarin itself (Duanmu, 2011), though it only occurred as part of a diphthong in some of the samples.

Syllable Structure

I took the liberty to base my analysis regarding glossomimetic syllable structure around my prosodic convention that stress designates a word boundary and words branch from left to right with the stress on the first syllable. I do this purely to extract phonotactic information, since without established word boundaries one cannot draw any conclusions regarding syllable structure.

Based on my parsing convention, the glossomimetic examples above exhibited a preference for CV syllable structure. CV syllables comprised roughly 54% of all the syllable types in the samples, CVC syllables 37%, V 6%, and VC 3%. Although Mandarin statistically prefers CVC (Duanmu, 2011), glossomimetic Chinese does not possess an irregular array of syllable~~-~~-type frequency. According to the *World Atlas of Language Structures,* glossomimetic Chinese falls within the range of moderately complex syllables and is in line with the majority of the languages of the world (Maddieson, 2011).

Conclusions

I would like to pursue further research and refine my analysis by accounting for some major variables. A speaker's preexisting knowledge of a targeted language is important, and further study would require a speaker to disclose his or her knowledge of the target language. A speaker's native language would also need to be accounted for since that would serve as a basis of comparison against the target language, and it would also establish some prosodic habits that could aid in parsing what would otherwise be a continuous utterance. Additionally, I would seek longer speech samples to establish if glossomimesis possesses language-like inventories and ratios as defined in Maddieson (1988).

This was a preliminary survey of the phonology of glossomimesis. I wished to analyze the means by which an individual mimicked, without actually communicating in, a foreign language. Glossomimesis' non-communicative nature is congruent with glossolalia, though its practice has one notable difference--glossomimesis has a target set of segments that the speaker must use to accurately mimic a language. The inventory of segments the speaker chooses to use will approach that of the target language. The individual segments the speaker will choose depends on the perceived frequency of the approximated segments in the target language and the segments the speaker identifies as alien to his or her own sound inventory. In conclusion, I refer to glossomimesis as exceptional language behavior, although its phonology is not.

Sources

12Medbe. "Whose Line: Foreign Film Dub." *youtube*. n.p., 28 November 2008. Web. 16 April 2013. < http://www.youtube.com/watch?v=fLJbGL3NZ6Y>.

Bredin, H. "Onomatopoeia as a Figure and a Linguistic Principle." *New Literary History* 27.3, Literary Subjects (Summer, 1996), pp. 555-569. Electronic.

Comrie, B. (ed). *The World's Major Languages*. New York: Oxford University Press, 1990. Print.

DannyRebus. "Rosie O'Donnell ching chong." *youtube*. n.p., 25 December 2010. Web. 28 April 2013. <http://www.youtube.com/watch?v=cbwPu_LuTZs>.

DeMoss, M. S. *Pocket Dictionary for the Study of New Testament Greek.* Downer's Grove, Illinois: InterVarsity Press, 2001. Print.

Duanmu, S. *The Phonology of Standard Chinese*. New York: Oxford University Press, 2000. Print.

Duanmu, S. (2011). Chinese syllable structure. In *The Blackwell Companion to Phonology, volume 5*, ed. by Marc Van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice, 2151-2777. Malden, MA & Oxford: Wiley-Blackwell. Retrieved from <http://www-personal.umich.edu/~duanmu/10ChineseSyllable.pdf>.

golocalise. "Catherine Tate Interpreter." *youtube*. n.p., 28 Nov. 2009. Web. 16 April 2013.

Goodman, F. D. "Phonetic Analysis of Glossolalia in Four Cultural Settings." *Journal for the Scientific Study of Religion* 8.2 (1969): 227-239. Electronic.

"Language." *The Oxford Greek Dictionary*. American Edition. 2000. Print.

Liberman, M. "Yaourter." *Languagelog*. University of Pennsylvania. 21 July 2009. Web. 16 April 2013. < http://languagelog.ldc.upenn.edu/nll/?p=1593>.

Maddieson, I. (2011). "Consonant Inventories." In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 1. Available online at <http://wals.info/chapter/12>. Accessed on 6 May 2013.

Maddieson, I. (2011). "Syllable Structure." In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 12. Available online at <http://wals.info/chapter/12>. Accessed on 23 April 2013.

Mueller, T. "Linguistic Analysis of Glossolalia: a Review Article." *Concordia Theological Quarterly* 45.3 (1981): 185-191. Electronic (.pdf).

Stahl, J. "What Does the Swedish Chef Sound Like to Swedes." *Browbeat: Slate's Cultural Blog*. The Slate Group, 28 Aug. 2012. Web. 16 Apr. 2013. <http://www.slate.com/blogs/browbeat/2012/08/28/what_is_the_swedish_chef_actually_saying_one_swede_translates_.html>.

thuncochin. "How to Speak a Fake Asian Language." *youtube*. n.p., 29 Aug. 2012. Web. 16 April 2013.

Zhang, Q. "A Chinese yuppie in Beijing: Phonological Variation and the Construction of a New Professional Identity." *Language in Society*, 34.3 (2005). Electronic (.pdf).

John Kinney

The L2 English production of [ð] in word-initial onset and intervocalic onset position – A pilot study

In this study I investigate the production of the English interdental fricative [ð] in word-initial onset and intervocalic onset position by Russian, Mongolian, Thai, and Amharic learners of English. Eight participants completed word and story reading tasks. The results show that all of these learners produce [ð] at a high rate in intervocalic onset (V.CV) position while often substituting [d] for [ð] in word-initial onset (.CV) position, illustrating one example of the universal process of spirantization, in this case stop to fricative spirantization: /d/ → [ð]. Data from the Russian, Mongolian, Thai and Amharic learners of English are argued to be exemplars of the emergence of the unmarked, a process in which L2 learners utilize a process that is underdetermined in the L1 or L2. The interlanguage data are analyzed and discussed within the framework of Optimality Theory. Within this framework, it is shown that these language learners rank a constraint prohibiting stops in intervocalic onset position higher than markedness constraints prohibiting interdental fricatives altogether.

Keywords: interdental fricative, L1 transfer, markedness, spirantization, intervocalic, differential substitution, Optimality Theory, emergence of the unmarked

1. Introduction

The central question this paper seeks to answer is whether second language learners of English (L2ers) have access to the universal process of /d/ → [ð] spirantization when acquiring the English segment [ð]. Spirantization, operable in many languages (Kirchner, 1998), is a type of lenition (or weakening) in which a stop becomes a fricative or approximate. Below are selected examples of stops becoming fricatives in Badimaya (Australian), Dahalo (Afro-Asiatic), Gajarati (Indo-European), Spanish (Indo-European) Basque (Basque), Lama (Niger-Congo), and Hausa (Afro-Asiatic) (Lewis, 2009, noting language families) as reported in Kirchner (1998).

(1) L1 Spirantization Examples

Badimaya	d, d^j → ð, ʒ / V___V	(p.7)
Dahalo	b, d → ß, ð / V___V	(p.7)
Gujarati	b^h, d^h, g^h → ß, ð, ɣ / V___V	(p.7)
Spanish	b, d, g → ß, ð, ɣ / non-initially, except after an [n] or [l]	(p.7)
Basque	k → ɣ / word-finally	(p. 8)
Hausa	b, d, g → w, r, w / coda position	(p. 9)

As seen in (1), spirantization can be applied to various place features (e.g., labial, coronal, velar) and in various syllable positions (e.g., intervocalic-onset, coda). Kirchner (1998) reviewed 272 lenition patterns and found that intervocalic onset position is the most preferred position for spirantization in L1 grammars. Based on this data, it can be assumed that, typologically, intervocalic onset position is the most common position for the process of spirantization. Thus, intervocalic onset position is an unmarked position for the feature +Continuant.

These typological facts about spirantization raise the question of whether L2ers' production rates of interdental fricatives in intervocalic position follow the universal pattern in L1 grammars. In order for spirantization to be illustrated in L2 English, [ð] would need to be produced less in word-initial onset than in intervocalic onset position. However, there is nothing in the target English input that would suggest that [ð] should be produced in intervocalic onset position and *not* in word-initial onset position.[1] Thus, if spirantization were illustrated in L2 English, its origin would be universal grammar or L1 transfer.

Maddieson (1984) reports that [ð] is present in only 7% of the languages in the UCLA Phonological Segment Inventory Database. On top of being a rare sound typologically, due to the marked place (+interdental) and manner (+fricative) features of [ð] (Ladafoged, 1993), the segment [ð] is difficult for L2 English learners to acquire. Since [ð] is marked segment with regards to typology and articulation, L2ers often substitute another segment for [ð]. Most researchers have found that L2ers substitute either [d] or [z] for [ð] (Kohmoto, 1965 (L1 Japanese L2ers)); (Weinreich, 1968 (L1 Russian L2ers)); (Michaels, 1973 (L1 Sinhalese L2ers)); (Altenberg &Vago, 1983 (L1 Hungarian L2ers)); (James, 1986 (L1 German L2ers)). Additionally, most L2 studies on [ð] (and [θ]) primarily analyze productions in onset and coda position (Weinberger, 1990); (Hancin-Bhatt, 1994); (Flege et al., 1995); (Lombardi, 2003); (Wester et al., 2007); (Rau et al., 2009), leaving the L2 production of [ð] in intervocalic position a relatively open area of research. In the domain of English L2 acquisition, the typological and articulatory markedness of [ð] allow L2 researchers to investigate the role of L1 transfer and language universals in an L2 grammar.

To determine if the universal process of spirantization is operable in an L2 English grammar, I investigate the production of [ð] in word-initial onset position and intervocalic onset position by Russian, Mongolian, Thai, and Amharic L2ers. Learners with these L1 backgrounds were chosen for two reasons: 1) because they lack [ð]; 2) they lack spirantization of the stop /d/.[2]

The data collected in this empirical study reveal that in word-initial onset position,

---

[1] In African-American Vernacular (AAVE), there is fortition of [ð] in word-initial onset position and [v] can be a substitute for [ð] in intervocalic onset position: they [ðej] → [dej]; mother [mʌðɚ] → [mʌvɚ] (Yavas, 2011, p. 65). Although exposure to AAVE might explain why L2ers substitute [d] for [ð] in word-initial onset position, it does not explain why [ð] is produced regularly in intervocalic onset position. With this said, however, since the participants of this study are exposed to a standard variety English in college ESL classes, it is doubtful that AAVE plays a role in their production of [ð].

[2] Stops in Amharic do undergo spirantization. However, spirantization in Amharic excludes the stop [d], so there is no /d/→ [ð] spirantization. It is argued here that lenition processes that affect other segments in these L1s will not be applied to stops (e.g. /d/) in the English L2. It might be expected that the Amharic learners would illustrate positive transfer of spirantization; however, the data from the Amharic speakers is similar to the data of the other participants.

[d] is substituted for [ð] more often than not. The data also show that [ð] is produced more, by a considerable margin, in intervocalic onset position than in word-initial onset position. I argue here that the data illustrate the universal process of spirantization.

I analyze the process of spirantization within the framework of Optimality Theory (OT) (Prince & Smolensky, 1993/2004). OT is a linguistic model that implements universal language constraints. The constraints hinge on linguistic typological universals and these linguistic universals' direct relationship to markedness. That is, the main force behind OT is the antagonistic relationship between markedness and faithfulness. Markedness constraints "compete" against faithfulness constraints, which seek faithfulness to the input, even if the input includes marked forms. The interaction between these two sets of constraints determines the surface forms of the phonological grammar. One aspect of phonological grammars that has been incorporated into OT is the so called *Emergence of the Unmarked* (McCarthy & Prince, 1994). This is a process in which an unmarked aspect of universal grammar, underdetermined by the L1 or the L2, "emerges" in the L2 grammar. In the case of the data here, spirantization, in which stops become fricatives in intervocalic onset position, is the emergent/unmarked process. I argue that this process is dictated by the positive markedness constraint SPIR (spirantization of stops in intervocalic position), and that this constraint competes against negative markedness constraints banning the highly marked segment [ð] (i.e., *ð) and segmental faithfulness constraints mandating that the input and output match.

2. Literature Review

2.1 Implicational Markedness and the Emergence of the Unmarked

Studies such as Colantoni and Steele (2007) focus on positional (or implicational) markedness. Colantoni and Steele (2007) also couch their study of the acquisition of the French [ʁ] in terms of positional markedness by supporting the claim that intervocalic position is the least marked position for fricatives. They analyze the acquisition of [ʁ], a voiced "dorsal" fricative (their description rather than "uvular"), in intermediate and advanced English learners of French with a word-reading and a passage-reading task. Calantoni and Steele evaluate several factors regarding the acquisition of [ʁ] such as voicing and manner, phonetic environment, and voicing by position. It is the last factor, voicing by position, which concerns us here. They argue that there is a developmental hierarchy with the acquisition of voiced fricatives. Generally speaking, [ʁ] in onset position is acquired before [ʁ] in coda position. However, the major claim made by Calantoni and Steele is that [ʁ] in intervocalic position (V.CV), which is an onset, is the position that will facilitate the highest production rate of [ʁ]. The results of their study partially support this hierarchy. Their advanced group produced [ʁ] in intervocalic position correctly 89% of the time versus word-onset position at 85% and coda position at 75%. However, their intermediate group's results do not fully support the hierarchy. This group had a 73% accuracy rate in coda position, specifically word-medial pre-consonantal position. However, their accuracy rate in word onset and intervocalic position was 52% and 69%, respectively.

In this paper, I argue that the L2 data reflect spirantization and that spirantization has emerged in the L2 grammar without positive evidence from the L2 or the L1, making this an example of the emergence of the unmarked (McCarthy & Prince, 1994). One example of the emergence of the unmarked is presented in Broselow et al. (1998), which investigated the simplification of English codas by Mandarin L2ers. Their argument is that universal markedness constraints prompt English coda simplification. Codas in Mandarin are only permitted to have glides and alveolar and velar nasals. Because of this L1 fact, Mandarin speakers have difficulty producing English obstruent codas which permit a wide variety of obstruents and consequently "repair" codas, for instance, with epenthesis, devoicing, or deletion in order to maintain their L1 licensed syllable structure. Broselow et al. (1998) argue that the OT constraints WD BIN (words must have two syllables) and NO OBS CODA (obstruents are disallowed in codas) conflict in the L1 grammar. The constraint WD BIN, a low ranked/violated constraint in L1 English and L1 Mandarin, *emerges* in the IL when the target L2 English violates the constraint NO OBS CODA, a constraint that is ranked high in L1 Mandarin. In this syllable context, the constraint WD BIN outranks NO OBS CODA and the faithfulness constraints relating to epenthesis (DEP – do not insert a segment not present in the input), deletion (MAX – do not delete a segment present in the input.), and devoicing (IDENT(VOI) — input and output must match). Like Broselow (1998), this study argues that L2 learners utilize a process (e.g spirantization) not found in the L1 or the L2, suggesting that a low ranked constraint has emerged.

2.2 Role of L1 Transfer[3]

Previous studies have claimed that L1 transfer explains why certain L2ers use the substitutes they do for [θ] and [ð]. Weinberger (1990), Hancin-Bhatt (1994), and Lombardi (2003) argue that features in the L1 determine whether [θ] and [ð] are substituted with [t] and [s] or [d] and [z]. Weinberger claims that a solution for differential substitution, at least between Japanese and Russian English language learners, is found by determining the most underspecified obstruent in each L1. Weinberger argues the segment

[3] Rau et al (2009) suggest that studies only focusing on L1 transfer have limited value in that they only account for the data in a general manner: e.g., Thai speakers substitute [θ] with [t], but Japanese speakers substitute [θ] with [s]. Rather, Rau et al. propose that the variable factors of frequency, markedness, speech style, and others are necessary to categorize and explain the variable L2er production data of interdental fricatives. According to the authors, participants had a higher accuracy rate with the formal tasks (word list, passage reading) and had a lower accuracy rate with the informal tasks (interview, story retelling). Participants' accuracy was also higher with high frequency words and lower with low frequency words. I refer the reader to the article for specifics as frequency and speech style are beyond the scope of this paper.

/s/ is underspecified in Japanese whereas /t/ is underspecified in Russian. Thus, Japanese speaking learners of English substitute [θ] with [s], and Russian speaking learners substitute [θ] with [t]. Lombardi (2003) departs from underspecification theory and posits that differential substitution can be explained by the transfer of L1 phonological rules. For instance, in L1 Japanese, an underlying /t/ becomes [ts] before [u]. This shows that manner features such as [stop] and [continuant] can be teased apart in the L1, prompting Japanese L2ers to be faithful to the [continuant] feature of [θ] with the substitute of [s]. Lombardi (2003) claims the opposite is true for L2ers with a Thai L1 background. In Thai, an underlying fricative in coda position is always realized as a stop. Thus, these L2ers substitute [t] for [θ] in L2 English.

Hancin-Bhatt (1994) also claims that L1 transfer is a determining factor in the L2 acquisition of English interdentals; however, unlike Weinberger and Lombardi, her study centers on the perception of interdentals rather than their production. Specifically, Hancin-Bhatt (1994) studies the L2 perception of the English interdentals by L2ers with Japanese, German, and Turkish L1 backgrounds with varying degrees of English proficiency. In this study, Hancin-Bhatt proposes the Feature Competition Model (FCM), which claims that prominent L1 features are transferred in L2 perception and learning. For instance, if the feature [continuant] is found to be a prominent distinctive feature in the L1 inventory, the feature will obscure other relevant L2 features in L2 perception. To test the FCM, participants listened to a total of 168 pseudowords with /θ, ð/. Six other segments, /f,v/, /t,d/, and /s,z/ were also tested since they can be mistaken for /θ, ð/ in perception. These segments were incorporated into three contexts: 1) word-initial 2) intervocalic and 3) word-final position.

Hancin-Bhatt (1994) predicts that the features that are prominent in the L1 will determine what segment will act as the substitutes for the English interdentals. For example, those from German and Turkish L1 backgrounds were predicted to use [s,z]. Following the predictions, the German group preferred /s,z/ as substitutes for the English interdentals; however, the Turkish group preferred to substitute English interdentals with stops in all three contexts. Hancin-Bhatt argues this discrepancy suggests that the FCM cannot predict prominence feature patterns of all L1s correctly.

Looking specifically at the production of the spirant/fricative [ð] and the spirantization process of /d/ → [ð] in monolingual native Spanish and German/Spanish bilinguals, Lleo and Rakow (2005) give evidence that L1 transfer plays a role in the production of [ð]. They showed that monolingual Spanish-speaking children had a high rate of spirantization from 1;3 years of age to 3;0 in words such as [deðo] "finger" (/dedo/ → [deðo]). In fact, the production rate of spirants never goes below 60% and reaches above 80% by 3;0 years of age. Lleo and Rakow found that the production of spirantization in German/Spanish bilingual children when speaking Spanish was lower than the monolingual Spanish speaking children. They conclude that L1 transfer of German is a factor here since German does not have the spirantization process of /d/ → [ð].

3. The Empirical Study

I investigate the interaction between the L2 English production of the segment [ð] and the syllable position [ð] is in by Russian, Mongolian, Thai, and Amharic L2ers. The syllable positions analyzed in this paper are word-initial onset position and intervocalic onset position. If L2ers make no distinction regarding the permissibility of the segment [ð] in various syllable positions, we might predict that the production rates of [ð] in word-initial onset, intervocalic onset, and coda syllable positions would be equally high or equally low for L2ers. In this case, it could be argued that L2ers are "blind" to syllable position when acquiring a new segment not in the L1 and do not illustrate spirantization. On the other hand, if production rates were inconsistent—say some high and some low—across various syllable positions, the implication in this case would be that certain syllable positions may facilitate the production of [ð] and that other positions may impede the production of [ð], suggesting that L2ers are *not* "blind" to syllable position. If the segment [ð] is produced more in intervocalic position than in word-initial onset position, spirantization would be illustrated.

Research Question

The empirical study presented here poses the following question: Do L2ers have access to the universal process of spirantization when acquiring the segment [ð]?

3.1 Linguistic Background for the Empirical Study

The following subsections present an overview of the pertinent linguistic facts regarding English, Mongolian, Russian, Thai, and Amharic. Since this study centers on the English segment [ð], the articulatory features of this segment are discussed as well. Additionally, since it is argued here that the universal process of spirantization is illustrated in the data, the following subsections take care to show that the segment [ð] and the process of spirantization is not present in the L1s of Mongolian, Thai, Russian, and Amharic.

a. English Interdental Fricatives

The interdental fricatives [θ] and [ð] (along with the palatal fricative [ʒ]) are some of the last segments to be acquired by native English speakers (Sander, 1961 qtd. in Ingram, 1989). In general, stops, nasals, laterals, and glides are mastered before fricatives by normal monolingual, English-speaking children and children with common phonological disorders. For example, children undergoing standard phonological development master stops such as [p], [b], [k], [g], and [d] by age 3;0 to 4;0; and children master fricatives such as [θ], [ʃ], [v], [ð], [s], and [z] by age 7;0 to 8;0 (Stoel-Gammon & Dunn, 1985, p. 31).

One major difference between fricatives and other classes of sounds such as stops is that fricatives have the [+continuant] feature. Based on a study with 90 children ranging in age from 40 months to 120 months, Singh and Frank (1979, p. 263) found that stops replace fricatives more than any other sound. Such a process is called "stopping" and is employed because stops are less complex than fricatives. Ladefoged (1993) describes the manner of articulation of a stop as a "complete closure of the articulator involved so that the airstream cannot escape through the mouth" (p. 8); he describes the manner of articulation of a fricative as a "close approximation of two articulators so that the airstream is partially obstructed and turbulent airflow is produced" (p. 10). Simply put, the "narrowed approximation" of fricatives demands great muscle control, whereas stops "involve very straight forward contact of the articulators" (Yavas, 1998, p. 138).

As noted, English has interdental fricatives and the stop [d]. However [d] and [ð] are separate phonemes in English, and there is no spirantization process of /d/ → [ð] with these segments in English. However, English does have other weakening processes. For instance, /t/ becomes the tap [ɾ] in intervocalic position when the proceeding syllable is stressed and the following syllable is unstressed as in /bɛtɚ/ → [bɛɾɚ] (Ladefoged, 1993, p. 92).[4]

b. Mongolian background

Svantesson et al. (2005) states that Mongolian has five fricatives: [x, χ, x□, s, ʃ]; Campbell (1995) notes that Mongolian has the aforementioned fricatives as well as [v]. There are differing accounts for stops as well. Svantesson et al. (2005) reports that Mongolian has [p, t, g, G] but no [d]; however, Campbell (1995) notes Mongolian has [d]. In the examples with [d] in Campbell (1995), [d] is only in word initial onset or coda position. Neither of these sources report that Modern Mongolian has spirantization. The sample words below confirm that stops do not become fricatives intervocalically.

(2) No intervocalic spirantization of stops in Mongolian

a. [vagon] 'coach' (Russian Loan word) (Svantesson et al 2005 p. 31)
b. [juutəŋ] 'hood' (Svantesson et al 2005 p. 60)
c. [igəm] 'collar bone' (Svantesson et al 2005 p. 82)
d. [saɢəm] 'buckwheat' (Svantesson et al 2005 p. 81)

However, Svantesson et al. (2005) does report that Old Mongolian spirantized velar and

---

[4] Additionally, English has a spirantization process that morphs stops into fricatives when derivational suffixes are added as in /ɪkspænd/ → [ɪkspænsɪv]; /dəsajd/ → [dəsajsɪv]; /pɹ̥mɪt/ → [pɹ̥mɪsɪv]. Although English does have spirantization, it is doubtful that L2 English learners would apply this particular spirantization process to the production of [ð] for two reasons: 1) the English spirantization process mentioned here is prompted by morphology; 2) there is sufficient evidence in the target language of English that [ð] is its own separate phone and not a result of allophonic variation. Thus, spirantization process of /d/ → [ð] / V __ V by L1 learners from L1 backgrounds lacking [ð] altogether suggests that this process is still underdetermined in the target L2 of English.

uvular stops in onset and intervocalic position, but there is no report that alveolar stops were weakened intervocalically. Additionally, as best I can determine, velar and uvular spirantization is not retained in Modern Mongolian.

Unfortunately, to my knowledge, there are no previous studies on Mongolian L2ers' production of interdental fricatives; thus, there is no precedent on what segments these L2ers will use as substitutes for [ð].

c. Thai Background

Thai has very few fricatives: [f, s, h] (Campbell, 1995; Tingsabadh & Abramson, 1999). Additionally, there is no report by Campbell (1995), Tingsabadh and Abramson (1999), or Kirchner (1999) of Thai having a spirantization process.

Smyth (2001) claims that L2 English learners from a Thai L1 background often substitute the stops /t, d/ or the fricative /s/ for English interdental fricatives in onset position and generally substitute /t/ for interdental fricatives in coda position.

(3) a. Thai L2er Onset Position Substitutes for Interdental Fricatives

/θ/ → [t, s]
/ð/ → [d, t, s]

b. Thai L2er coda position substitutes for interdental fricatives

/θ, ð/ → [t]

Additionally, in English loan words, Thai keeps intervocalic stops, as the examples in (4) show.

(4) No intervocalic spirantization of stops in English loan words in L1 Thai

a. /lejdi/ → [leedii] "lady"
b. /sajdɚ/ → [sajdəə] "cider"
c. /bejɾə/ → [beetaa] "beta" (From Kenstowicz & Suchato, 2006)

With these L2 pronunciation and loanword facts in mind, it appears that Thai speakers have no spirantization of stops in their L1.

d. Russian Background

The fricatives in Russian are [f, v, s, z, ʃ, ʒ, x] (Monk and Burak, 2001; Campbell, 1995). Timberlake (2004) notes that Russian also has [Ɣ], in addition to the previously mentioned fricatives. Russian stops do not undergo spirantization, but they do undergo palatalization, as examples from loanwords show below.

(5) Russian palatalization of intervocalic stops in loan words

a. German *Flugel* ‘wing’ /fly:gəl/ → [fi^{j}u:g^{j}ɛr]
b. Turkish *bituk* ‘hooligan’ /bityk/ → [bit^{j}uk]
c. French *bordure* ‘edge’ /bɔrdyr/ → [bɔrd^{j}yr]

Monk and Burak (2001) report that Russians learning English will substitute /s/ and /z/ rather than /t/ or /d/ for /θ/ and /ð/, respectively. With this in mind, we expect that Russian speakers will be faithful to the feature of +continuant when producing interdental fricatives~~,~~ but still err with regards to place of articulation. Thus, English L2 learners with a Russian L1 background are expected to perform equally to other L2 English learners who lack interdental fricatives in their L1.

e. Amharic Background

The fricatives in Amharic are [f, s, s’, z, ʃ, ʒ, h] (Hayward & Hayward, 1999). Campbell (1995) notes that Amharic also has [v]. The stops /b/ and /k/ do undergo spirantization in intervocalic position. The stop /b/ becomes [ß] in intervocalic position within words void of morphology and at word boundaries where affixes are added to the root.

(6) a. /ababa/ → [aßaßa] ‘a flower’ (“Archive Phonetics,” 1996)

b. /bet/ ‘house’ → [kəßet] ‘from the house’ (van Oostendorp, 2011, p. 2227)

The stop /k/ becomes [h] intervocalically when affixes are added to the root.

a. /nəka/ ‘touch’ → [yə.nahal] ‘he touches’ (Leslau, 1995, p. 17).

However, /k/ → [h] intervocalically does not appear to be an across-the-board process, as the examples below illustrate.

(7) a. /hakim/ → [hakim] ‘a doctor’ (“Archive Phonetics,” 1996)
b. /täkus/ → [täkus̥] ‘shooting’ (“Archive Phonetics,” 1996)

Geminate /kk/ or geminate /k’k’/ do undergo weakening to one segment.

(8) a. /täkkälä/ → [täkː^{h}älä] ‘he planted’ (“Archive Phonetics,” 1996)
b. /bäk’k’älä/ → [bäk’ːälä] ‘it grew’ (“Archive Phonetics,” 1996)

However, the segment /d/ does not undergo any spirantization (or weakening) in intervocalic position (Clavavin, 2010), which is an important distinction since the L2 data suggest that L2ers with an L1 Amharic background weaken /d/ in intervocalic position.

(9) No intervocalic spirantization of [d] in L1 Amharic

a. [awwədə] "perfume" (Colavin et al., 2010)

b. [bədəbbədə] "beat" (Colavin et al., 2010)

Although Amharic segments, such as /b/ and /k/, as shown in example (7) above, can undergo spirantization in intervocalic position, it is not certain that this L1 spirantization process would be generalized to the English segments /d/ and /ð/.

On the other hand, at minimum, it could be argued that these Amharic speakers would have slightly higher production rates of [ð] in intervocalic position since they have L1 spirantization. In fact, one Amharic speaker has 100% accuracy rate of [ð] in intervocalic position. However, this is not all that different from the Russian speakers who have 91-96% accuracy rate of [ð] in intervocalic position. Since the production rates are high for the Russian and Amharic speakers, I assume that the data of the participants of this study with an Amharic L1 background do not illustrate L1 transfer of spirantization. (Figure 3 in the results sections illustrates the results mentioned here.)

3.2 Methods for the Empirical Study

A. Participants

a. ***Non-native Speakers of English***: A total of eight nonnative speakers participated in the study. Three participants were native speakers of Russian, two were native speakers of Mongolian, two were native speakers of Amharic, and one was a native speaker of Thai. All were current or former ESL students at Northern Virginia Community College (NVCC). The current ESL students were high intermediate to advanced learners of English based on their current ESL courses.

Since all participants of the study were ESL students at NVCC, the following gives background on the ESL program for a better understanding of the participants' English proficiency. The credit-ESL program emphasizes academic writing and reading skills at all levels. Most credit ESL students in this study were in class 10-15 hours a week. There are four levels in this program: level 2, 3, 4, and 5. Level 5 is the exit level for ESL at NVCC that permits students to register for freshman English composition at the college.

The primary placement test NVCC uses is the English Accuplacer Test. This placement test has a reading skills, sentence meaning, and language use section, each with 20 questions ("Accuplacer"). The highest possible score on the test is 360. The following are Accuplacer test score ranges and proficiency rates: 225-274 = Beginning; 275-299 = intermediate; 300-324 = high intermediate; 325-349 = Advanced; 350 = placement into Comp 101. NVCC also asks students to write an essay after completing the Accuplacer. Their Accuplacer score, in combination with their essay test results, determine their ESL level.

The figure below summarizes pertinent background information for all participants. Each participant is given a code, and specific participants will be referred to by the code in the remainder of the paper. The first letter of the code matches the first letter of the speakers' L1: R = Russian; M = Mongolian; T = Thai; A = Amharic. In the figure below, the following abbreviations are used: AOE: Age of onset English; LOR = length of residence in the US; Acc = Accuplacer. All participants were female, so gender is not noted.

Figure (1): Participant background information

Code	Origin	L1	Age	AOE	LOR	Acc. Placement	Acc. date	Current English level	Daily % of English use
R1	Ukraine	Russian	39	30	8	Intermediate	2003	Completed ESL & English freshman comp	90
R2	Ukraine	Russian	24	21	3	Advanced	2011	ESL Lvl 5	50
R3	Ukraine	Russian	24	20	1	Advanced	2011	ESL Lvl 5	50
M1	Mongolia	Mongolian	38	36	1.6	Beginner	2010	ESL Lvl 4	70
M2	Mongolia	Mongolian	22	18	3	Beginner	2010	ESL Lvl 4	70
T1	Thailand	Thai	46	7	8	High Intermediate	2010	Completed ESL & English freshman comp	60
A1	Ethiopia	Amharic	36	9	8	Beginner	2008	ESL Lvl 4	50
A2	Ethiopia	Amharic	27	12	5	Beginner	2008	ESL Lvl 4	50
Median			**31.5**	**21**	**4**	--	--	--	**55**

Although all of participants have differing initial placement test scores, at the time of the study, they were all are intermediate or advanced learners of English. The overall median age of English onset for all participants is 31.5.

As figure (1) shows, the participants' daily use of English varies from 50% to 90%. The high usage of English in participant R1 can be explained by the fact the participant has a job requiring her to speak English. Participants who use English 50-60% of the day have friends and family members with their same L1 with whom they speak regularly. The median percentage that the participants use English daily is 55%.

*b. **Native English Speaking Control Group***: Two native English speakers participated in the study. Both native speakers were female and from Virginia. One speaker was 21 years old and the other was 65. Both speakers had studied French and Spanish in college but reported they were not proficient in either language.

B. Materials

All participants read a word list and a story. (See the appendix for the word list and story.) The word reading task had 100 words. Twelve words had [ð] as an onset (e.g., *there*) and thirteen words had [ð] intervocalically (e.g., *neither*), for a total of 25 instances of [ð]. Three words had [ð] as a coda (*bathe*, *breathe*, *teethe*), but due to apparent unfamiliarity with these words, these were excluded from the study. The other 72 words were distractors.

The story reading task had multiple paragraphs and [ð] was in the onset position thirty-three times and [ð] was in intervocalic position thirty-two times for a total of 65 instances of [ð]. There are a total of 340 words in the reading task. There is relative repetition of words such as *the*, *this*, *that* and *brother*, *mother*, *father*.

Altogether, there were a total of 90 instances of [ð]: 45 in word-initial onset position and 45 in intervocalic position.

C. Recording Procedures

Participants were told that they would have two reading tasks to complete and that they would be recorded. Participants were instructed to read the word list first and then read the story directly after in the same recording session. Before recording began, the investigator stated that the words in the word list reading task needed to be said in the carrier phrase: "Now say __________ again." The first two words of the reading list were modeled by the investigator and the investor said, "*Now say <u>big</u> again*; *Now say <u>four</u> again*." After the investigator explained the tasks, he asked participants if they had any questions. The only questions related to the carrier phrases. Most participants double-checked that they were to put each word in a phrase. The investigator did not model any part of the story reading task.

Participants were recorded with a SONY ICD-SX712 digital voice recorder. The recorder was set to interview mode and noise reduction mode. All recordings took place in a quiet room. The participants were alone in the room while completing the tasks.

D. Coding

The investigator reviewed each data sample three times. Most coding errors were due to an oversight on a particular word not being counted rather than a mistake with transcription.

All determiners such as *the*, *these*, and *that* were considered to be a word-initial onset environment for [ð] regardless of the preceding word. There were only three instances where a determiner occurred directly after a vowel sound: *by the, be the*, and *to their*. This coding choice was determined because there was no clear pattern with the production of [ð] in these environments. Most participants substituted [d] in each of these positions, matching their production in standard onset [ð] environments. This suggests most learners treat the [ð] in determiners as onsets.

A deletion of a determiner[5], schwa insertion before [ð] in an onset position, or an unnaturally long pause before [ð] in onset position were all considered a non-production of [ð] in onset position.

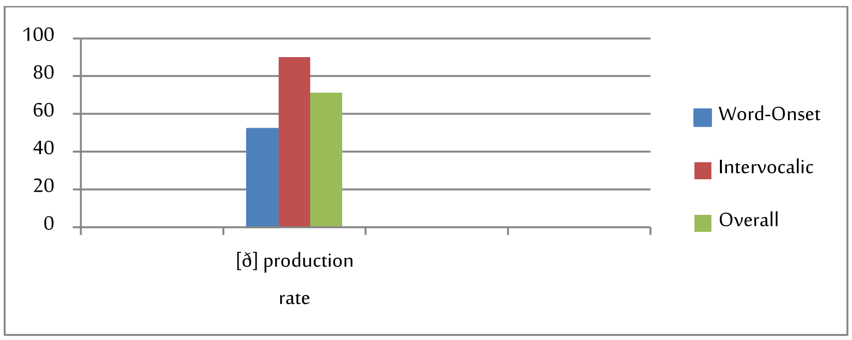
[5] The deletion of a determiner only occurred four times and it occurred in all L1 language groups. It appears this deletion was a reading error.

4. Results

The two native speaker controls performed at 100% accuracy in the production of [ð] on the word and story reading task. No other errors regarding other segments were made.

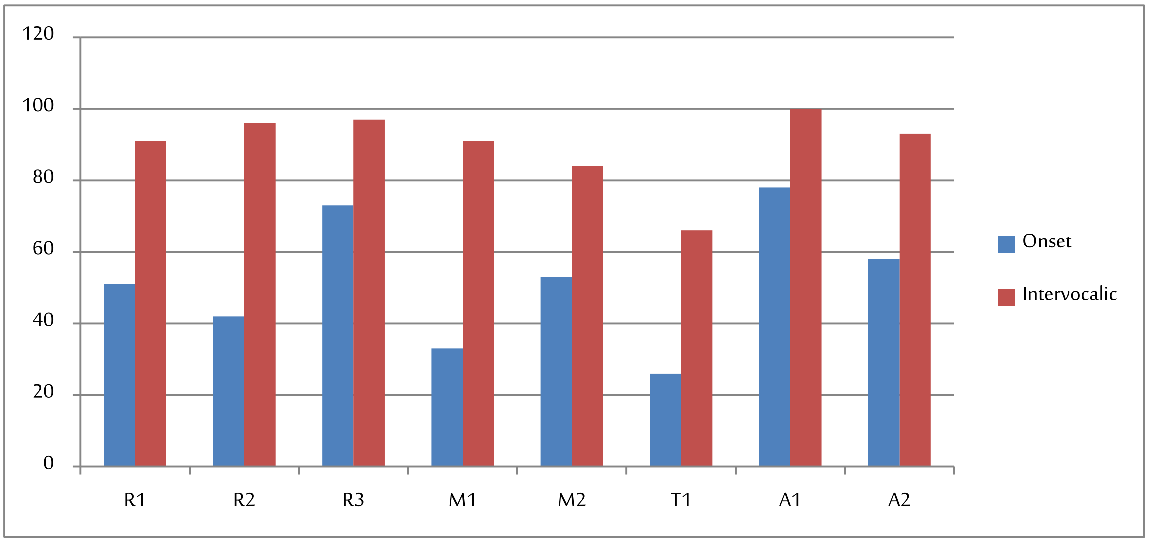
The overall accuracy rate of fricatives in onset and intervocalic position in both the word and story reading task by the L2ers was 71% (512/720). The production rate of [ð] in onset position was 52% (187/360), and the production rate of [ð] in intervocalic position was 90% (325/360). Figure (1) below gives an overview of the accuracy rate of [ð] in word-onset, intervocalic position, and the combined overall accuracy rate in both word-onset and intervocalic position.

Figure (2): Accuracy rates of [ð] in word onset and intervocalic position



These overall results show that [ð] is a relatively marked segment that is somewhat troublesome for L2 English learners, but that it is still produced more in intervocalic position than onset position. This is also true with each participant. In every case, the accuracy rate of [ð] in intervocalic position is higher than in word-initial onset position.

Figure (3): Participant Accuracy rates of [ð] in word onset and intervocalic position



If an accuracy rate criterion for the segment [ð] was set at 79%, as in other L2 studies (see (Anderson, 1978); (Cancino et al., 1975); (Eckman & Iverson, 1993); (Carlisle, 1998), only intervocalic position illustrates stable accuracy rate.

Figure (4), below, presents the median, average, and standard deviation for the accuracy rate of [ð] in onset and intervocalic position.

Figure (4): [ð] accuracy rate in onset and intervocalic position

	Onset	**Intervocalic**
Median	23.5	42
Average	20.7	38.4
Standard deviation	7.53	11.9

In addition, a two-tailed t-test was used to determine if the L2 learners' accuracy rate differs significantly in onset and intervocalic position.

Figure (5): Significant differences of accuracy rate of [ð] in onset and intervocalic position

Condition	t-statistic	df	two-tailed p-value	Significant?
Onset vs. Intervocalic	5.40	7	<0.0001	Yes

The results from this t-test reveal a probability of 0.0001, which is well below the threshold for significance of $p < .05$. In fact, the probability of this result being due to chance is less than .1 percent ($p < .001$), and the results are significant to the level of 99.9 percent. Moreover, the effect size (Cohen's d) is 2.35. Generally, effect size values of .2, .5, and .8 are thought to be small, medium, and large effect sizes, respectively (Mackey & Gass, 2005). The effect size of 2.35 calculated here is, therefore, well over the range for a large effect. Taking all of this into consideration, the results from this experiment support the hypothesis that L2ers utilize the universal process of spirantization.

Theoretical Implications

5. OT Analysis

The data presented in the previous section leads to several theoretical implications regarding the interplay between SLA and Optimality Theory (OT). OT is a framework that seeks to offer analyses of natural grammars with universal language constraints. By incorporating empirical SLA data in an OT analysis, the claim that L2 grammars are, in fact, constrained by UG, will be strengthened. Additionally, as mentioned earlier, the data reported here cannot be attributed to positive evidence from the L1 or L2. Hence, OT is an appropriate framework in which to analyze the data, presenting insights that may not be best illustrated with derivational rewrite rules.

5.1 Input—Output Representations

In phonology, one of the main goals is to analyze the difference between competence and performance. In the case of L2 phonologies, the gap between competency and performance can be wide. One way to explain this gap is in terms of mental representations and surface representations. Speakers have mental representations for segments that differ from the surface representations. Derivational phonology (Chomsky & Halle, 1968; Goldsmith, 1990; Kiparsky, 1982) assumes that derivational rewrite rules "chart" the mental representations to the surface representations, and these rewrite rules are argued to explain the difference between competence and performance.

OT argues, however, that grammars are organized differently from derivational phonology. The concept of mental and surface forms still remains, but in OT these are referred to as *input* and *output representations* (I use the terms input and output representation from here on out). The primary difference between derivational phonology and OT is that in OT, constraints, not rewrite rules, "chart" the input representations to the mental representation. Specifically, the ordering or ranking of constraints determine the output representation. When an output representation best abides by the ranking of the constraints, the output is considered to be optimal.

OT diverges from derivational in theory in other ways as well. Two of OT's abstract components are the Generator (GEN) and Evaluator (EVAL). Before the optimal output representation can surface, candidates (possible outputs) must be generated by the generator (GEN), and those candidates must be evaluated by the evaluator (EVAL) with the proper constraint ranking of the language. Derivational phonology, in contrast, only considers the actual output and does not consider other "candidates" (hypothetical outputs) as OT does. Also, GEN can give more than one hypothetical candidate and "is free to generate any conceivable output candidate for some input" (Kager, 1999, p. 20).

Depending *how* the input and output representations are categorized for the data collected for this paper, different explanations of the data can be presented. In the IL, if the underlying form for [ð] is /ð/ then a fortition argument can be presented, but if the underlying form of [ð] is /d/, then a lenition (or spirantization) argument can be presented (Shea & Curtin, 2006). (See Bauer (1988) for a discussion on the difficulty determining the difference between fortition and lenition.)

(10) Input / Output Representations of Fortition and Lenition

Lenition: Input representation of /d/ in intervocalic position = output representation of [ð]

Fortition: Input representation of /ð/ in intervocalic position = output representation of [d]

I adopt a lenition argument for the OT analysis that follows. Specifically, I argue that even at the intermediate stage of acquisition, the L2ers in this study have the input representation of /d/ for [ð]. This, of course, is crucial to the argument here as I claim the L2er data illustrate the universal process of spirantization.

However, I recognize the alternative view that if these L2ers do produce the segment [ð] that this segment is available as an input representation. In this case, a fortition argument, rather than a lenition argument, can be made. This would call for a constraint that bans continuants in intervocalic position instead of a constraint banning stops in intervocalic position, contrasting the argument posed here. Although the argument that /ð/ is available to the L2ers as an input representation is reasonable, I argue that /ð/ is not available as an input representation for the following reasons.

The first reason relates to segment [ð] itself and its relation to L1 transfer. As noted earlier, this segment is highly marked with regards to typology and articulation, strongly suggesting it will be a difficult segment to acquire. Consequently, it is assumed here that the L2ers retain L1 input representation even at an intermediate stage of acquisition. This is in line with Shea and Curtin's (2006) analysis of L2 Spanish. They report data from L1 English/L2 Spanish learners that show that the L2ers' input representation of the target L2 [ß] is, in fact, /b/. Shea and Curtin (Ibid) argue that the input representation of /b/ for [ß] remains even at the intermediate stage of acquisition at which L2ers begin to produce [ß]. The L2ers with Mongolian, Russian, Thai, and Amharic L1 backgrounds in this study are still at an intermediate level as well, and so I claim /ð/ is not yet available as an input representation.

Second, I argue that these L2ers illustrate spirantization because the results of the data mirror typological facts related to spirantization. The data here show that /d/ weakens to [ð] intervocalically rather than weakening to other phonetically similar segments such as [z], [f], and [v], which are other possibilities noted in the literature. For instance, a /d/ to [z] spirantization process is noted in the Tahltan language (Reported in Kirchner, 1998). And although I still assume that /d/ is the input representation of [ð], previously noted substitutions of the input representation /ð/ should be considered as well in order to be thorough. In L2 grammars, researchers have noted that [z] is often a substitute for /ð/ (Weinberger, 1990; Hancin-Bhatt, 1994; Lombardi, 2003), and the output of [z] is believed to be a result of L1 transfer. Finally, [f] and [v] are substitutes for /ð/ in AAVE due to the shared phonetic features of [ð], and [f] and [v] (Bailey & Thomas, 1998; Rickford, 1999).

A review of the spirantization data in Kirchner (1998) shows that when /d/ is weakened to a fricative in intervocalic position, the fricative is [ð] more often than not. This is true in the following languages: Badimaya, Dahalo, Danish, Guayabero, Gujarati, Ladakhi, Mexico City Spanish, Pennsylvania German, Purki, and Proto-Germanic. Additionally, [t] becomes [ð] intervocalically in Yindjibardndi, Uradhi, and Taiwanese. Furthermore, the typological data of spirantization show that a voiced segment never becomes voiceless in intervocalic position. However, voiceless segments can remain voiceless or become voiced in intervocalic position (e.g. British English: /t/ → [ʔ]; Gondi /k/ → [h]; Maori /k/ → [x]; Basque /k/ → [ɣ]) (as reported Kirchner, 1998), implying that the feature specification of [–voice] is the marked specification in intervocalic spirantization. Altogether, these typological facts suggest that when /d/ becomes a fricative in intervocalic position, the tendency is for the fricative to be a voiced non-alveolar coronal (i.e., [ð]), ruling out other possible segments in intervocalic position.

(11) Dispreferred Segments for Spirantization in Intervocalic Position

a) *[f]: not possible because it is a labial and voiceless
b) *[v]: not possible because it is labial
c) *[z]: not possible because it is alveolar

These facts strengthen the argument that constraint SPIR emerges and that the L2er production reported in this paper replicates the typological tendencies of spirantization.

Finally, /d/→ [ð] spirantization essentially illustrates that the segments [d] and [ð] are in complementary distribution, making [ð] an allophone of /d/, suggesting that /ð/ is unavailable at the phonemic level to the L2ers in this study. The same is true, for instance, in L1 Spanish. Although there is dialectal variation with regards to spirantization in Spanish, Lleo and Rakow (2005) conclude that *intervocalic* spirantization of stops is mandatory in all dialects. I have informally tested L1 Spanish speakers to see if they do, in fact, treat [ð] as an allophone of /d/, disallowing /ð/ phonemically. When I have asked linguistically naïve L1 Spanish speakers whether the consonants in the word [deðo] ("finger") are similar or different, they all say that the segments are pronounced the same and that each segment is a [d].[6] This is analogous to the fact that native English speakers do not perceive the difference between aspirated and unreleased voiceless stops in onset and coda position (e.g., [pʰip̚] vs. [pip] "peep"). For native English speakers, the underlying representation for an [pʰ] and [p̚] is /p/. Therefore, if L1s have allophones in complementary distribution, it is presumed that L2 grammars can as well.

Since I adopt an input representation of /d/, I use the hypothetical input of /dada/ for illustrative purposes because this input has word-initial and intervocalic position for

[6] It is possible that the Spanish speakers make this determination based on spelling as the orthographic representation is *dedo*.

the data from the intermediate stage reported in the data section of the paper. I also argue this same representation for the L2 initial state. The input of /dada/ is also used for the L2 initial state analysis as well.

5.2 OT Constraints

Rather than describing phonological grammars with derivational rules, OT's descriptive force is driven by violable constraints (constraints that can be violated). Constraint rankings determine the optimal output of a phonological form, and this fundamental characteristic of OT is applicable to the data presented here. Again, although it is possible to describe an L2er's process of spirantization with a classical derivation rule, derivational rules are assumed to be acquired through positive evidence. In the case of the L2ers with a Russian, Mongolian, Thai, and Amharic L1 background, no spirantization rule of /d/ → [ð] has been demonstrated in the L1. Moreover, these learners have not received positive evidence of spirantization from the English target language since English does not spirantize voiced stops. Without positive evidence, it would be theoretically unsound to argue that these learners have acquired a spirantization rewrite rule. Thus, the claim here is that the L2ers have reranked constraints. The constraints needed for an OT analysis of the data reported in this paper are as follows:

- **SPIRANTIZATION (SPIR):** This is a *positional markedness* constraint mandating that stops be weakened in intervocalic position (Kirchner, 1998; Shea & Curtin, 2006). This constraint is noted formally as LAZY(STOP)V_V in Kirchner (1998) and as LENITION in Kennedy (2008). I use SPIR to reflect the specific process the data in this paper illustrate. Although the descriptor/constraint LENITION (Kennedy, 2008) would not be incorrect, it is a more general description as lenition could refer to a number of processes (e.g. flapping, degemination). In this paper, the constraint SPIR reflects that the stop /d/ becomes the fricative [ð] in intervocalic position. Kirchner (1998) emphasizes that the constraint LAZY(STOP)V_V (or SPIR as I have labeled it) is particularly focused on the "minimization of articulatory effort" and is therefore a constraint that promotes lenition. As McCarthy (2002) points out, the LAZY constraint is "functionally motivated" as the constraint embodies a measurable physical event of weakening a segment (p. 222). This constraint (SPIR) is argued to be the markedness constraint that becomes highly ranked due to the trigger of the L2 input, illustrating the emergence of the unmarked.

- **MAX-IO**: This is a *faithfulness* constraint disallowing the deletion of segments: "every segment in the input has a correspondent in the output" (McCarthy & Prince, 1997, p. 6). This constraint is labeled "MAX" to emphasize the requirement that "input segments be MAXIMALLY expressed in the output" (McCarthy, 2008, p. 24). This constraint is used here to show that an IL grammar is more faithful to the input than a child L1 grammar: child L1 grammars delete more than IL grammars. I use

"MAX" in the OT analysis as this is commonly done in the OT literature.

- **IDENT (Manner)**: This is a *faithfulness* constraint that mandates that the feature(s) of the output match the feature(s) of the input regarding manner of articulation (McCarthy and Prince 1997). The term "IDENT" is short for "Identity." In the OT literature, this constraint has been "exploded" to indicate specific features: IDENT(Place), IDENT(syllabic); IDENT(round), and so on. See McCarthy (2008) for an expanded list. (See Lombardi, 2003 for an illustration of the necessity of constraint "explosion." Generally, the data in this paper violate the "identity" of place and manner features. However, I only highlight the manner feature here because it is the most prominent feature this paper discusses.

- ***ð**: This is a *segmental markedness* constraint prohibiting the segment [ð] in any syllable position (adapted from Lombardi, 2003). Note the constraint *ð is an extrapolation of Lombardi's (2003) constraint *θ. I assume here that if there can be a markedness constraint for [θ], there can also be a markedness constraint for [ð]. Like the constraint SPIR, the constraint *ð can be motivated typologically and articulatorily. Recall that interdental fricatives are rare among the world's languages (Maddieson, 1984) and that the articulation of a fricative is more difficult than the articulation of a stop (Yavas, 1998, p. 138).

  When discussing the initial state of the L2ers' grammar, I use the constraint *ð because it is assumed that L2ers begin L2 acquisition with their L1 constraint ranking. Since interdental fricatives are not part of L1 Russian, Mongolian, Thai, or Amharic, the constraint *ð must be highly ranked.

- ***ð(Onset)**: This is a *segmental markedness* constraint prohibiting the segment [ð] in a word-initial onset position. This constraint is a proposed subset constraint of the more general constraint *ð presented above. This subset constraint is used in this paper because it accounts for the data that show the segment [ð] is not consistently produced in word-initial onset position. The constraint *ð(Onset) also directly relates to the sonority of the onset segment. If we adopt Selkirk's (1984) sonority scale, for instance, which states that fricatives are more sonorous than stops:

  (12) Selkirk's (1984) Consonant Sonority Scale – least to most sonorous:

  p,t,k < b,d,g < f,θ < v,z,ð < s < m,n < l < r

  and we assume that onsets prefer the least sonorous segment in order to have the highest possible peak between the onset and vowel, the segment [ð] is dispreferred in onset position because it is more sonorous than stops such as [d]. Of course, a more general constraint pertaining to sonority such as, *ONSET/FRICATIVE (i.e., onsets do not have the sonority level fricative), could account for why more sonorous onsets are dispreferred over less sonorous onsets (Prince & Smolensky, 1993: §8.1; Zec, 1988; Clements, 1990). However, I retain the constraint *ð(Onset) here as this paper's focus is the segment [ð].

Additionally, the constraint *ð(Onset) is also necessary to avoid stringency violations with the more general (or less stringent) constraint *ð. I discuss stringency violations in section 5.4.

In the analysis that follows, the constraints above will have different rankings depending on the L2ers's particular stage of acquisition. At the initial state it is presumed that the markedness constraint *ð is highly ranked, prohibiting production of [ð] altogether. Once the L2ers begin to acquire the L2, the target L2 triggers constraint re-rankings. The data in this paper reflect an intermediate stage of L2 English acquisition. In this stage, the markedness constraint SPIR "emerges" and permits [ð] in intervocalic position. However, in word-initial onset position, the markedness constraint *ð (Onset) is ranked higher than the faithfulness constraint IDENT(Manner), disallowing [ð]. Consequently, this intermediate stage of acquisition reflects an IL grammar that is partially producing the segment [ð]. In this sense, the data reflect partial faithfulness to the L1 grammar. As the L2ers progress, we expect that they will be more accurate in their production of [ð] in all contexts, coming close to the accuracy rates of native English speakers. This shift would illustrate total faithfulness to the L2 grammar.

5.3 L2 Initial State Constraint Ranking

It is generally assumed that the initial state of the L2 incorporates most, if not all, the features/representations of the L1, and presumably, the L2 initial state prohibits [ð] in all contexts. With the Russian, Mongolian, Thai, and Amharic L1s, a constraint ranking regarding the fricative [ð] is straightforward in that [ð] is prohibited in the L1 altogether. Recall that, as far as fricatives are concerned, Russian permits [f, v, s, z, ʃ, ʒ, x] (Monk & Burak, 2001); Mongolian permits [x, χ, x[illegible], s, ʃ] (Svantesson et al., 2005); Thai only permits [f, s, h] (Campbell, 1995; Tingsabadh & Abramson, 1999); the fricatives in Amharic are [f, s, s', z, ʃ, ʒ, h] (Hayward & Hayward, 1999). None of the permissible fricatives are interdentals and so the constraint *ð is *highly* ranked in these L1s. At the same time, these L1s do not demonstrate the spirantization process of /d/ → [ð], so a constraint mandating the spirantization of stops (e.g. SPIR) is ranked *low* in these grammars. Tableau (1) below represents the L2 initial state constraint ranking regarding the absence of [ð] and spirantization in L1 Russian, Mongolian, Thai, and Amharic.

In tableau (1), below, it is assumed that at the initial state, these L2ers'output representation will not have [ð]. I use the input representation of /d/ for [ð] because the L2ers presumably transfer their L1 input representations to the L2 initial state. With this in mind, I use the hypothetical input of /dada/ rather than actual English words for illustrative purposes as the input /dada/ has all the necessary syllable positions needed for the analysis.

Tableau (1): L2 Initial State Constraint Ranking

/dada/	*ð	IDENT(Manner)	SPIR
→ a. dada		*	
b. daða	*!	*!	
c. ðada	*!	*! *	

Candidate-a is the preferred surface form as there is no violation of the markedness constraint *ð, nor is there a violation of the constraint IDENT(Manner). The non-violation of INDENT(Manner) illustrates L1 transfer: the L2ers are faithful to their transferred L1 input representations. Since [ð] is not permitted in the output in intervocalic position, the constraint SPIR is the only constraint violated in the optimal candidate; thus, it is ranked lowest here. The constraints *ð and IDENT(Manner) are separated by a broken line to signify they are equally ranked as candidate-a does not violate either constraint. Although neither constraint can be officially noted as dominating the other, intuition suggests that *ð is the highest ranking constraint since the segment [ð] is prohibited in all the L1s noted in this study. Since the constraint the *ð is violated, this assumes that its subset constraint *ð(Onset) is also violated.

Regarding the losing candidates -- b and c -- they each have a "fatal" violation of the equally ranked constraints *ð and IDENT(Manner), producing dispreferred outputs. Candidate-b does not violate SPIR because [ð] is present in intervocalic position. Candidate-b violates the markedness constraint SPIR by permitting [ð] intervocalically. These losing candidates show that [ð] is dispreferred altogether.

5.4 IL Intermediate Stage Constraint Ranking

The overall results of the data from the IL grammars gathered for this paper report [ð] is produced more in intervocalic-initial position (90%: 325/360) than in word-onset position (52%: 187/360).[7] Production of [ð] suggests there has been a shift in constraints from the initial state constraint ranking proposed above in tableau (1). One constraint ranking shift is that SPIR has become promoted, permitting [ð] in intervocalic position. Thus, the constraint *ð can now be violated, but this constraint is violated by the separate subset constraint of *ð(onset). Tableau (2) below shows the constraint ranking for this IL intermediate stage. Again, I use the hypothetical input of /dada/ for illustrative purposes.

[7] These data illustrate variable output by L2ers. Sometimes the L2ers produce [ð] correctly in all contexts and sometimes one context has a higher production rate than another. Thus, the constraint rankings are not necessarily fixed. Stochastic OT (Boersma, 1998; Boersma & Hayes, 2001) is an amendment to "classic" OT theory in that it argues that constraints can be variable and do not have to have a fixed ranking. So, for instance, Constraint A could dominate Constraint B at times and vice versa at other times. Although the data here could be incorporated into a Stochastic OT model, I have chosen to explain constraint rankings in terms of what most often happens in the production data.

Tableau (2): IL Intermediate ranking

/dada/	*ð(Onset)	SPIR	IDENT(Manner)
→ a. daða			*
b. ðada	*!	*!	*
c. ðaða	*!		**
d. dada		*!	

As tableau (2) reflects, the constraints *ð(Onset) and SPIR are equally ranked as these constraints are not violated in optimal candidate-a; the constraint IDENT(Manner) is violated in the optimal candidate, so it is ranked the lowest. The non-optimal candidates b through c either have a fatal violation of *ð(Onset) and/or SPIR.[8] The segment [ð] is still dispreferred in word-initial onset position and the segment [d] is dispreferred in intervocalic position. Again, I use the constraint *ð(Onset) instead of the constraint *ONSET/FRICATIVE as the specific segment [ð] is the dispreferred segment in onsets. Other fricatives, such as [s], [f], [v], and [z] are available to these L2ers.

The constraint *ð is not in tableau (2) because of the so called "stringency relationship" (McCarthy, 2008). I illustrate this below in tableau (3). Tableau (3) is the same as tableau (2) with the constraint *ð added on the far right.

Tableau (3): Stringency Violation(SV)

	(SV)			(SV)
/dada/	*ð(Onset)	SPIR	IDENT(Manner)	*ð
→ a. daða			*	*
b. ðada	*!	*!	*	*
c. ðaða	*!		**	**
d. dada		*!		

A stringency relationship is a relationship between a general and specific constraint in which a violation of the specific constraint is *also* a violation of the general constraint. Thus, such constraints cannot conflict and, as a consequence, cannot be ranked (McCarthy, 2008, p. 67). In this case, the specific constraint (more stringent) is *ð(Onset) and the general constraint (less stringent) is *ð. The constraint *ð cannot be ranked below *ð(Onset) due to a stringency violation. Because of this stringency violation, the constraint *ð is not included in intermediate ranking in tableau (2).

The most important shift noted in tableau (2) is the re-ranking of SPIR. The constraint SPIR permits a violation of *ð, illustrating the process of the emergence of the unmarked in which fricatives are preferred in intervocalic position. This positional markedness constraint has "emerged" (been promoted) in the IL grammar. As the section above pointed out, the constraint SPIR is ranked the lowest in the initial state ranking as [ð] is not predicted to be produced in any context at that stage. However, due to the trigger of the L2 input at the intermediate stage, SPIR is ranked the highest, facilitating the production of [ð] in intervocalic position. This constraint shift, in which SPIR has emerged,

[8] It is assumed that other markedness constraints such as *θ and *ɣ are also active, but I have omitted them since the focus here is on [ð].

facilitates the positionally conditioned production of [ð] (or "positional asymmetry" in which [ð] is more marked in word-initial onset position than in intervocalic position). The positional asymmetry that the constraint SPIR prompts in an L2 grammar is certainly curious. It could not have come from L1 or L2 input, as there is no evidence of such a process in either of these grammars. At the same time, the data that show [ð] is not produced in word-initial onset is unsurprising since this illustrates L1 transfer.

Keeping the facts of these data in mind, we are given clues as to what roles L1 transfer and UG play in the acquisition of an L2. There are a variety of proposals concerning the relationship between UG and L1 transfer in an L2 grammar. The data here support, for example, the Full Transfer-Full Access proposal (FTFA), which states that the L1 is the initial state of the L2 and that the L2 grammar conforms to the properties of UG (Schwartz & Sprouse, 1994, 1996). In short, at least with this phonological data, L1 transfer affects acquisition and UG is accessible. Moreover, not only does the constraint SPIR illustrate access to UG, but it also helps codify the characteristics of an IL phonological grammar and how it differs from an emerging L1 and adult L1 grammar. One such difference is the positionally conditioned production of [ð] (or "positional asymmetry"). The role markedness and faithfulness play is another difference between an IL grammar, an emerging L1 grammar, and an adult L1. I expand on these differences in the next section.

5.5 Emerging L1 grammars and adult L1 grammars

In an emerging L1 grammar and an adult L1 grammar, this positional asymmetry illustrated in the L2 grammar is absent. In an emerging L1 child grammar, the constraint *ð will be ranked above MAX or IDENT. For instance, Moskowitz (1970) presents data from a two-year-old child, Erica, that show the accuracy rate of [ð] to be at zero. In word-initial and intervocalic position, Erica deletes [ð] altogether or uses [d] as a substitute. It is the substitution that I wish to focus on here. In tableau (4), below, we see that the preferred surface output, candidate-a, has substituted [d] for [ð]. It is assumed here that the input representation for an emerging L1 grammar would be /ð/ (I assume L1ers interpret L1 segments at "face value," which is, of course, divergent from the input representations of an L2er). I use the hypothetical input of /ðaða/ for illustrative purposes.

Tableau (4): Child L1 Grammar: Constraint Ranking for [ð]

/ðaða/	*ð	IDENT(Manner)
→ a. dada		**
b. ðaða	**!	

The constraint *ð(Onset) is not present in tableau (4) because the segment [ð] is deleted in all syllable positions, and so this constraint does not apply here (Also, having both *ð(Onset) and *ð would be a stringency violation). The indication here is that the child L1er treats all syllable positions equally. Moreover, the emerging L1 grammar ranking in tableau (4) is similar to the L2 initial state ranking (see tableau (1), above) in that interdental fricatives are prohibited in the grammar. The ranking in tableau (4) illustrates the preference of markedness over faithfulness (or unmarked structures over marked structures) in a child L1 grammar. That is, the markedness constraint *ð dominates the faithfulness constraint IDENT(Manner). Thus, Erica's data show unmarked surface forms trumping marked surface forms. (See Gnanadesikan (1995) for an in-depth OT account of a child L1 grammar.)

In an L1 adult grammar, the ranking presented in tableau (4) would be reversed, illustrating that faithfulness dominates markedness.

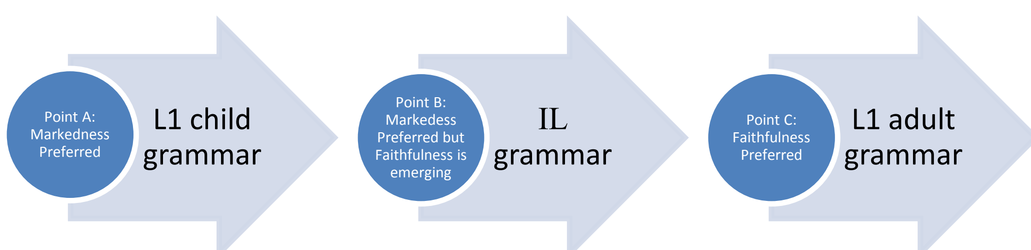
Tableau (5): Adult L1 grammar: Constraint Ranking for [ð]

/ðaða/	IDENT(Manner)	*ð
→ a. ðaða		**
b. dada	**!	

Also, in the adult L1 grammar, as in the emerging L1 grammar, there is no positional asymmetry regarding the segment [ð]. Presumably, this would be the same constraint ranking of a final state L2 grammar that reflected accurate production of [ð] in word-initial onset and intervocalic position. Also, the input representation for an L1 adult grammar and a final state L2 grammar production accuracy of [ð] across the board would each have the input representation of /ð/.

In sum, a comparison of the L1 emerging grammars, the L2 intermediate grammar, and the L1 adult grammar reveal that the IL intermediate grammar could be metaphorically described "as one step ahead" of child L1 grammar and "one step behind" an L1 adult grammar.

Figure (6): Markedness to Faithfulness Continuum



The IL intermediate grammar is one step ahead in the sense that the syllable position of [ð] matters in the IL grammar whereas in L1 child grammars [ð] can be deleted in all positions as illustrated in the L1 data presented in Moskowitz (1970). The IL grammar is one step behind an L1 adult grammar in the sense that there is still a preference for unmarked forms; however, marked forms such as [ð] are emerging in the IL grammar.

With this continuum in mind, it is assumed that an end state L2 grammar that produces [ð] in all contexts has the same constraint ranking as an adult L1 grammar.

5.6 Summary of OT Rankings

In sum, the following stages of [ð] production were presented:

(1): At the initial state, L2ers transfer their L1 ranking and L1 input representations which prohibit interdental fricatives.
(2): At the intermediate stage of acquisition the input representation of [ð] is still /d/, and L2ers promote the constraint SPIR to permit [ð] in intervocalic position, illustrating the emergence of the unmarked. The constraint *ð(Onset) prohibits [ð] in onset position.
(3): At the final state of acquisition, L2ers have a context-free input representation of /ð/ and demote *ð(Onset) and SPIR, prompting production of [ð] across the board.

These stages move from disallowing marked segments to permitting marked segments in the output. So, over time, an L2er's grammar shifts from being dominated by markedness constraints to be being dominated by faithfulness constraints.

The re-ranking in stage (3) noted directly above, of course, is hypothetical. L2ers show variability, and it is plausible that some learners would become more inconsistent than stated in stage (2), illustrating a U-shaped acquisition curve. It is also plausible that certain learners would *never* shift the input representation from /d/ to /ð/ and/or produce [ð] correctly, illustrating fossilization.

6. Conclusion

This empirical study has shown that [ð] is a troublesome segment for L2 English learners. More interestingly, however, this study shows that the universal process of spirantization is accessible in L2 acquisition. The high accuracy rate of [ð] in intervocalic position is the primary support for L2 spirantization.

Although this study makes claims related to markedness and UG in an L2 context, there are limitations to the study. First, it could be argued that the word list and story reading tasks only prompt a formal speech style, possibly inflating the results presented here. Second, frequency was not evaluated here. An analysis of frequency may reveal that more frequent words with [ð] intervocalically may have a higher accuracy rate than less frequent words with [ð] intervocalically. However, the lower accuracy rates with [ð] in word-onset position suggest that frequency is most likely not a factor for this position. Moreover, this study did not incorporate a contrastive analysis between function and content words (See Bybee (2001) for a discussion on function vs. content words).

Since the data reflect a pattern that is underdetermined by the L1s or the target L2 English, a derivational rewrite rule accounting for the data cannot be presented. Thus, I have shown that an OT analysis is an appropriate framework to invoke here. The data reported here suggest that the ranking of the constraint SPIR illustrates the emergence of the unmarked. The emergence of SPIR creates a positional asymmetry between word-initial onset and intervocalic position in which [ð] is inconsistently produced in word-initial onset, but consistently produced in intervocalic position.

Overall, the investigation of the L2 production of [ð] permits researchers to evaluate the role markedness and UG play in L2 acquisition. Consequently, L2 studies focusing on interdental fricatives remain a fruitful area of research.

References

Altenberg, E. and Vago, R.M. (1987). Theoretical implications of an error analysis of second language phonology production. In G Ioup, and S. Weinberger (Eds.), *Interlanguage phonology* (pp. 148–164). Cambridge, MA: Newbury House.

Archive Phonetics. (2012). UCLA Phonetics Lab: http://archive.phonetics.ucla.edu/. Retrieved from Oct. 1, 2012, from http://archive.phonetics.ucla.edu/Language/AMH/amh_word-list_1966_01.html

Asher, J & Garcia, R. (1969). The optimal age to learn a foreign language. *Modern Language Journal*, *53*, 334-341.

Bailey, G. & Thomas, E. (1998). Some aspects or African-American Vernacular English phonology. In S. Mufwene, J. Rickford, G. Bailey, and J. Baugh (Eds.), *African-American English* (pp. 85-109). New York: Routledge.

Barlow, J. A. & Gierut, J. A. (1999). Optimality theory in phonological acquisition. *Journal of Speech Language & Hearing Research, 42,* 1482-1498.

Bauer, L. (1988). What is lenition? *Journal of Linguistics* , *24*, 381-92.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech, Perception and Linguistic Experience: Issues in Crosslanguage Research (pp.* 171–203). Baltimore: York.

Boersma, P. (1998). *Functional Phonology: Formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.

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

Broselow, E., Chen, S.-I., & Wang, C. (1998). The Emergence of the Unmarked in Second Language Phonology. *Studies in Second Language Acquisition, 20*, 261-280.

Bybee, J. (2001). *Phonology and language use*. Cambridge Studies in Linguistics 94. Cambridge: Cambridge University Press.

Campbell, G.L. (1995). *Concise Compendium of the World's Languages.* New York: Londres.

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

Chomsky, N., & Halle, M. 1968. *The sound pattern of English*. New York: Harper & Row.

Clements, G. N. 1990. The role of the sonority cycle in core syllabification. In J. Kingston and M. Beckman (Eds), *Papers in Laboratory Phonology I: Between the*

*Grammar and Physics of Speech (*pp. 283-333). Cambridge: Cambridge University Press.

Colavin, R., Levy, R. & Rose, S. (2010). Modeling OCP-Place in Amharic with the Maximum Entropy phonotactic learner. Chicago Linguistic Society 46.

Eckman, F.R. & Iverson, G.K. (1993). Sonority and markedness among onset clusters in the interlanguage of ESL learners. *Second Language Research 9*, 234–52.

Elliott, R. E. (1995) Field independence/dependence, hemispheric specialization, and attitude in relation to pronunciation accuracy in Spanish as a foreign language. *The Modern Language Journal*, *79*, 356-371.

Flege, J. E. & Fletcher, K. L. (1992). Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America*, *91*, 370-389.

Flege, J., Munro, M. & MacKay, I. (1995). The effect of age of second language learning on the production of English consonants*. Speech Communication*, *16*, 1-26.

Flege, J. E., Yeni-Komshian, G. & Liu, H. 1999. Age constraints on second language acquisition, *Journal of Memory & Language*, *41*, 78-104.

Gnanadesikan, A. (1995). Markedness and faithfulness constraints in child phonology. Ms., University of Massachusetts, Amherst.

Goldsmith, J. A. (1990). *Autosegmental and metrical phonology*. Cambridge, MA: Blackwell.

Hancin-Bhatt, B. 2008. Second language phonology in Optimality Theory. Phonology and Second Language Acquisition. Ed. M Zampini and J Hansen.
Amsterdam/Philadelphia:
John Benjamins, 2008. 117-146.

Hayward, K. & Hayward, R. (1999). Amharic. In International Phonetic Association (Ed.), *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet* (pp. 45-49). Cambridge, UK: Cambridge University Press.

James, A. (1986). *Suprasegmental Phonology and Segmental Form.* Niemeyer: Tübingen.

Kahn, D. (1976). Syllable-based Generalizations in English Phonology, Doctoral dissertation, MIT.

Kiparsky, P. (1982). Lexical morphology and phonology. In The Linguistic Society of Korea (Ed.), *Linguistics in the morning calm* (pp. 1–91). Seoul, Korea: Hanshin.

Kirchner, R. (1998). An effort-based approach to consonant lenition. Doctoral Dissertation. UCLA. [Available on Rutgers Optimality Archive, ROA-276.]

Kirk, C, & Katherine D. (2003). Onset/coda asymmetries in the acquisition of clusters. In B. Beachley, A. Brown, & F. Conlin (Eds.), *Proceedings of the 27th Annual Boston University Conference on Language Development* (pp. 437-448). Somerville, MA: Cascadilla Press.

Kohmoto, S. (1965). *Applied English Phonology: Teaching of English Pronunciation to the Native Japanese Speaker*. Tanaka Press.

Kuhl, P. K. (1995). 'Linguistic experience and the perceptual magnet effect in W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Crosslanguage Research* (pp. 121–154). Timonium: MD.

Ladefoged, P. (1993). *A Course in Phonetics*, 3rd ed. Fort Worth TX: Harcourt Brace & Company.

Leslau, W. (1995). *Reference Grammar of Amharic.* Wiesbaden: Harrassowitz.

Lewis, M. P. (Ed.). (2011). Ethnologue: Languages of the World, Sixteenth edition. Dallas, Tex.: SIL International. Online version: http://www.ethnologue.com/.

Lleó, C. & Rakow, M. (2005). Markedness Effects in the Acquisition of Voiced Stop Spirantization by Spanish-German Bilinguals. In Cohen, J., McAlister, K. T., Rolstad, K. & MacSwan, J. (Eds.), *Proceedings of the International Symposium on Bilingualism 4* (pp. 1353–1371.). Cascadilla Press, Somerville, Massachusetts.

Lombardi, L. (2003). Second language data and constraints on Manner: explaining substitutions for the English interdentals. *Second Language Research*, *19*(3), 225-250.

Maddieson, I. (1984). *Patterns of Sounds*. Cambridge: UP Press.

Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. Mahwah, NJ: Lawrence Erlbaum.

McCarthy, J. (2002). *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.

McCarthy, J. (2008). *Doing Optimality Theory*. Malden, MA and Oxford: Blackwell.

McCarthy, J., & Prince, A. (1994). The emergence of the unmarked: Optimality in prosodic morphology. *Proceedings of the North East Linguistics Society 24*, 333–379.

McCarthy, J. & A. Prince. (1997). Faithfulness and Reduplicative Identity. [Available on Rutgers Optimality Archive, ROA-216.]

Michaels, D. (1974). Sound replacements and phonological systems. *Linguistics* , *126*, 69–81.

Missaglia, F. (1999). Contrastive prosody in SLA: an empirical study with adult Italian learners of German. *Proceedings of the 14th International Congress of Phonetic Sciences*, *1*, 551-554.

Monk, B. & Burak, A. (2001). Russian speakers. In M. Swan, & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (pp. 145– 161).

Moskowitz, A. I. (1970). The two-year-old stage in the acquisition of English phonology. *Language, 46*, 426–41.

Moyer, A. (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition, 21*, 81-108.

Oyama, S. (1982). The sensitive period for the acquisition of a nonnative phonological system. In R. Scarcella & M. Long (Eds.), *Child-Adult Differences in Second Language Acquisition* (pp. 20-38). Newbury House, Rowley, MA

Piper, T. & Cansin, D. (1988). Factors influencing the foreign accent. *The Canadian Modern Language Review, 44*, 334-342.

Purcell, E & Suter, R. (1980). Predictors of pronunciation accuracy: A reexamination. *Language Learning, 30*, 271-287.

Prince, A. & Smolensky, P. (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar.* Malden, MA, & Oxford: Blackwell. [Revision of 1993 technical report, Rutgers University Center for Cognitive Science. Available on Rutgers Optimality Archive, ROA-537.]

Rickford, J. R. (1999). *African American Vernacular English: Features and Use, Evolution, and Educational Implications*. Oxford: Blackwell.

Selkirk, E. (1984). On the major class features and syllable theory. In Aronoff & Oehrle (Eds.) *Language Sound Structure: Studies in Phonology* (pp. 107-136). Cambridge: MIT Press.

Shea, C and Curtin S. (2006). Learning Allophonic Alternations in a Second Language: Phonetics, Phonology, and Grammatical Chang. In M. O'Brien, C. Shea, & J. Archibald, (Eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)* (pp. 124-131). Somerville, MA: Cascadilla Proceedings

Svantesson, J., Tsendina, A., Karlsson, A., & Franzén, V. (2005). *The phonology of Mongolian.* Oxford: Oxford University Press.

Tahta, S.,Wood, M ., & Lowenthal, K. (1981). Foreign accents: Factors Relating to transfer of accent from the first language to the second language. *Language Speech, 24*, 265-272.

Tingsabadh, M.R., & Abramson, A. (1999). Thai. In International Phonetic Association (Ed.), *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet* (pp. 147-150). Cambridge, UK: Cambridge University Press.

Yavas, M. (2011). *Applied English phonology.* Malden, MA: Blackwell Publishing.

Yavas, M. (1998). *Phonology development and disorders.* San Diego: Singular Publishing.

Weinberger, S. (1990). Minimal segments in second language phonology. In J. Leather & A. James (Eds.), *New Sounds 90: Proceedings of the 1990 Amsterdam Symposium on the acquisition of second language speech*. University of Amsterdam.

Weinreich, U. (1953/1968). *Language in contact.* The Hague: Mouton.

Wireback, K. (1997). *The role of phonological structure in sound change from Latin to Spanish and Portuguese*. New York: Peter Lang.

Zec, D. (1988). *Sonority Constraints on Prosodic Structure*. PhD dissertation, Stanford University.

APPENDIX: READING TASKS

For each word, insert the word into the following phrase: "Now say ______ again": "Now say 'big' again."

1)	big	four	etch	maze	legal
2)	earth	when	this	hurts	another
3)	hope	brother	vision	then	jolly
4)	thinks	your	stripe	mother	bath
5)	van	personal	principle	neither	cannot
6)	green	these	bathe	fresh	apples
7)	speak	sunny	rather	author	than
8)	those	sprint	phrase	joke	weather
9)	math	bread	pants	father	hopes
10)	dials	them	played	games	hotel
11)	other	birth	though	locks	doors
12)	changes	other	admit	curse	there
13)	even	they'll	eggs	leather	fox
14)	burst	breathe	now	therefore	joy
15)	strange	weather	thumb	fingers	Athens
16)	the	offer	ouch	angry	either
17)	healthy	school	stereo	spend	pencils
18)	essay	notes	bother	teacher	they
19)	orange	teethe	commuting	drives	passion

20) choices whether maybe oceans first

Story reading task

This story is about two brothers, Othello and Keith, from Athens, Georgia. Othello was always thought of by the family to be the one who did everything right while Keith was always thought to do things wrong.

Thanks to a winning lottery ticket, Keith was able to buy a house with his own money, and decided to move southwest of Athens to Oglethorpe County.

Othello's mother thought that Othello would go see Keith often after he moved to Oglethorpe County, but Othello never went to see his brother there.

One winter the weather was really bad and Keith had no heat or running water, but Othello still did not visit or help Keith either. Keith called his mother, who at that time, lived in another state to tell her about this situation between him and his brother. She was really confused about everything that Keith told her and she said that she planned to talk to their father about the situation.

A week went by, and then another, and finally the mother spoke to the father about the situation with the brothers. Later in the evening after Keith called, his mother was a little nervous about telling her husband about the situation. She thought about this and she knew that this would either bother him a little or really make him angry. After the mother told the father about the brothers, she was surprised because he laughed. He laughed and laughed and he did not bother to tell her why. The mother just threw her hands up in the air during this strange display from the father. Therefore, the mother did not know whether the father had understood her or whether he was keeping something from her.

The mother had to find out what was going on between the father and the brothers, Othello and Keith. She had some questions: Why was neither brother speaking to each other? What did her husband think was so funny? Why on earth did Keith think it was a good idea to live in Oglethorpe County?

The Distinction between Noun-Phrase Premodifiers: Nouns are not Adjectives

Justin Voigt

George Mason University, Fairfax, Virginia

The Distinction between Noun-Phrase Premodifiers: Nouns are not Adjectives

Determining the syntactic category of an out-of-context lexical item is often impossible. For example, *fight* may be either a noun (e.g., *The fight was enjoyable*.) or a verb (e.g., *They fight for freedom*.). The category of these lexemes is only distinguishable when they enter the syntax in a phrase or clause thereby exhibiting the distribution characteristic of their category. However, even in context, nouns seem to function like adjectives quite often and quite productively. For example, *medicine* may function like a noun (*The medicine has expired*.) or an adjective (*The medicine cabinet was closed*). Thus, both nouns and adjectives seem to be able to function as a premodifier (or adjunct) of some head noun (see Adger, 2003, p. 275-277).

Of course, one might simply suppose that *rigid* and *granite* in (1) and (2) below are both adjectives.

(1) *His rigid resolve won the battle.*
(2) *His granite resolve won the battle.*

Both words ascribe some property to a head noun—doing so in a very similar way, and they both seem to share the same syntactic position. Moreover, both constituents allow for recursion as demonstrated by (3) and (4) below:

(3) *this old national treasure*
(4) *the toy car collection* (from Giegerich, 2009)

Nevertheless, a number of constituency tests will demonstrate that these two premodifiers maintain the distinguishing characteristics of their category even in this shared syntactic position.

Already, the recursion from (3) and (4) above expose a distinction between the two premodifiers. Notice that *old* does not premodify *national*; instead, it seems to describe *treasure* or perhaps *national treasure*. In contrast, *toy* does seem to be premodifying *car* in that it is a collection of [*toy cars* $]_{N}$, not a collection of cars that are toy. While adjectives are certainly allowed to modify another adjective in a similar way as *toy* and *car* do (see (5) below), that ability seems much more limited than with nouns (consider (6)).

(5) [*icy cold purple soft drink*]$_N$
a. [[[*icy*]$_A$ [*cold*]$_A$]$_A$ [*purple*]$_A$ [*soft*]$_A$ *drink*]$_N$
b. [[A+A] [A] [A] [N]] or, alternatively,
c. [[A+A] [A [A+N]]]]
d. *[[A+A] A] A] [N]]

(6) [*toy-car-collection catalog page number*]$_N$
a. [[[[[*toy*]$_N$ [*car*]$_N$]$_N$ [*collection*]$_N$]$_N$ [*catalog*]$_N$]$_N$ [[*page*]$_N$ [*number*]$_N$]$_N$
b. [[N+N] N] N] [N+N]]

The adjectives can only premodify another adjective once, not recursively, and the relationship between those adjectives seems to be unique (i.e. coordinated). In contrast, nouns can premodify nouns without limit (theoretically). Thus, the first distinction between these two premodifiers is that nouns can adjoin recursively in a much more productive and unrestricted way than adjectives.

Distinct constituencies demonstrate the second difference between these two premodifiers. Adjectives can be premodified by degree words while nouns are generally premodified only by adjectives or other nouns. In (7)-(10) below, the degree word *very* shares a constituency with the adjective *rigid*, but not the noun *granite*, and *almost* shares a constituency with *musical*, but not *music*.

(7) *The very rigid laws of the king could not be changed.*
(8) **The very granite laws of the king could not be changed.*
(9) *The* [[*almost musical*]$_A$ *museum*]$_N$ *was alive with sound.*
(10) **The* [[*almost music*]$_N$ *museum*]$_N$ *was alive with sound.*

Likewise, adjectives have a limited ability to share a constituency with adjectives and nouns, explaining (in part) the limited recursion in the discussion directly above and further suggesting difference between adjective and noun premodifiers. In sum, the second distinction between these two premodifiers is that adjectives share a constituency of adjectives (as one would expect). Nouns do not share this constituency, even in a noun-premodifier position.

Coordination demonstrates the third distinction between these two premodifiers. Generally, only like constituents can be coordinated, and, although many noun premodifiers cannot be coordinated, those that can only seem to allow coordination with other nouns.

(11) [_strong and rigid_]$_{\text{A}}$
a. _The strong and rigid will of the king won the battle._
b. ?*_The strong and granite will of the king won the battle._
c. *_The granite and strong will of the king won the battle._

(12) [_medicinal and herbal_]$_{\text{A}}$
a. _The medicinal and herbal advice in this book is valuable._
b. ?*_The medicinal and herb advice in this book is valuable._
c. ?*_The medicine and herbal advice in this book is valuable._

(13) [_medicine and herb_]$_{\text{N}}$
a. _This medicine and herb cabinet should have a lock on it._
b. *_This medicine and herbal cabinet should have a lock on it._
c. *_This medicinal and herb cabinet should have a lock on it._

Thus, the third distinction between adjective and noun premodifiers is that they cannot be coordinated with each other.

Finally, thematic roles demonstrate the last distinction between these two premodifiers. Adjectives cannot be assigned thematic roles (Adger, 2003, p. 275), but nouns can. We would expect this to hold between adjective and noun premodifiers. Below, (14)-(17) demonstrate this reality:

(14) [_street seller_]$_{\text{N+N}}$
a. _someone who sells on the street_ <locative>
(15) [_cake baker_]$_{\text{N+N}}$
a. _someone who bakes cakes_ <theme>
(16) [_alpine hiker_]$_{\text{A+N}}$
a. *_someone who hikes in alpine_ <locative>

b. *someone who hikes in the mountains* <locative>

(17) [*lunar observer*]$_{A+N}$

a. **someone who observes lunar* <theme>

b. *someone who observes the moon* <theme>

Adjectives can only get a thematic role if they are derived from a noun and they revert back to the underlying noun. We would expect non-derived adjectives to not have any recourse of repair for thematic role assignment. Below, (18)-(19) show just that:

(18) [*fast hiker*]$_{A+N}$

a. **someone who hikes in fast* <locative>

b. *someone who is fast at hiking*

(19) [*bad observer*]$_{A+N}$

a. **someone who observes bad* <theme>

b. *someone who is bad at observing*

Although adjectives and nouns share the same syntactic position as a premodifier of some head noun, they exhibit the distribution unique to their distinctive category. This should come as little surprise, as adjectives share other syntactic positions with nouns; see (20)-(21) below:

(20) *The car is* [*a lemon.*]$_{NP}$

(21) *The car is* [*blue.*]$_{AP}$

Using various tests such as recursion, constituency, coordination, and thematic-role assignment, we have demonstrated a distinct, categorical difference between two noun-phrase premodifiers: adjectives and nouns. These lexical items are categorically distinct, and they remain distinct even when they share a syntactic position.

Can the Syntactic Burden on Working Memory Account for Island Constraints?

Justin Voigt

George Mason University, Fairfax, Virginia

Can the Syntactic Burden on Working Memory Account for Island Constraints?

The prevailing account of filler-gap dependencies in the psycholinguistic literature posits a limited workspace of memory for processing utterances, called working memory (Just & Carpenter, 1992; although see MacDonald & Christiansen, 2002, for the connectionist account). Such an account correlates the limitedness of working memory to processing difficulties; processing difficulties arise as an individual's working memory capacity fills up—sometimes filling up to the point that information is shunted from working memory (Dickey, 1996).

For example, in filler-gap dependencies, greater distances[1] from filler to gap require a filler to be held in working memory longer while intervening words (with their syntactic nodes) are processed. As things pile up in working memory, processing slows. The distance between filler and gap is demonstrated in (1) below.

1) a. *The pole*$_1$ *that* Ø$_1$ *struck Sydney...*
   distance: n syntactic nodes
   b. *The pole*$_1$ *that Sydney hit* Ø$_1$*...*
   distance: n + ≥10 syntactic nodes
   c. *The pole*$_1$ *that Sydney walked the dog into* Ø$_1$*...*
   distance: n + ≥18 syntactic nodes

The greater distances associated with the filler-gap dependency in (1 b.) predicts that it will be more difficult to process than (1 a.) (see King & Just, 1991).

An increased syntactic burden on working memory can also account for the distinction among (2-4).

2) **The child that the dog that the man owned bit cried.*
   **The child*$_1$ [*that the dog*$_2$ [*that the man owned* Ø$_2$]$_{R\text{-}CL2}$ *bit* Ø$_1$]$_{R\text{-}CL1}$ *cried.*
   greatest distance: n + ≥45 syntactic nodes
3) *The child that the dog bit cried.*
   The child$_1$ [*that the dog bit* Ø$_1$]$_{R\text{-}CL1}$ *cried.*

[1] Distance can alternately be defined as the number of words between filler and gap (often called just "distance") or as the number of relevant syntactic nodes separating a filler from its gap (often called "depth"). Although distance and depth often happen to correspond, evidence suggests that depth has a stronger correlation to processing load (e.g. McKee & McDaniel, 2001; O'Grady, Lee, & Choo, 2003; also Dickey, 1996). I use "distance" to mean only the number of intervening syntactic nodes (i.e. "depth").

greatest distance: n + ≥10 syntactic nodes

4) *I saw the man that owned the dog that bit the child that cried.*

I saw the man$_1$ [*that* Ø$_1$ *owned the dog*$_2$ [*that* Ø$_2$ *bit the child*$_3$ [*that* Ø$_3$ *cried*]$_{R\text{-}CL3}$]$_{R\text{-}CL2}$]$_{R\text{-}CL1}$.

greatest distance: n syntactic nodes

Consider, first, the difference in filler-gap distance associated with the fillers in (2) and the fillers in (3-4). As I demonstrate in Diagram 2 (Appendix 1), the first filler (*the child*) in (2) is separated from its gap by a much greater distance than any of the fillers in (3-4). In fact, *the child* is separated from its gap by at least twice as many syntactic nodes as the other fillers.

In addition to the greater distance associated with its filler-gap dependency, (2) also requires multiple fillers to be held in working memory simultaneously. For example, *the dog* must be held in working memory until after *owned* is processed; *the child that the dog that the man owned* must be held in working until after *bit* is processed. Both of these fillers will be held in working memory at the same time until at least *that the man owned* is finished being processed (see Figure 1 below).

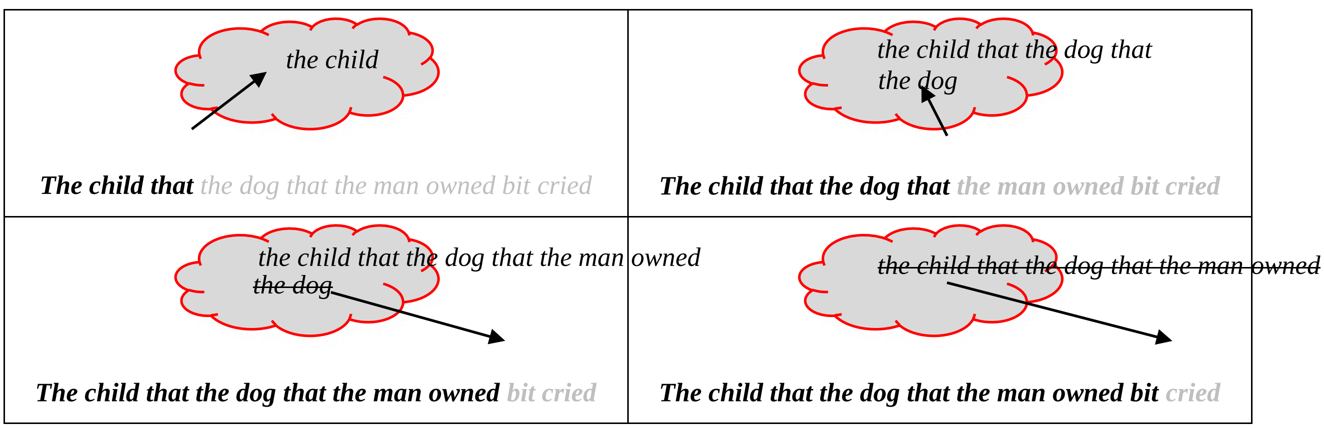


Figure 1. The processing of (2) demonstrating how two fillers must be held in working memory. (Bold text has been processed by the parser, and the gray cloud represents working memory.)

Contrastingly for both (3-4), only one filler is ever kept in working memory at any given time. Thus, two factors might together overload working memory and cause the unacceptability of (2):

factor one—greater distances between filler and gap;

factor two—multiple fillers held in working memory simultaneously.

Such syntactic burdens on working memory can account for the processing difficulty and even unacceptability associated with filler-gap dependencies like those,[2] but can they also account for the unacceptability of island constraints? Consider Ross' (1967) Complex Noun Phrase Constraint demonstrated by (5) below (borrowed from Hofmeister & Sag, 2010).

5) **What did he know someone that has?*
   **What did he know someone*₁ [*that has* Ø₁]R-CL1*?*
   greatest distance: n + ≥

Here again, the first filler (*what*) is separated from its gap by a distance analogous to the tremendous distance associated with the filler-gap dependency in (2) above. Moreover, *what* and the second filler (*someone*) are both simultaneously held in working memory at least until *has* has been processed. In other words, island filler-gap dependencies seem to always be associated with the same two syntactic factors that are associated with the potential overload on working memory in (2) above.

Of course, these factors parallel nicely with Chomksy's Subjaceny Condition. That is, Subjacency prohibits movement across two cyclical nodes (DPs and CPs). The syntactic burden of crossing two such nodes is equated with a tremendous distance between the filler and gap as well as, at some point, two fillers being held in working memory simultaneously. Moreover, Chomsky's Barriers account seems to further support the effect of these two syntactic factors on working memory. That is, Barriers suggests that an increase in the number of barriers crossed will decrease the acceptability of a filler-gap dependency or, in terms used here, an increase in the syntactic distance increases the burden on working memory thereby degrading acceptability.

The effect of these syntactic burdens on working memory during processing of center embedding, *wh*-islands, and similarly difficult filler-gap dependencies seems to cover similar

[2] Of course, this is not to suggest that other accounts cannot also explain the differences associated with (2-4).

empirical ground as generative constraints, and the working-memory account outlined here need not be mutually exclusive to a generative account. In fact, this account seems to be simply a variant on the generative constraints for filler-gap dependencies, a variant that adopts psycholinguistic terminology to address processing data. For example, the effect of syntactic distance on working memory might better explain the gradations of processing speeds and even acceptability judgments for many speakers. That is, while Barriers can only count to 1, syntactic distance presents a potentially infinite gradation of distinction, which might account for any levels of acceptability beyond good, marginal, or bad—should there turn out to be more than these three levels (see Hofmeister & Sag, 2010).

Appendix 1

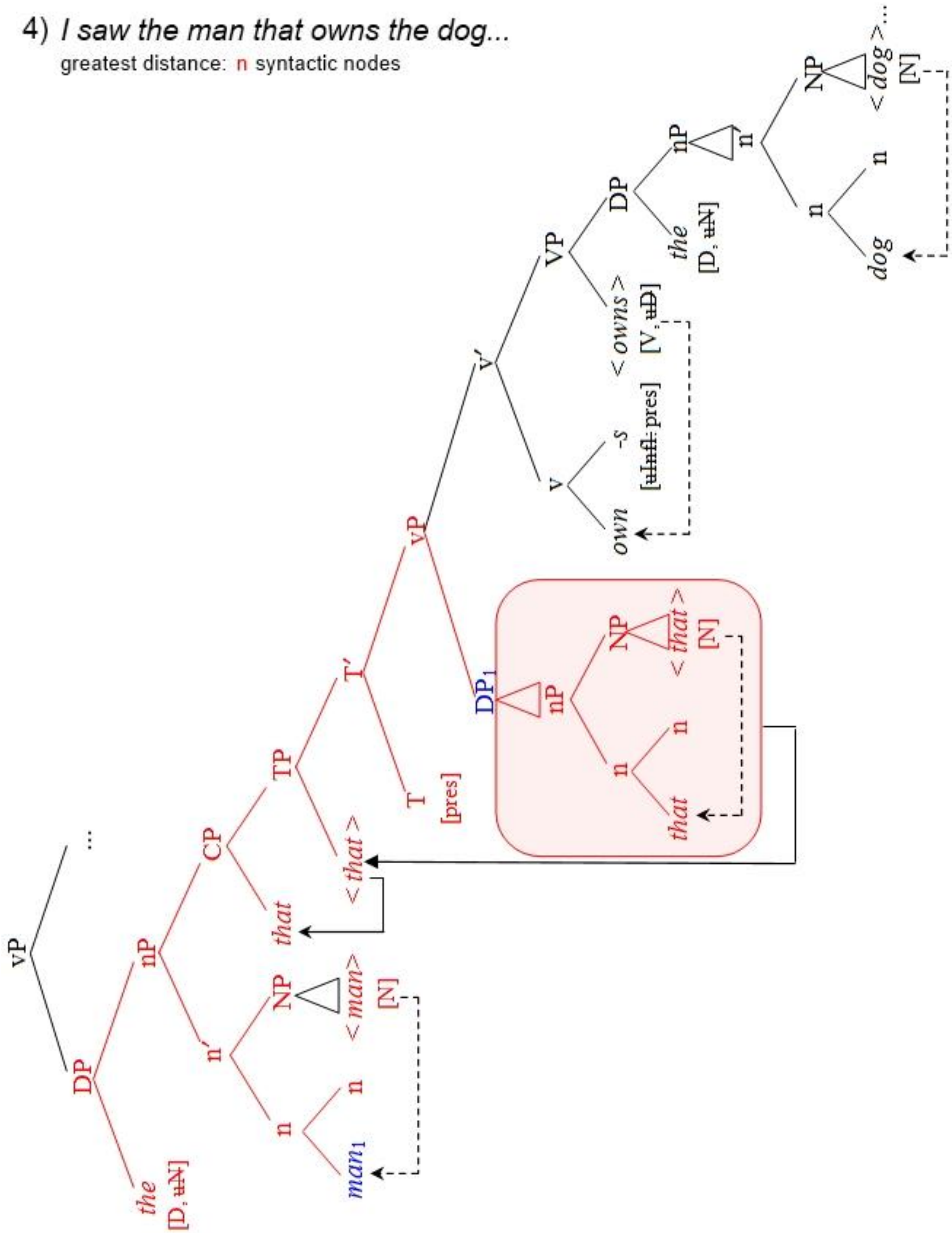


Diagram 1. Diagram for example (4) above. Distance from filler (*the man*) to gap in red; greatest distance = n.

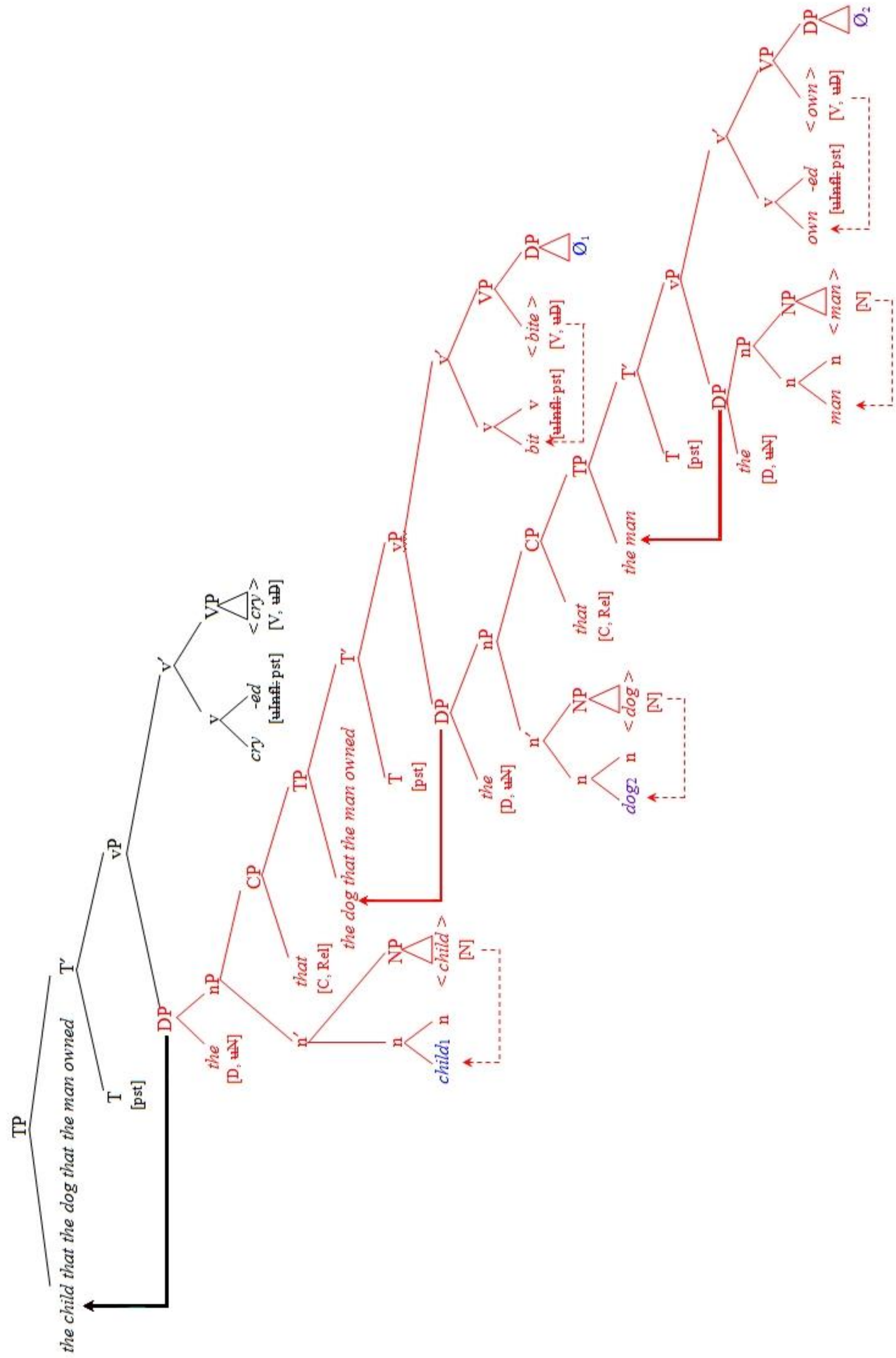


Diagram 2. Diagram for example (2) above. Distance from first filler (*the child*) to gap in red; greatest distance = n ≥45.

References

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122-149.

Dickey, M. W. (1996). Constraints on the sentence processor and the distribution of RPs. *Linguistics in the Laboratory, 19*, 157-182.

Hofmeister, P., & Sag, I. (2010). Cognitive constraints and island effects. *Language, 86,* 366–415.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30*, 580–602.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*(1), 35–54.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language*, *47*(1), 69–90.

A binary features approach to person and number asymmetry

Kelly Enochson

George Mason University

Abstract

Much syntactic research addresses the asymmetry between person features and number features. Notably, many languages exhibit a Person Case Constraint, while no languages exhibit a Number Case Constraint. In addition, some languages demonstrate Omnivorous Number, while no languages demonstrate Omnivorous Person. Nevins (2011) identifies a relationship between these phenomena, arguing that the conditions of Matched Values and Contiguous Agree fully explain the data. Nevins assumes that person features are fully specified binary features, whereas number features are privative with the unmarked value syntactically underspecified. I argue that there is sufficient evidence of the unmarked value of number, namely singular, being active in the syntax that it is inaccurate to represent number with privative feature specification. By assuming binary features for both person and number, it is possible to derive the relationships that Nevins addresses. Specifically, I suggest that the condition of Matched Values is an optional condition on the already relativized probe.

Keywords: features, person complementarity, omnivorous number, underspecification

A binary features approach to person and number asymmetry

1. Introduction

The syntactic difference between person features and number features has been widely discussed in the literature (Preminger, 2011; Baker, 2011; Nevins, 2011; Wechsler, 2011, *inter alia*). The Person Case Constraint (Bonet, 1991) prohibits unmarked (third) person values from dominating marked (first or second) person values. However, there is no equivalent phenomenon for the number feature. Additionally, several languages exhibit Omnivorous Number (Nevins, 2011), in which the presence of a plural feature on either subject or object (or both) will result in the presence of the plural agreement morpheme, causing potential ambiguity. However, there is no equivalent phenomenon for the person feature. In addition, Baker (2011) identifies that adjective agreement occurs for number and gender, but not for person.

Nevins (2007, 2011) argues that the conditions of Matched Values and Contiguous Agree account for phenomena such as the Person Case Constraint and Omnivorous Number, as long as one assumes that person features are fully specified in the syntax, and that number features are privative. I argue that number must also be fully specified in the syntax, and that the syntactic difference between person and number features can be attributed to Nevins' Matched Values principle, which I suggest is an optional condition on the probe.

In this paper, I begin by reviewing Nevins' (2007, 2011) conditions of Matched Values and Contiguous Agree (section 2). Then I show that number cannot be syntactically underspecified (section 3). Finally, I attempt to explain the asymmetry between person and number features, assuming only binary feature specification (section 4).

2. Matched Values and Contiguous Agree

Nevins (2011) invokes feature specification to explain the difference in syntactic behavior between person features and number features. Person features are binary, meaning that they are fully specified in the syntax using the features [±participant] and [±author]. On the other hand, number features are privative, meaning that the number value [plural] is specified in the syntax but that singular is not. Nevins' system (2007, 2011) is based on two conditions: Matched Values and Contiguous Agree. Nevins (2007) defines Matched Values as a condition that requires that all elements in the domain of relativization contain the same value for whatever feature is being agreed with. Contiguous Agree is a constraint prohibiting intervening elements in an agree domain. Nevins argues that these two conditions, along with the underspecification of number features, yield the attested linguistic results.

The Person Case Constraint (Bonet, 1991) prohibits unmarked (third) person values from dominating marked (first or second) person values. For example, in a language such as French that has pronominal clitics for direct and indirect objects, it is not possible to have a third person indirect object with a first person direct object. There does not exist a comparable rule for number values.

In order to rule out the possibility of a Number Case Constraint, Nevins relies on the number feature being syntactically privative. As he explains, plural number is marked [plural] in the syntax, but singular number is underspecified. A Number Case Constraint would prohibit an unmarked value for number dominating a marked value. For instance, a singular subject with a plural object would be prohibited. The principle of Matched Values explains the Person Case Constraint: violations of the PCC are also violations of Matched Values (3a). There is no Number Case Constraint because an unmarked number value does not appear overtly in the syntax, thus avoiding the problem of violating Matched Values (3b). Nevins makes a distinction between weak PCC, strong PCC, and ultrastrong PCC. For explanatory purposes, I will explain how Nevins' method addresses the weak PCC. The method is the same for the other two versions of the PCC; the difference is in the relativization of the probe.

(1) a. *Person Case Constraint*: the probe is relativized to the marked value of the participant feature [part], in this case βF. NP1 violates Matched Values so the probe cannot agree with NP2. The probe does not show simple agree with NP1 because NP1 has a value of αF rather than βF. Agreement fails.

___________X___________

P	NP1	NP2
[uF]	[αF]	[βF]

b. *no Number Case Constraint*: The probe does not agree with NP1 because its underspecified features make it invisible to the probe; therefore, it does not violate Matched Values. The probe agrees with NP2.

__________X__________

P NP1 NP2

[uF] Ø [βF]

Matched Values also explains the presence of Omnivorous Number and the absence of Omnivorous Person. Omnivorous Person would violate Matched Values, as shown in (2a); this is not a problem for number because unmarked number is underspecified, as shown in (2b).

(2) a. no Omnivorous Person: the probe is relativized to marked[part], in this case βF. The outcome is the same as (1a).

__________X__________

P NP1 NP2

[uF] [αF] [βF]

b. *Omnivorous Number*: three possible scenarios can result in omnivorous number. NP1 can be underspecified while NP2 is [plural], NP1 can be [plural] while NP2 is underspecified, or both can be [plural]. In the first two cases, one NP is invisible to the probe; therefore, there is no violation of Matched Values. In the third case, the feature values are the same.

__________X__________

P NP1 NP2

[uF] Ø [βF]

[uF] [βF] Ø

[uF] [βF] [βF]

The fact that three different scenarios all result in the same feature values on each NP means that each of these scenarios is potentially ambiguous.

Although Nevins' system accounts for the attested data, there are some problems with assuming privative number features, which I will address in the following section.

3. Problems with privative number

Nevins (2011) presents several phenomena suggesting that singular number must be underspecified, each of which I will address in section 4. However, it is necessary to address one before I continue. Nevins points to attraction effects as an indication that singular number must be invisible in the syntax. Attraction effects occur when a verb shows spurious plural agreement due to the presence of a nearby DP marked with the plural feature. Nevins argues that attraction effects occur for plural number, but not singular, because singular number, being underspecified, is not visible to the probe. He cites the following examples on page 945.

(3) The key to the cabinets are missing.

(4) *The keys to the cabinet is missing.

(5) *The story about me am interesting.

According to Nevins, the first example is attested in casual speech because the presence of [plural] on the DP *cabinets* causes the presence of [plural] on the verb, despite the fact that the verb should agree with *key*, which is singular. This is an example of Omnivorous Number. The ungrammaticality of (4) suggests that [singular] is not specified in the syntax, and therefore cannot be "attracted" to the agreement probe. The ungrammaticality of (5) suggests that attraction effects happen only with number, not with person.

There are occasional productions in English that do not follow Nevins' claims; there are attested examples of attraction effects occurring with singular number. For instance, consider the following.

(6) The views of the Potomac from the house is just perfect.[1]

Example (6) shows what Nevins claims to be impossible; the feature [singular] is attracted to the verb from a DP modifier. The sentence in (6) suggests that number features are not underspecified in the syntax. Bock and Miller (1991) sought to induce agreement attraction effects in an experimental setting, and their results show a significantly greater number of agreement errors with plural number, but they do find a small number of errors with singular number. In order for this type of attraction effect to occur, the feature [±singular] must be a binary feature that is overtly specified. It may be more likely that the value

[1] http://www.yelp.com/biz/george-washington-memorial-parkway-mclean

[-singular] results in attraction effects because this is the marked value of the feature (Preminger, 2011; Nevins, 2007), and some probes prefer to value themselves with a marked feature. However, since it is possible for [+singular] to participate in attraction effects, it must be fully specified in the syntax.

An additional problem for Nevins' account of number, as Béjar (2011) points out, is that Nevins' system does not account for dual number. Béjar (2011) reworks some of Nevins' examples, assuming a privative system for number including relative underspecification, which allows for singular, dual, and plural number. Béjar points out that under Nevins' system, Number Case Constraint effects in dual and plural contexts is predicted to arise even though these effects are not attested. Because privative number specification that includes dual number does not rule out NCC effects with a complex probe, and because attraction effects can occur for singular number, I assume a fully binary specification of both person and number features. Following Harbour (2006), I use the features [±singular], [±augmented] to specify binary number features. Singular number is specified as [+sing] [-aug]. Dual number is represented by [-sing] [-aug], and plural number is [-sing] [+aug]. I will use these binary features to explore the asymmetry in person and number.

4. Person and number asymmetry

Nevins (2011) identifies several instances of asymmetry in the syntactic behavior of person features and number features. He uses the conditions of Matched Values and Contiguous Agree, along with the assumption of underspecification of number, to explain these asymmetries. I will address each of the points that Nevins raises and attempt to show that a binary specification of both person and number features will result in the same behaviors. I assert, additionally, that the condition of Matched Values is a condition on the probe, and need not always apply.

Nevins identifies the presence of a Person Case Constraint, alongside the absence of a Number Case Constraint, as evidence of privative number features. I argue that a binary specification of number allows for the same predictions. A Person Case Constraint arises in the same situation that Nevins describes, shown in (1a), repeated below as (7).

(7) *Person Case Constraint*: the probe is relativized to marked[part], in this case βF. NP1 violates Matched Values; the probe does not show simple agree with NP1 because NP1 does not have the correct feature value to match the relativized probe. Agreement fails.

___________X___________		
P	NP1	NP2
[uF]	[αF]	[βF]

My system does not alter Nevins' analysis of the PCC.

Under my system of binary specification of number, it seems at first glance that a Number Case Constraint would be predicted to occur. However, by relativizing the probe in a particular way, this can be avoided. In her response to Nevins (2011), Bejar (2011) states, "no conclusions can be drawn independently of one's assumptions about the feature structure of the probe [F]," (p.988). Nevins' system relies on probes being relativized to different features in different situations. Probes can be relativized to look for all features, for marked features only, or for contrastive features only (Nevins, 2007). He suggests that different probe relativization partly accounts for the parametric variation between languages. My system adds a condition to certain probes, which is well within the scope of probe relativization. I suggest that on certain probes, Matched Values does not apply.

Matched Values applies quite rigidly to person features; hence the PCC and the lack of Omnivorous Person. However, in some situations, Matched Values does not apply to number features. This explains the lack of Number Case Constraint as well as the presence of Omnivorous Number. There is theoretical backing for treating person features differently from other features. Baker (2011) derives the Structural Constraint on Person Agreement (SCOPA) to address the ways in which person features behave differently from number and gender. Preminger (2011) argues that person and number have separate probes. Restricting Matched Values only to person features is a principled theoretical move.

Consider (8), which illustrates the lack of Number Case Constraint.

(8) *no Number Case Constraint*: the probe is relativized to marked[sing], in this case βF. Additionally, Matched Values is not a condition on the probe.

__________X__________		
P	NP1	NP2
[uF]	[αF]	[βF]

The probe does not agree with NP1 because the probe is seeking the feature value βF. NP1 does not violate Matched Values because Matched Values is not a condition on the probe. Ignoring Matched Values results in the mismatching feature value on NP1 being invisible to the probe. The result is simple agree with NP2. This is an instance of Omnivorous Number.

By the same condition on the probe, I am able to explain all three examples of Omnivorous Number discussed in (2b).

(9) *Omnivorous Number*: the probe is relativized to marked[sing], in this case βF. Matched Values is not a condition on the probe.

__________X__________		
P	NP1	NP2
[uF]	[αF]	[βF]
[uF]	[βF]	[αF]
[uF]	[βF]	[βF]

The first instance is the same situation as the example discussed in (8). In the second instance, simple agree occurs for the same reason as in the first: the probe searches for the feature value βF and is not concerned with Matched Values. In the third instance, both NP1 and NP2 agree with the probe. All three occurrences result in Omnivorous Number and, consequently, in potential ambiguity.

Restrictions on Omnivorous Person are the same under my system as under Nevins' system, discussed in (2a) and repeated here in (10).

(10) *Omnivorous Person*: the probe is relativized to marked[part], in this case βF. The probe also has the condition Matched Values. NP1 violates Matched Values; therefore, agreement fails.

__________X__________		
P	NP1	NP2
[uF]	[αF]	[βF]

By assuming that Matched Values is an optional condition on the probe, I am able to account for the presence of the PCC and Omnivorous Number as well as the lack of a Number Case Constraint or Omnivorous Person.

Although Nevins (2011) focuses much of his paper on the PCC and Omnivorous Number, he also cites three additional examples illustrating the asymmetrical behavior of person and number features. He identifies attraction effects, which I addressed in section 3, predicate adjective agreement, and the expressive use of number agreement as examples of asymmetrical syntactic behavior. He suggests that these examples illustrate the need to underspecify number. I will address each of these in turn to show how specifying number with binary features does not yield a different result from Nevins' system.

As I discussed in section 3, Nevins asserts that attraction effects occur only for plural number. I have shown that singular number can also show attraction effects, indicating that it is syntactically specified. Nevins also demonstrates that attraction effects occur only for number features, not for person features. The condition of Matched Values operates on the probe relative to person features, prohibiting attraction effects for person. It is not a condition on the probe relative to number features, allowing attraction effects to occur for number.

Nevins identifies partial agreement on predicate adjectives as demonstrating further evidence that person and number features must be specified differently. Nevins references Baker's (2011) observation that adjectives show number agreement but not person agreement. Baker observes that gender features tend to behave the same way as number features, in that adjectives often show gender agreement as well as number agreement. Person agreement is specifically blocked, suggesting that something about person features is different from other features. Consider Baker's (6) from p.879, repeated here as (11). The predicate adjective 'fat' can agree in gender and number, but not in person.

(11). a.	Est-as	mujer-es	son	gord-*as*.	(Spanish)
	these-F.PL	women(F)-PL	are.3pS	fat-*F.PL*	
	'These women are fat.'				
b.	El	hombre	*es*	gord-*o*.	
	the.M.SG	man(M.SG)	is.3pS	fat-*M.SG*	

'The man is fat.'

c. (Nosotras) somos gord-*as* / *gord-*amos*

we.F.PL are.1pS fat-*F.PL* / *fat-*1p*

'We (a group of females) are fat.

Wechsler (2011) attributes the difference between person agreement and number agreement on adjectives not to the features, but to the goals. He suggests that adjectives show agreement with concord phi features, while nouns show agreement with index phi features. Person is only part of the index feature bundle, not the concord feature bundle. Although the feature bundles are different, the two types of features need not be specified differently. Using binary features instead of privative features does not affect this analysis.

Lastly, Nevins addresses the issue of expressive use of agreement. Plural number can be used on a verb expressively along with a syntactically singular but semantically plural subject (Reid, 2011). For example, a collective noun can be used with a plural verb, as in (12).

(12) The team are thrilled to be in the championship.

According to Nevins, "Arguably, this sort of mechanism is allowed precisely because there is no "overwriting" of a singular feature. The [plural] feature may be added or subtracted without the need to manipulate or specify singular in the syntax" (p.945). This phenomenon can also be accounted for under my system. In the case of expressive agreement, the probe is relativized to the marked value of singular, [-sing]. Failing to find that value on the NP, the probe matches *features* with the NP, but does not match values. Instead, the probe, which is relativized to the marked value, becomes valued with the marked value.

5. Conclusion

Nevins (2011) identifies several phenomena that illustrate an asymmetry between person and number features. He relies on the underspecification of number, along with the binary specification of person, to account for these phenomena. In this paper, I have shown that number features cannot be underspecified and must be binary. However, assuming that both person and number features are fully specified in the syntax, I am able to achieve the same predictive results as Nevins. I achieve this by assuming an additional property of the already relativized probe; namely, that Matched Values is an optional condition of the probe and need not apply for number features.

References

Baker, M. 2011. When agreement is for number and gender but not person. *Natural Language and Linguistic Theory 29*: 875-915.

Béjar, S. 2011. Remarks on omnivory and complementarity: a commentary on the paper by Andrew Nevins. *Natural Language and Linguistic Theory 29*: 973-997.

Bock, K. and Miller, C. 1991. Broken Agreement. *Cognitive Psychology* 23, 45-93.

Bonet, E. 1991. Morphology after syntax: Pronominal clitics in Romance. Doctoral Dissertation, MIT.

Den Dikken, M. 2011. Phi-feature inflection and agreement: an introduction. *Natural Language & Linguistic Theory 29:* 857-874.

Harbour, D. 2006. *Morphosemantic number: From Kiowa noun classes to UG number features*. Dordrecht: Springer.

Nevins, A. 2011. Multiple agree with clitics: person and complementarity vs. omnivorous number. *Natural Language and Linguistic Theory 29*: 939–971

Nevins, A. 2007. The representation of third person and its consequences for person-case effects. *Natural Language and Linguistic Theory 25*: 273-313.

Preminger, O. 2011. Asymmetries between person and number in syntax: a commentary on Baker's SCOPA. *Natural Language and Linguistic Theory 29:* 917-937.

Reid, W. 2011. The communicative function of English verbal number. *Natural Language & Linguistic Theory 29*: 1087-1146.

Wechsler, S. 2011. Mixed agreement, the person feature, and the index/concord distinction. *Natural Language and Linguistic Theory 29*: 999-1031.

Aspir(at)ing to Speak Like a Native: Tracking Voice Onset Time in the Acquisition of English Stops

Mike Klein
George Mason University

Abstract

Much attention has been paid in recent decades to the development of new phonetic categories in the process of second language acquisition. Empirical studies, many of them focusing on voice onset time (VOT) in the acquisition of stops, have abounded, and a number of theories have been advanced to account for the data uncovered in these studies. In this paper I present the results of a corpus study focusing on VOT in the production of English stops by native Mandarin, Spanish, and French speakers who acquired English at different ages and possess differing lengths of residency in an English-speaking area. Data for the native Mandarin and French speakers (although not the native Spanish speakers) strongly suggest that learners with a significant period of residency tend to produce more authentic English voiced stops. Three theories of phonological learning – the Speech Learning Model (Flege, 1995), Full Transfer/Full Access (Escudero & Boersma, 2004; Schwartz & Sprouse, 1996), and Statistical Learning (McMurray, Aslin, & Toscano, 2009) – are evaluated in light of these new data; the results strongly favor Full Transfer/Full Access and present a new challenge to the Speech Learning Model.

1. Introduction

When speakers use a language they learned after childhood, native-speaking interlocutors generally have no trouble detecting their “foreign accents.” This is true even of advanced learners with flawless syntax and a robust lexicon: phonology, it seems, gives them away every time (Scovel, 1995). While there are undoubtedly numerous acoustic cues in a given speech signal that allow us to detect a foreign accent, it has long been recognized that voice onset time (VOT) in the production of stops is particularly salient – and particularly amenable to measurement and analysis. VOT is a measure of the relative timing of two articulatory gestures associated with stop consonants: the release of the stop and the start of vocal fold vibrations (in other words, voicing). Positive values of VOT are found when voicing begins after a stop is released and thus correspond to aspirated plosives; likewise, negative VOT values indicate that voicing begins before a stop is released and are therefore associated with (partially or fully) voiced stops.

Experimental studies have shown that an L2 (second language) speaker’s realization of VOT is closely correlated with native speakers’ perceptions of foreign accent. Major (1987) measured VOT values for items read from a word list by adults learning English as a foreign language in Brazil and also had native English speakers rate their accents using recorded speech samples. He found a tight correlation: “the higher the accent score [i.e. the more native-like a speaker is rated], the closer the VOT conforms to the American English norm” (p. 199). Using almost the same methodology, Flege and Eefting (1987) achieved similar results using a group of 50 adult learners of English from the Netherlands, again finding a correlation between VOT and perception of foreign accent. These findings were confirmed yet again by Riney and Takagi (1999), this time by way of a long-term study of Japanese speakers of English as a foreign language. While a full understanding of all the factors that lead to the perception of a foreign accent remains an ongoing research program, it is safe to say that VOT is a major contributor; as Riney and Takagi conclude, “in L2 pronunciation there is a basic correlation between GFA [global foreign accent] and VOT” (Riney & Takagi, 1999, p. 298). Correlation, of course, does not prove causation; other factors could be just as important in the detection of a foreign accent. However, these studies suggest that VOT can stand as a convenient standard-bearer for foreign accent in general.

The central role of VOT in studies of the acquisition of L2 phonology is the result of at least three factors. First, as mentioned above, it is a relatively straightforward matter to measure VOT using a laptop computer armed with acoustic analysis software like Praat (Boersma & Weenink, 2012). Second, as we have seen, the study of VOT in L2 production can serve as a convenient, more concrete proxy for the study of foreign accent in general, a much trickier thing to evaluate or even define. Finally, the phonological systems of some very commonly studied languages differ significantly in the arrangement of their stops with regard to VOT, making it easy to find subjects whose L1 and L2 feature phonemes are not likely to match up in VOT. In the United States, for instance, it is typically not difficult to find native Mandarin and Spanish speakers who are learning English, and all three of these languages feature systems of stop phonemes with very different VOT values (see Figure 1). This is fortunate for researchers who want to study the role of VOT in phonological acquisition: without such variety, it would be difficult to pose and answer interesting questions about what language learners do when faced with discrepancies between L1 and L2 phonemes.

Figure 1 shows the arrangement of stop contrasts with regard to VOT in four languages. Since the actual realization of stop phonemes depends on many factors (including, but probably not limited to, place of articulation, register, phonetic environment, and individual variation), Figure 1 is intended to give an impression of the kinds of contrasts that are possible, rather than a guide to actual VOT values in these languages. Each language, for a given place of articulation, contrasts two or three phonemes, but the contrasts are different in each case. Despite their traditional label, English voiced stops /b/, /d/, and /g/ are typically only partially voiced, with small negative VOT values (light blue box); unvoiced stops, at least in word-initial position, are aspirated, with moderately large positive VOT (dark blue box). In Mandarin, all stops (indeed all obstruents) are unvoiced, and the contrast is instead between unaspirated /p/, /t/, and /k/, with small positive VOT (orange box), and aspirated /p^h/, /t^h/, and /k^h/, with large positive VOT (brown box), slightly higher than English voiceless stops. The Spanish system contrasts fully voiced /b/, /d/, and /g/, with large negative VOT (light green box), with unvoiced and unaspirated /p/,

/t/, and /k/ (dark green box), similar to those in Mandarin. (The French system, also relevant for this study, is essentially the same as Spanish.) Finally, Thai boasts a three-way contrast between fully voiced /b/, /d/, and /g/ (pink box), unvoiced and unaspirated /p/, /t/, and /k/ (lavender box), and heavily aspirated /pʰ/, /tʰ/, and /kʰ/ (purple box).

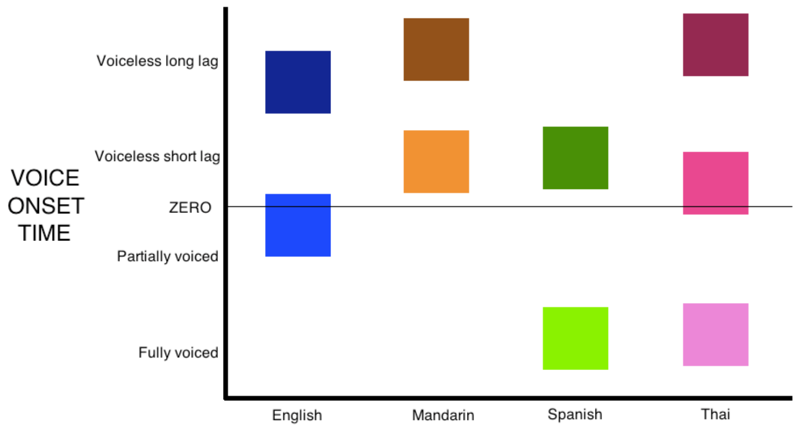


FIGURE 1. Stop contrasts in English (Lisker & Abramson, 1964), Mandarin (Chao & Chen, 2008), Spanish (Williams, 1977), and Thai (Gandour, Petty, Dardarananda, Dechongkit, & Mukngoen, 1986).

The general research question addressed in this paper is this: what happens when adults learn a language with a phonological system featuring stop phonemes with VOT values not found in their native language? More specifically, can adult learners improve their L2 pronunciation with experience, producing L2 stops with VOT values more closely approximating native norms? This paper, of course, is not the first to address VOT in second language acquisition; on the contrary, there has been a thriving literature on the subject for several decades. To the best of my knowledge, however, no study has focused on the possibility that production of L2 stops might change systematically over time as adults learn a second language. As we will see, the answer to this question has important ramifications for second language phonological theory, since different theories of phonological acquisition make different predictions with regard to the possibility that VOT values change over time for language learners.

In the next section of this paper, I will review several experimental studies that have been published on VOT and on second language phonology in general. I will present these studies in the context of three theories of phonological acquisition: Flege's Speech Learning Model (Flege, 1995), probably the gold standard in the field; Full Transfer/Full Access, introduced by Schwartz and Sprouse (1996) in the domain of syntactic acquisition but since extended to phonology as well (Escudero & Boersma, 2004); and the Statistical Learning Model as formalized by McMurray et al. (2009), who provide a model of L1 phonological acquisition that is readily extendible to L2 acquisition. In Section 3, I present the results of an extensive corpus study of English speech samples recorded by native speakers of Mandarin, Spanish, and French. The corpus used here is the George Mason University Speech Accent Archive (Weinberger, 2012). With more than 1,700 speech samples accompanied by demographic information on each contributor (including age of English onset and length of residency in an English-speaking country), the Archive is ideally suited to address, by way of a semi-longitudinal study, the possibility that learners can improve their pronunciation by producing VOT values closer to native English norms as they gain experience. In Section 4, I present a general discussion of the experimental results, evaluating each of the previously discussed models of L2 acquisition in light of these new data. Finally, Section 5 provides a brief conclusion to the paper.

2. Three Models of Phonological Acquisition

In this section, I will introduce three prominent theoretical models of phonological acquisition. In the course of discussing these models, I will review a number of experimental studies that have been published regarding second language phonological acquisition, with special emphasis on studies that have shaped our thinking on VOT in the acquisition of L2 stop contrasts. In a few cases I will highlight the advantages of my experimental design compared to previous studies that did not attempt to systematically deal with the possibility that learners' realization of L2 stops can change over time. I will end this section with a discussion of the differing behavioral predictions of each model with regard to the corpus study presented later in this paper.

2.1. Speech Learning Model

The Speech Learning Model (SLM), discussed and developed throughout the decades-long research program of James Emil Flege and fleshed out most fully in Flege (1995), attempts to explain why L2 learners are not as successful as L1 learners at acquiring the phonetic details of a particular language and why certain sounds are harder to acquire than others. The SLM ultimately attributes these differences to the auditory processing limitations of late L2 learners, rendering them less able than early learners to perceive and internalize certain phonetic contrasts. Young learners are able to perceive all of the phonetic contrasts that are relevant to the phonological systems of the world's languages. Accordingly, these learners have no problem acquiring the phonetic categories of any language they learn, no matter how similar or different these categories are to those of a previously learned language. On the other hand, later learners are increasingly unable to perceive (without explicit training; see Tees & Werker (1984)) some of the finer contrasts between L1 and L2 sounds that they encounter, by hypothesis making it difficult for them to produce native-like phonetic categories in their second language. Where L2 sounds differ significantly from sounds found in their L1, no learners are predicted by the SLM to have difficulty in forming new categories. It is precisely those L2 sounds that are similar but not identical to L1 sounds that present problems for adult learners, who are likely to assimilate them into their existing L1 categories rather than form new L2 categories. According to the SLM, fossilization is a good possibility here: due to these auditory processing limitations, learners will never be able to break free from the phonetic details of their first language phonetic category, resulting in a persistent "foreign accent."

Flege (1980) was the first study within the budding SLM model to tackle the question of VOT in the acquisition of L2 stops. In this study, Flege examined various aspects of the production of voiceless stops (/p/, /t/, and /k/) among Saudi Arabic speakers learning English in the United States. He divided them into two groups of six participants each; every member of the first group (Ar1) had spent less than one year in the U.S., while members of the second group (Ar2) had all spent more than a year in the U.S., with an average of 39 months of residency. He asked them, as well as a native English-speaking control group (Am), to produce CVC words in a carrier sentence and measured VOT for the initial stops in these words. His results are summarized in Table 1.

	Am	Ar2	Ar1
Voice-onset time			
/p/ in *pat*	46	21	14
	(4)	(11)	(10)
/t/ in *tab*	62	29	32
	(11)	(14)	(10)
/k/ in *cab*	67	47	41
	(12)	(11)	(7)

TABLE 1. Summary of VOT results from Flege (1980).

As can be seen from Table 1, which shows the average VOT and the standard deviation (in parentheses) for each group and each phoneme, the two native Arabic groups differed markedly from the American control group, but did not differ significantly from each other. Native Arabic norms for VOT are smaller than English values, around 37 ms for /t/ and 52 ms for /k/ (Flege, 1979); the Arabic speakers in this study were therefore producing sounds much more like L1 Arabic than L2 English. Note that the very low VOT values for /p/ in Table 1 are likely a result of the fact that learners are actually associating English /p/ with Arabic /b/, since most varieties of Arabic have no voiceless bilabial stop /p/.

Although Flege's study does attempt to deal with the possibility of change over time by contrasting the two Saudi groups, his null result – that there is no significant difference between those who had been in the U.S. less than one year and those who had a longer length of residency – is not particularly persuasive. First, there were only six subjects per group, arguably not a large enough participant base to draw any reliable conclusions. Second, even those in group Ar2, with a longer length of residency, averaged only 39 months in the U.S.; consequently, if learners require more than a few years of exposure to shift their VOT towards native English norms, Flege would have been unable to detect those changes. Finally, acknowledging that the results were not statistically significant, the differences in the mean VOT for /p/ and /k/ were moving in the right direction with more experienced speakers producing longer (more English-like) VOTs; although this was not true for /t/, the difference between the two groups was very small in this case. In the corpus study to be presented in this paper, I hope to improve on Flege (1980) by tracking

VOT over a longer period of time (in some cases several decades) among a larger participant pool (about 50 subjects for each L1 group).

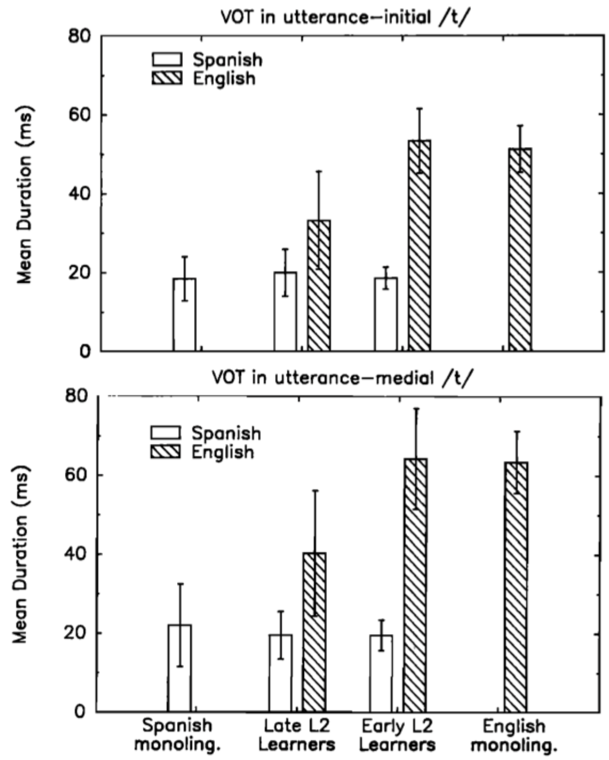


FIGURE 2. Summary of VOT results from Flege (1991).

Flege (1991) revisited the question of VOT in L2 acquisition, this time comparing a group of L1 Spanish early learners of English (age of onset less than seven years) with a group of late learners, also native Spanish speakers. Monolingual Spanish- and English-speaking control groups were also included in his experimental model. He focused only on the phoneme /t/, having participants read sentences like *take a textbook* (or *tengo un tigre*) and measuring VOT for /t/ in utterance-initial and utterance-medial position. His results are summarized in Figure 2.

As can be seen, monolingual Spanish speakers produced tokens of /t/ with an average VOT of around 20 ms, while monolingual English speakers averaged 50 to 60 ms. Early L2 English learners, in accord with the predictions of the Speech Learning Model, were able to fully acquire English /t/, with average VOT values indistinguishable from monolingual English speakers. Interestingly, however, late L2 learners produced VOT values that were, on average, intermediate between Spanish and English norms, hovering around 30 to 40 ms.

To explain this result within the SLM, Flege appeals to “realization rules” (Flege, 1991, p. 407), which allow an individual speaker to pronounce a given phoneme differently in different situations. These rules are independently needed to account for the apparently universal fact that speakers modify their pronunciation as a function of things like rate of speech and social context (or in experimental contexts; see Goldrick (2004); Nielsen (2011); Kirov & Wilson (2012)). Flege believes that these same kinds of rules allow L1 Spanish learners of English to supply longer (but not quite native-like) VOT values for /t/ when speaking English than they do when speaking Spanish, despite the fact that (by hypothesis within the SLM) there is only one phonetic category to cover both languages. While this explanation is not implausible, it is not the only possible explanation for the intermediate VOT values produced by late English learners in this study. If, in fact, VOT values steadily improved with experience so that inexperienced learners produced very Spanish-like stops while more seasoned learners produced more English-like tokens, we would expect precisely the kind of result shown in Figure 2. The average VOT for late learners would naturally fall between that of native English and native Spanish values, since it would reflect the behavior of speakers all along the continuum from L1-like to L2-like production of /t/. It is precisely this possibility that will be explored later in this paper.

Most recently, still working within the Speech Learning Model, Gonzalez Lopez (2012) examined a group of sixteen native English-speaking college students who were intermediate learners of L2 Spanish. These subjects were asked to read sentences in English, Spanish, and a mixture of the two languages. This last sentence type, illustrated in Figure 3, was included because part of Gonzalez Lopez’ research question (not directly related to the subject of the present paper) involved the interaction between L1 and L2 sounds when subjects were forced to switch quickly between them.

a. Spanish → English CS
Pre-switch /t/ Switch /t/ Post-switch /k/
/**t**/odos mis amigos | /**t**/alked Spanish as /**k**/ids.
'All my friends talked Spanish as kids.'

b. English → Spanish CS
Pre-switch /t/ Switch /t/ Post-switch /p/
The /**t**/yphoon damaged | /**t**/echos y /**p**/aredes.
'The typhoon damaged ceilings and walls.'

FIGURE 3. Bilingual sentences in Gonzalez Lopez (2012).

Gonzalez Lopez measured VOT for tokens of Spanish and English voiceless stops (/p/, /t/, and /k/) contained in these sentences. Her results for the monolingual Spanish and English sentences are shown in Figure 4; Figure 5 shows the results for the mixed sentences.

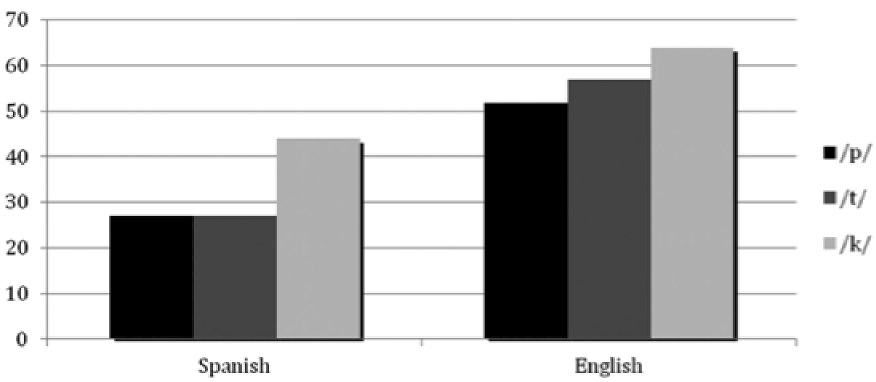


FIGURE 4. Results for monolingual sentences in Gonzalez Lopez (2012).

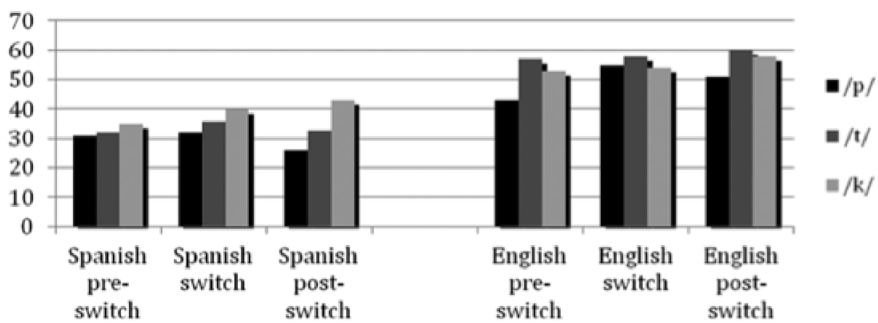


FIGURE 5. Results for mixed sentences in Gonzalez Lopez (2012).

As can be seen from both figures, but especially Figure 4, subjects produced notably lower values for voiceless stops in L2 Spanish than they did in L1 English. Gonzalez Lopez takes this as evidence that subjects are, in fact, able to form new phonological categories for Spanish voiceless stops, although – as can clearly be seen by comparing the configuration of bars in Figure 5 to those in Figure 4 – there was some interaction between L1 and L2 when subjects had to switch mid-sentence. While speakers are clearly treating Spanish and English sounds differently, they are not consistently producing native Spanish-like stops: average VOT for Spanish sounds in Figure 4 are still noticeably above the 20 ms found for native Spanish speakers in Flege (1991) (see Figure 2). Thus, neither Flege's explanation (one phonological category with different realization rules) nor the alternative hypothesized in this paper (a continuum of speakers with more or less native-like VOT values) can be entirely discounted.

2.2. Full Transfer/Full Access

The idea that second language learners initially transfer aspects of their L1 into their L2, but nonetheless are able to slowly approach L2 norms by tapping into the same learning mechanisms available in first language acquisition, was first proposed in the domain of syntax by Schwartz and Sprouse (1996). This theory, known as Full Transfer/Full Access, extends fairly naturally to phonological acquisition, where we know impressionistically that beginning L2 learners typically have very noticeable accents that reflect the phonological structure of their native language, but that pronunciation usually improves over time. Several studies have provided experimental confirmation of this qualitative impression about L2 phonological learning.

Escudero and Boersma (2004) examined the contrast between lax /ɪ/ and tense /i/ in words like *ship* and *sheep* among L1 Spanish learners of two varieties of English, those spoken in Scotland and Southern Britain. While members of both learner groups were ultimately able to acquire the contrast, those whose target language was Scottish English had a much easier time than those who were trying to learn Southern British English, acquiring the contrast more quickly. Escudero and Boersma showed that phonological differences between the two dialects could account for this result; starting with L1 Spanish, learners of Scottish English simply had less distance to cover in learning the contrast, with the direct result that they required less exposure to successfully master it. According to the authors, all learners should be able to acquire the contrast, but some phonological features require more time to acquire due to greater differences between L1 and L2. (Although

Escudero and Boersma do not discuss this possibility, it seems possible – and entirely consistent with Full Transfer/Full Access – that some phonological details might require so much time to acquire that typical adult learners never reach target-like norms.)

Halicki (2010) used a well-formedness judgment task to probe the intuitions of native English speakers with varying levels of experience in L2 French. She found that beginners did not possess native-like phonotactic knowledge, but that more experienced learners increasingly resembled native French speakers in their judgments of novel French-like and non-French-like vocabulary items. Halicki took this to indicate that L2 learners are able to access the same mechanisms as native learners in constructing the phonology of a language; if this is true, then lingering pronunciation problems among even advanced L2 learners may be more about insufficient executive control (Bialystok, Craik, & Ryan, 2006; Rodriguez-Fornells, Balaguer, & Münte, 2006) than incomplete acquisition of the phonological system itself.

As far as I am aware, no previous study has attempted to assess the predictions of the Full Transfer/Full Access model in the realm of VOT in the production of L2 stops. These predictions, however, should be amenable to evaluation using the methodology presented in this paper. We will return to this issue very shortly, in section 2.4.

2.3. Statistical Learning Model

Turning for the moment to the realm of first language acquisition, Maye, Werker, and Gerken (2002) demonstrated that the sound perception of infants was significantly affected by the statistical distribution of the sounds to which they were exposed (See also Maye & Gerken (2000) for an extension to adult participants).

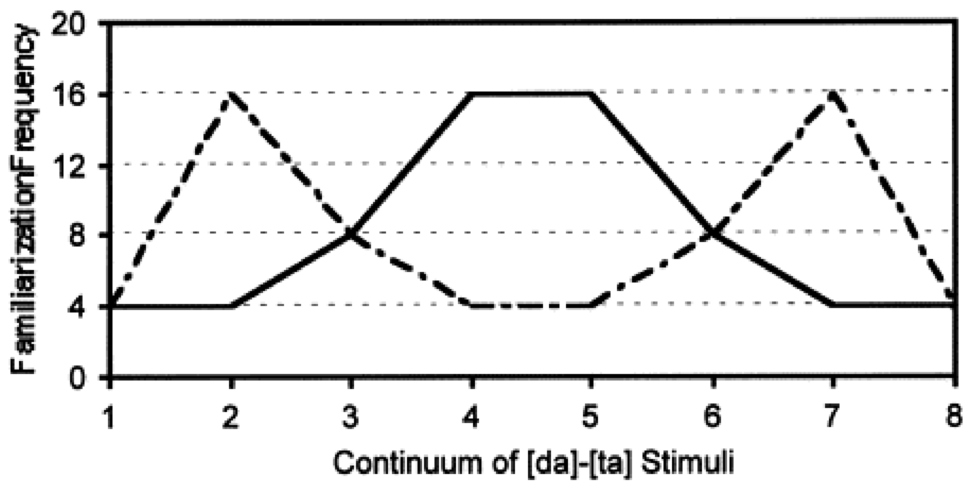


FIGURE 6. Unimodal and bimodal frequency distributions used in May et al. (2002).

Their experiment involved two phases: a familiarization phase, where 6- to 8-month-olds were exposed to either a bimodal or unimodal frequency distribution of tokens on the [da]-[ta] continuum (see Figure 6), and a testing mode, where looking times were used to find out if these young subjects were able to discriminate pairs of words at opposite ends of the continuum. The researchers found that looking time was significantly correlated with familiarization condition, indicating that infants who had been exposed to the bimodal distribution had a much easier time discerning between [d]-like sounds and [t]-like sounds. From a very young age, language learners are apparently highly sensitive to distributional patterns in their language input.

McMurray, Aslin, and Toscano (2009) built on the work of Maye et al. by developing a computational model that was successfully able to acquire the contrast between /t/ and /d/ in English. Their computational learner was supplied with data corresponding to the frequency distribution of VOT for these two sounds in English (Figure 7) and was able to discern two distinct categories, just as L1 learners of English can. This result lends credence to the notion that human language learners are also statistical learners of a sort.

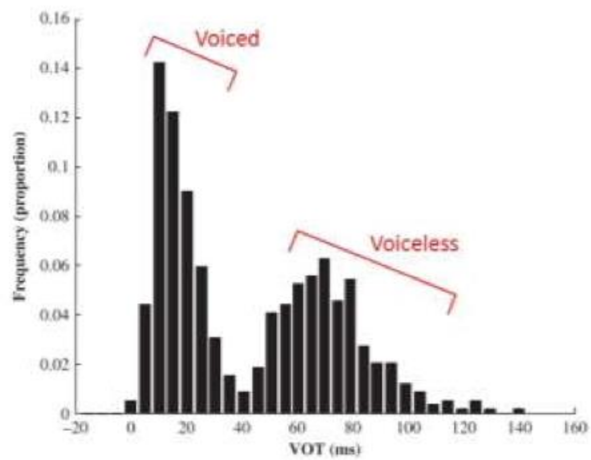


FIGURE 7. Frequency distribution of VOT for English /d/ and /t/, as employed in McMurray et al. (2009).

The Statistical Learning Model has not been explicitly applied to the formation of phonological categories in second language acquisition, but there is no reason it could not be extended to cover L2 as well as L1 acquisition. In the next section, we will do just that and explore the predictions of the Statistical Learning Model in terms of the corpus study to be presented in section 3 of this paper.

2.4. Experimental Predictions

Each of the three models considered in this section makes a different prediction with regard to the results of the corpus study to be presented in the next section of this paper. Recall that this study will focus on the question of whether adult L2 learners can improve their realization of English stops by producing more native-like VOTs with increased L2 experience, specifically with longer length of residency (LOR). Since each of the three models assumes that learners are endowed with different mechanisms for acquiring new phonetic categories, it should be possible to evaluate these models based on the behavior of a large group of adult learners with differing levels of English experience.

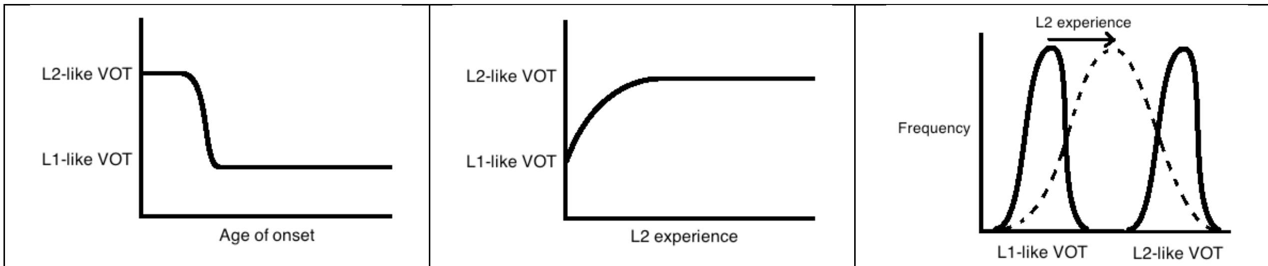
According to Flege's Speech Learning Model, late L2 learners are incapable of developing new phonetic categories for sounds similar to those found in their native language. Since this is exactly the task presented to the language learners to be considered in the corpus study, the Speech Learning Model predicts that age of English onset should be the major factor determining whether speakers are able to produce native-like values of

VOT when speaking L2 English. Early learners should do a very good job producing native English-like stops; later learners are expected to fossilize at L1-like values. This prediction is depicted in Figure 8(a).

The Full Transfer/Full Access model offers a rosier picture for language learners. Although beginners are expected to supply VOT values more in line with their L1 than with L2 English norms (full transfer), they should be able to approach authentic native English behavior with increased experience (full access). This is shown in in Figure 8(b).

Finally, if we extend the Statistical Learning Model to second language acquisition and assume that adult learners as well as young learners have the capacity to form categories based on the statistical distribution of sounds, then it is inevitable that speakers should eventually be able to discern distinct categories for L1 and L2 sounds. This may, however, take quite a bit of time, especially if the peaks in the distribution are fairly close together, as they will be in the cases to be studied here. Before learners are able to establish a new category for L2 sounds, we might expect them to attempt to assimilate them into their L1 category, essentially expanding the old category to accommodate the new sounds until a new category is formed.

The predictions of the Statistical Learning Model, depicted in Figure 8(c), may prove to be very difficult to tease apart from the steady improvement predicted by Full Transfer/Full Access, but there are at least two differences. First, while the Statistical Learning Model predicts that average VOT for L2 stops will, at first, progress steadily from L1-like towards target-like values as the L1 category expands to accommodate new L2 data, this will not be the uninterrupted march predicted by Full Transfer/Full Access. It cannot be, since the category still has to accommodate L1 data as well. Assuming an (obviously idealized) even split between L1 and L2 data, we would never expect average VOT to progress beyond the halfway point between native L1 and L2 values during this intermediate phase, before a new L2 category can be formed. Second, the Statistical Learning Model predicts that, during this initial phase with a single category for L1 and L2 sounds, the production of L1 sounds will be affected as well. This second prediction will not be addressed by the current study, but some previous research has indicated that L1 pronunciation may indeed be affected by the learning of a second language (Pavlenko, 2000).



(a) SPEECH LEARNING MODEL: Only early learners can achieve native L2-like VOT values, so age of onset is the main determiner of VOT in L2 stops.

(b) FULL TRANSFER/FULL ACCESS: VOT values improve steadily from L1-like to authentic L2-like values, given sufficient linguistic experience.

(c) STATISTICAL LEARNING MODEL: L1 category initially expands to accommodate L2 data (dotted line); eventually separate categories are acquired (solid lines).

FIGURE 8. Predictions of the phonological learning models discussed in Section 2.

3. Corpus Study

In order to test my research question – Can adult learners pronounce L2 stops with more authentic VOT values if they spend more time in an English-speaking region? – I collected acoustic and demographic data from the George Mason University Speech Accent Archive (http://accent.gmu.edu; Weinberger, 2012). The Archive contains speech samples from more than 1,700 individuals of varying geographical and linguistic backgrounds, all of whom have submitted an audio recording of the same English paragraph. Demographic information for each subject includes birthplace, native language, other language(s) spoken, age, sex, age of English onset, learning method (academic or naturalistic), country of English-language residence, and length of residency in that country. Many samples feature phonetic transcriptions and an explanation of the prominent features of each subject's pronunciation, although these will not be used in the present study.

My focus will be on L1 Mandarin, Spanish, and French learners of English. For the native Mandarin speakers, I will be interested in the pronunciation of English voiced stop phonemes /b/, /d/, and /g/, which Mandarin speakers are likely to identify with their native voiceless and unaspirated phonemes /p/, /t/, and /k/. The English phonemes feature average VOT values lower than those for the Mandarin phonemes, so my hypothesis (adopting for the sake of simplicity the assumptions of the Full Transfer/Full Access model) is that subjects with more L2 experience (i.e., longer length of residency) will produce lower, and therefore more native-English like, VOT values. For the native

Spanish and French speakers, I will focus on the realization of English voiceless stops /p/, /t/, and /k/ in word-initial position, where they are aspirated and therefore have longer VOT values than the corresponding unaspirated voiceless Spanish and French phonemes. My hypothesis will then be that VOT will increase with L2 experience for the native Spanish and French speakers pronouncing these phonemes. The elicitation paragraph used for each sample on the Speech Accent Archive is given below; the words that will be used in the analysis of L1 Mandarin speakers are shown in red, while those that will be used in the analysis of the Spanish and French speakers are shown in green:

> Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.

The general procedure for each L1 group was to collect relevant acoustic information from the Speech Accent Archive samples using Praat (Boersma & Weenink, 2012), and then construct a series of linear mixed effects models to determine whether the available measure of L2 experience, length of residency, makes a unique contribution to average VOT values. I used the statistical package R (R Development Core Team, 2012) to perform statistical analysis and generate relevant graphs and charts.

In the remainder of this section, I will first describe and analyze the data from Mandarin learners of English, then the data from Spanish learners, and finally the data from French learners.

3.1. L1 Mandarin, L2 English

3.1.1 Participants and Materials

A total of 48 native speakers of Mandarin have submitted samples on the Speech Accent Archive; all of these samples, and the attached demographic information, will be used in the analysis presented in this section. In addition, samples from a randomly selected group of twenty native English speakers will serve as a control group to verify that L1 Mandarin speakers indeed behave differently in the pronunciation of voiced English stops and to allow us to test the hypothesis that these Mandarin speakers behave more like native English speakers with increased L2 experience.

3.1.2 Procedure

Each of the relevant sound files (48 L1 Mandarin, 20 L1 English) from the Speech Accent Archive was captured, as well as each subject's demographic information, which was collected in a spreadsheet. Praat (Boersma & Weenink, 2012) was then used to measure the VOT for the initial stop in each of the four relevant words (*Bob*, *big*, *bags*, and *go*) in each sample; these values were entered into the same spreadsheet. Finally, the data was imported into R in preparation for the statistical analysis to be described in the next section.

When measuring VOT, there were three distinct possibilities for the realization of each stop consonant: negative VOT (i.e., voicing), positive VOT (i.e., aspiration), and zero VOT. Figure 9, a screen shot from Praat, shows an example of negative VOT.

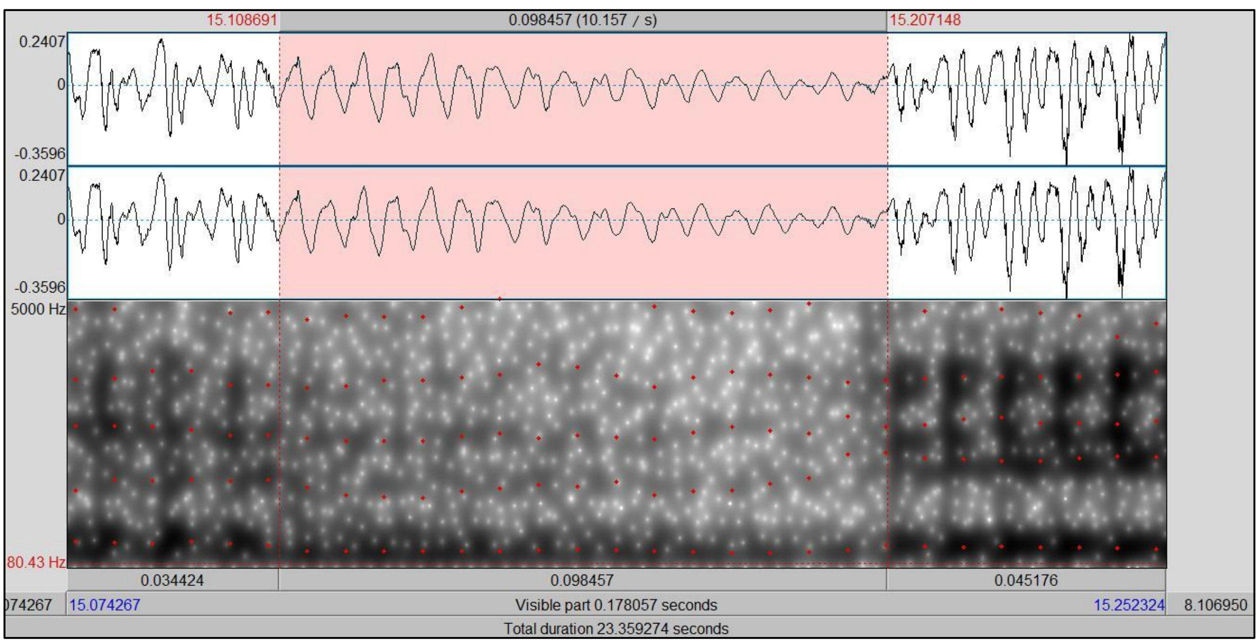


FIGURE 9. Waveform and spectrogram for *big* from subject mandarin33, illustrating negative VOT.

As can be seen by inspecting the waveform (top half of Praat display) in the pink shaded area, periodic vibrations continue throughout the closure between vowel sounds. In addition, the "voicing bar" – the region of low-frequency sound in the spectrogram (bottom half of Praat display) – is also visible throughout closure and is another sign of vocal fold vibration. In this case, the token of /b/ (in *big*) is fully voiced, and VOT is measured from the end of the previous vowel to the beginning of the next vowel (vowels are most easily identified from the more complicated waveform to the left and right of the pink shaded area). This VOT is assigned a negative value, since the start of voicing precedes the release of the stop.

Figure 10 shows an example of positive VOT.

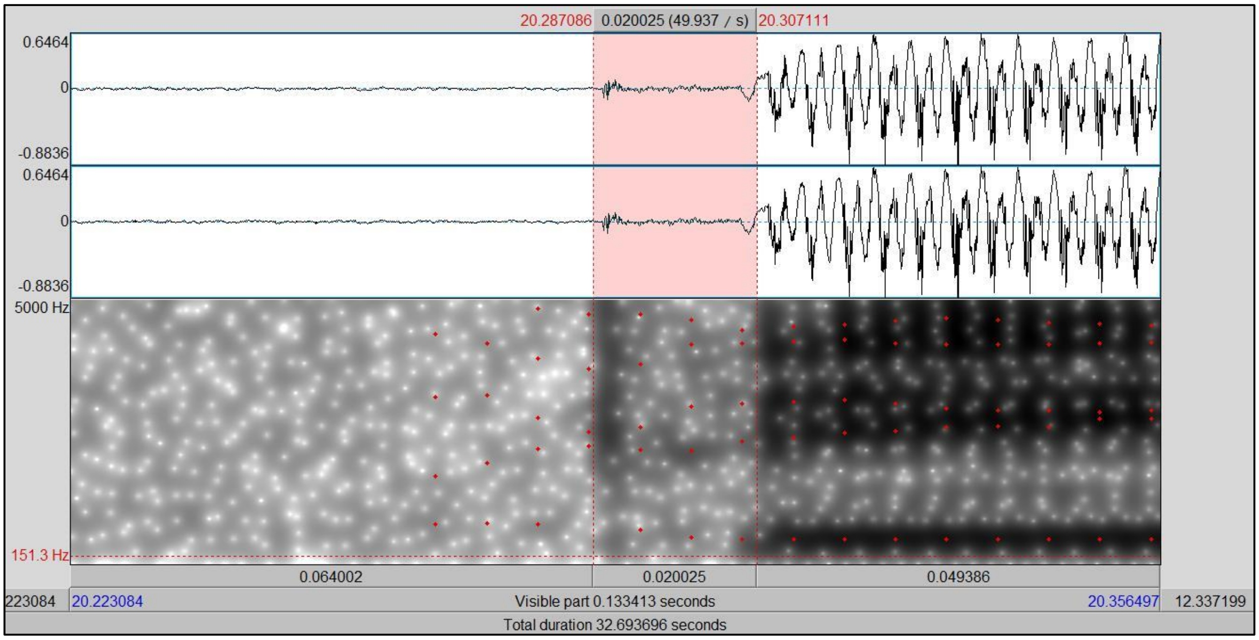


FIGURE 10. Waveform and spectrogram for *big* from subject mandarin18, illustrating positive VOT.

The release burst for /b/ is visible most clearly in the spectrogram as a brief period of "white noise" extending throughout the frequency range (but especially prominent in the higher frequencies); it is also accompanied by a slight perturbation in the waveform. The important matter here is that there is then a discernable lag between the release burst and the beginning of voicing for the vowel, clearly visible here as the mostly empty space after the release burst in the pink shaded region of the waveform and spectrogram. This lag was measured and assigned a positive VOT value, since voicing begins after the stop is released.

Finally, Figure 11 is an example of zero VOT.

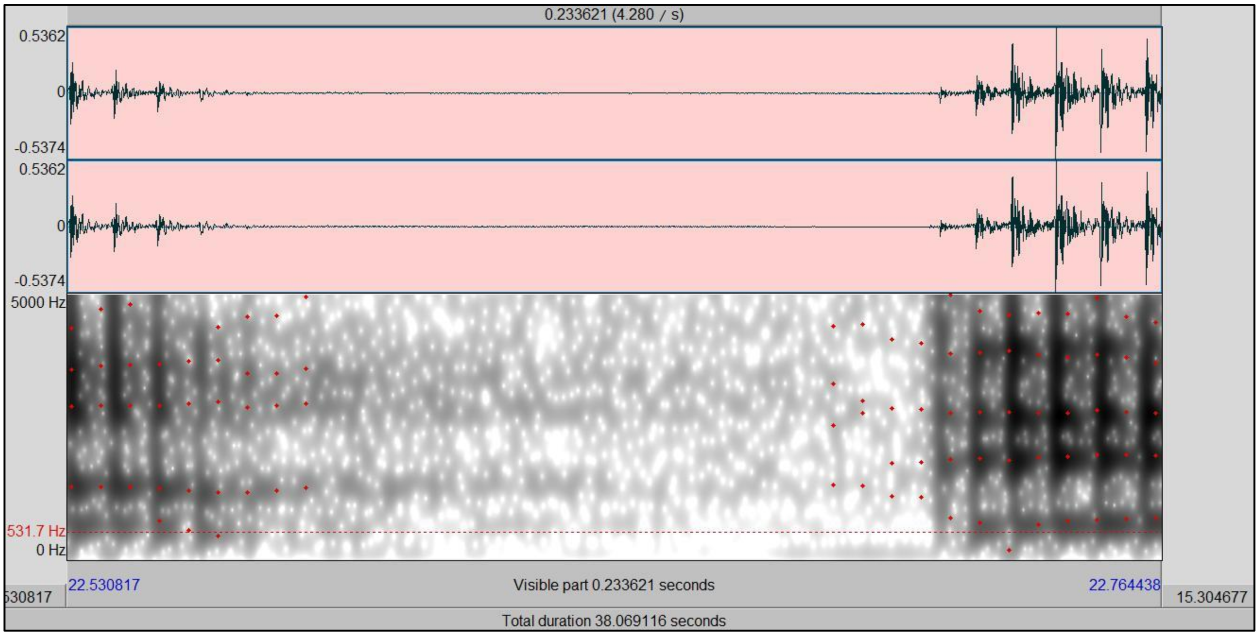


FIGURE 11. Waveform and spectrogram for *big* from subject mandarin30, illustrating zero VOT.

In this case, there is no indication of voicing leading up to the release of the stop: the waveform is completely flat between the two vowels, and no voicing bar is visible in the spectrogram. In addition, there is no evidence of a release burst; the acoustic calm of the unvoiced closure is disturbed only with the beginning of the vowel itself. For this reason, this sample (along with just a few other samples in the data set) was assigned zero VOT.

VOT measurement proceeded in this manner for all of the L1 Mandarin and English samples considered in this section, as well as for the L1 Spanish and French samples to be considered in Sections 3.2 and 3.3.

3.1.3 Results and Discussion

Figure 12 displays the mean VOT values produced by native English and native Mandarin speakers for the initial voiced stop in each word. Table 2 summarizes the descriptive and comparative statistics. (Appendix A contains all of the raw data used in this study.)

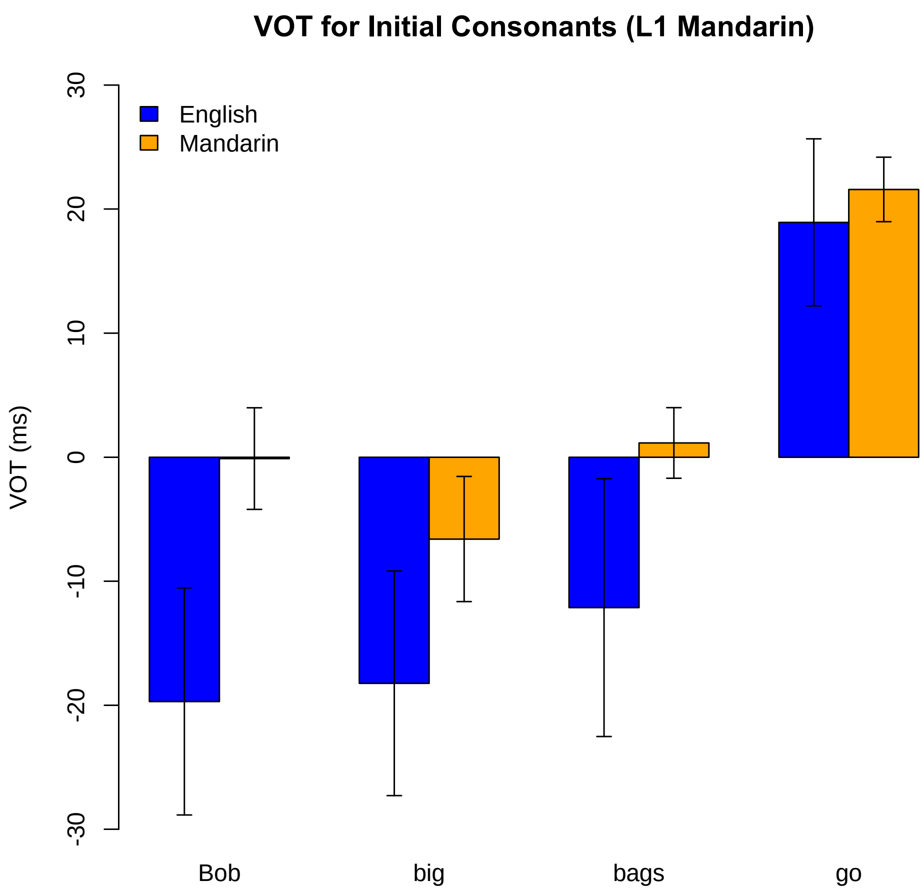


FIGURE 12. Average VOT for initial voiced consonants produced by L1 Mandarin and L1 English speakers.

	Mandarin		English		Welch's t-test		
Word	Mean VOT (ms)	Std. Dev.	Mean VOT (ms)	Std. Dev.	t	d.f.	p
Bob	-0.12	28.38	-19.70	40.93	-1.95	26.94	0.06 (n.s.)
big	-6.61	34.94	-18.23	40.50	-1.12	31.39	0.27 (n.s.)
bags	-1.14	19.75	-12.13	46.48	-1.23	21.92	0.23 (n.s.)
go	21.58	18.01	18.93	28.55	-0.37	22.27	0.72 (n.s.)
ALL WORDS	**4.00**	**28.05**	**-8.47**	**42.04**	**-2.41**	**105.97**	**< 0.02**

TABLE 2. Descriptive and comparative statistics for initial voiced stop data.

In each case, as expected under the assumption that L1 Mandarin speakers associate English voiced stops with their native unvoiced and unaspirated stops, these speakers produce longer average VOT values than native English speakers. Also as expected, there is a great deal of variation among the four words used here; this is most noticeable with *go*, which differs in place of articulation from the other three tokens, but there are differences among the words with initial /b/ as well, most likely due to phonological environment. As can be seen in the right-hand section of Table 2, although there is not sufficient data to find a statistically significant difference between L1 English and L1 Mandarin VOT values for each individual word, we do find a rather significant difference ($p < 0.02$) between the aggregate data sets. Based on this result and the fact that the difference was in the expected direction for each individual token, the analysis in the rest of this section will collapse data for all four words.

Figure 13 shows the relationship between age of onset and average VOT values for each L1 Mandarin speaker. From the figure it is clear that those who began to acquire English at an early age (earlier than about eight years) tend to produce more native English-like voiced stops (recall from Table 2 that the overall mean for L1 English speakers was -8.47 ms). Those who began to learn English later than age eight tend to produce more Mandarin-like stops, with higher average VOT values. To test the predictive power of age of onset, I performed a likelihood ratio test to compare two linear mixed effects models using the *lme4* package in R (Bates, Maechler, & Bolker, 2012). The first model was a "null" model, with just speaker and word (i.e., *Bob*, *big*, *bags*, or *go*) as random effects and no fixed effects at all; the second model added age of onset as a fixed effect. The difference between these two models was significant ($\chi^2 = 6.26$, $p < 0.02$), confirming that age of onset is a significant predictor of performance. (Detailed results for these linear mixed effects models, and those to be discussed below, are contained in Appendix B.)

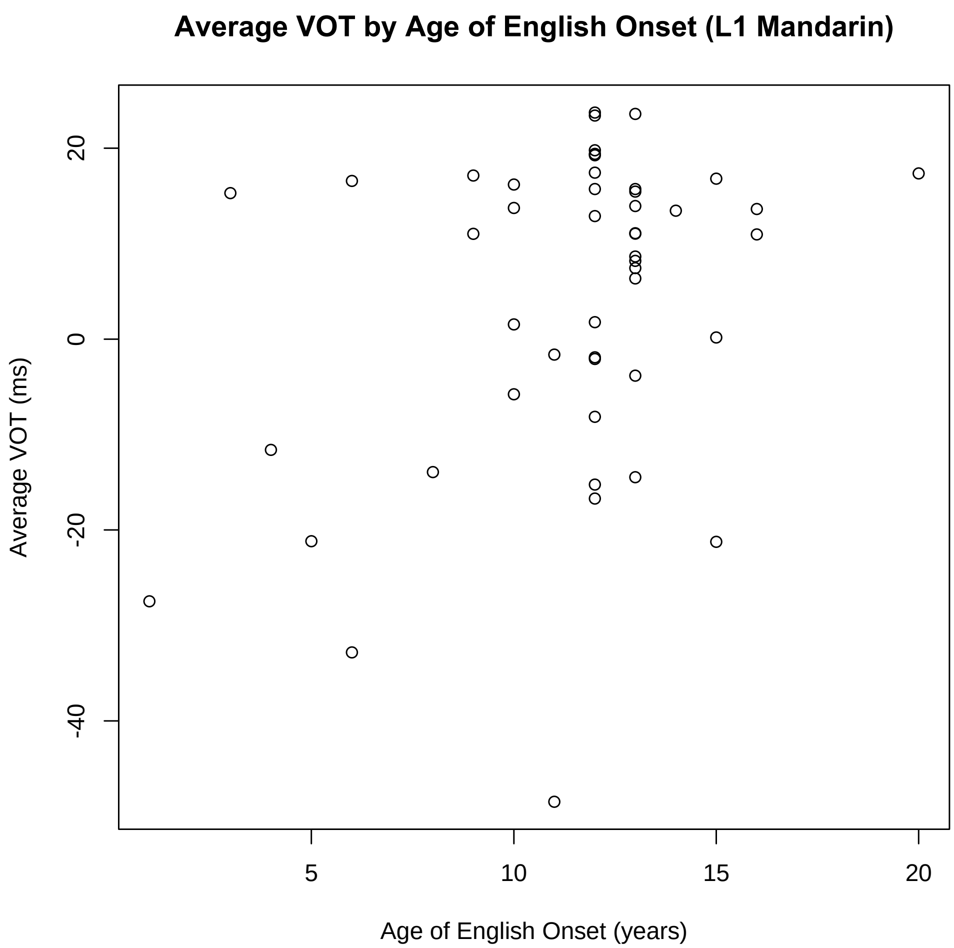


FIGURE 13. Age of onset and production of VOT by L1 Mandarin speakers.

Figure 14 shows the relationship between length of residency (LOR) and production of VOT for voiced English stops. Subjects with very long length of residency – more than about twenty years – tended to produce more English-like voiced stops with negative VOT, while those with shorter LOR produced more Mandarin-like stops. To test the predictive power of length of residency, I added LOR to the linear mixed effects model (along with age of onset) and compared it to the previous model, with only age of onset as a fixed effect. The likelihood ratio test confirms that LOR is a highly significant predictor above and beyond age of onset ($\chi^2 = 12.45$, $p < 0.001$).

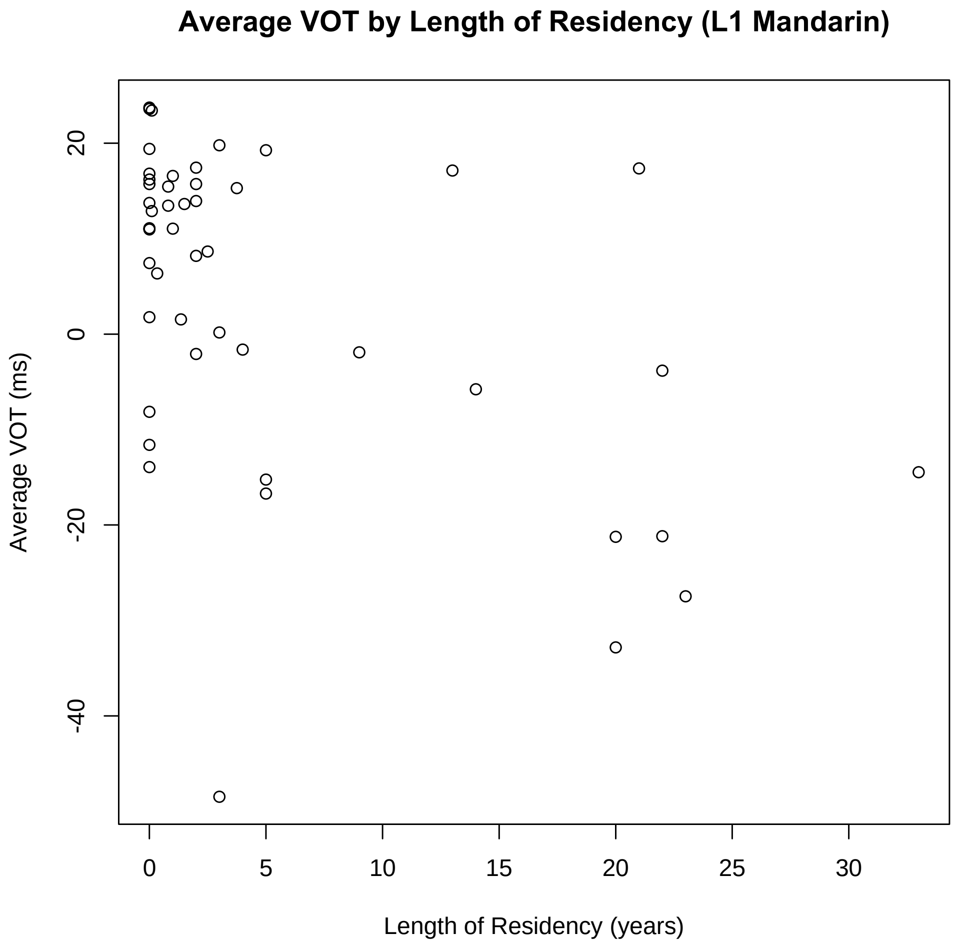


FIGURE 14. Length of residency and production of VOT by L1 Mandarin speakers.

3.2. L1 Spanish, L2 English

3.2.1 Participants and Materials

114 native Spanish speakers have submitted samples to the Speech Accent Archive, but the data set used here was limited to the most recently submitted 50 samples to make data collection more manageable and to keep the Mandarin, Spanish, and French data sets relatively even. In addition, samples from the same twenty L1 English speakers used in Section 3.1 were analyzed as a benchmark for comparison with the L1 Spanish samples.

3.2.2 Procedure

The procedure used here is identical to that described in Section 3.1.2, except that the phonemes of interest were the initial stops in the words *call*, *peas*, *toy*, and *kids*. VOT for the 50 L1 Spanish samples and the twenty L1 English control samples were measured using Praat, collected in a spreadsheet along with demographic information for each subject, and then imported into R for statistical analysis. Measurement of VOT using Praat proceeded in the same manner described in Section 3.1.2.

3.2.3 Results and Discussion

Presentation of data and analysis for the L1 Spanish speakers will proceed in the same manner as for the Mandarin speakers. Figure 15 and Table 3 give the basic results. VOT values for the initial voiceless English stop in each word differ in exactly the way we would expect if Spanish speakers associate aspirated (high-VOT) English stops with their native unaspirated (low-VOT) voiceless stops: the values, in each case, are larger for L1 English speakers than L1 Spanish speakers. Unlike the L1 Mandarin data, this difference is significant for each individual word as well as for the aggregate data. Data for all four words will again be collapsed in the analysis that follows.

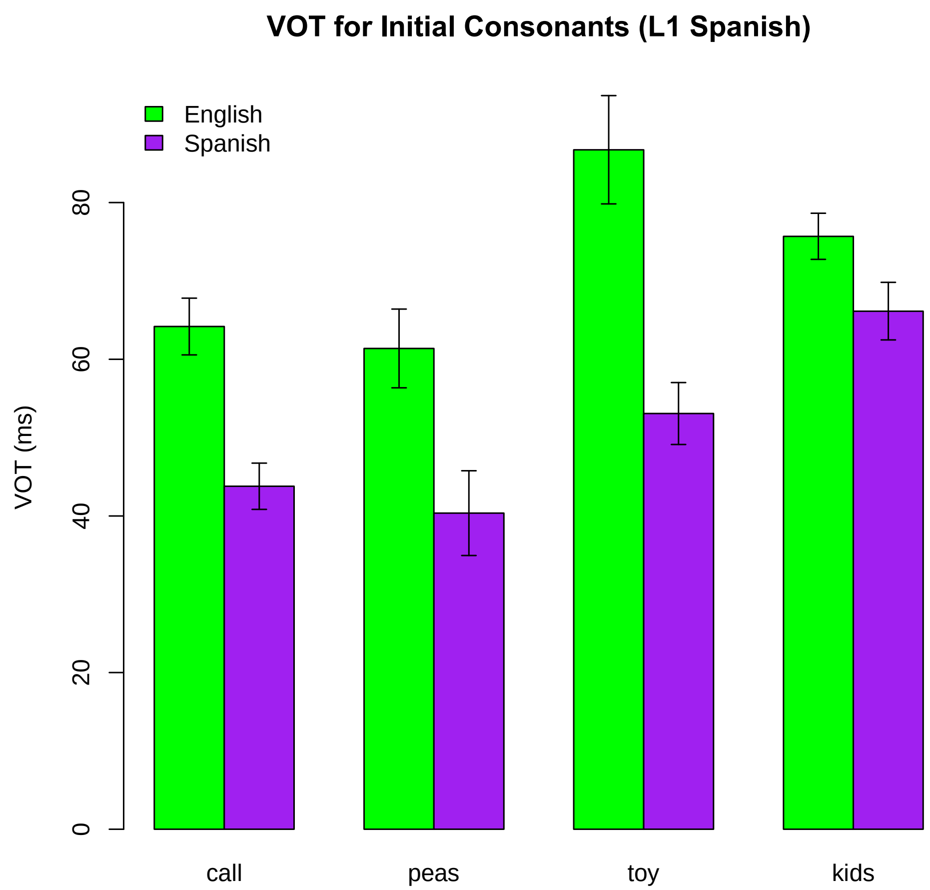


FIGURE 15. Average VOT for initial voiceless consonants produced by L1 Spanish and L1 English speakers.

	Spanish		English		Welch's t-test		
Word	**Mean VOT (ms)**	**Std. Dev.**	**Mean VOT (ms)**	**Std. Dev.**	**t**	**d.f.**	**p**
call	43.79	20.87	64.18	16.20	4.36	44.91	< 0.001
peas	40.36	38.24	61.37	22.46	2.85	58.24	< 0.01
toy	53.07	27.97	86.74	30.93	4.22	32.13	< 0.001
kids	66.14	25.97	75.69	13.14	2.03	64.09	< 0.05
ALL WORDS	**50.84**	**30.43**	**72.00**	**23.64**	**6.21**	**186.05**	**< 0.001**

TABLE 3. Descriptive and comparative statistics for initial voiceless stop data.

The effect of age of onset on production of voiced stops is shown in Figure 16. As might be expected, earlier learners produce stops with more English-like VOT values (recall from Table 4 that the native English mean VOT is 72 ms). A likelihood ratio comparison between a "null" model (with only random effects of speaker and word) and a model with an added fixed effect of age of onset confirms that this is a significant predictive factor ($\chi^2 = 4.35$, $p < 0.05$).

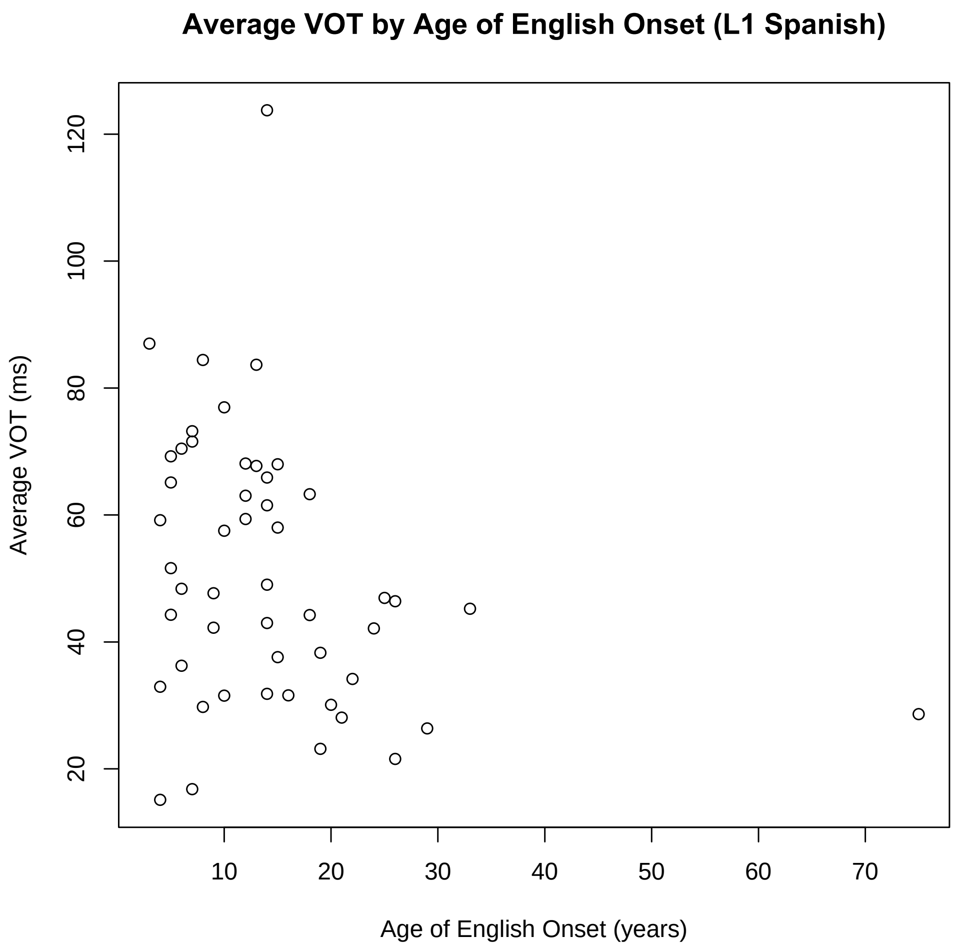


FIGURE 16. Age of onset and production of VOT by L1 Spanish speakers.

Figure 17 shows the relationship between mean VOT and length of residency. Impressionistically, it is difficult to discern a pattern from the graph; while there are fewer participants with longer LOR, their average VOT values do not seem to trend in any particular direction. The likelihood ratio comparison here confirms this impression: adding length of residency to the linear mixed effects model has no predictive value for the L1 Spanish speakers ($\chi^2 = 0.09$, $p = 0.76$ (n.s.)).

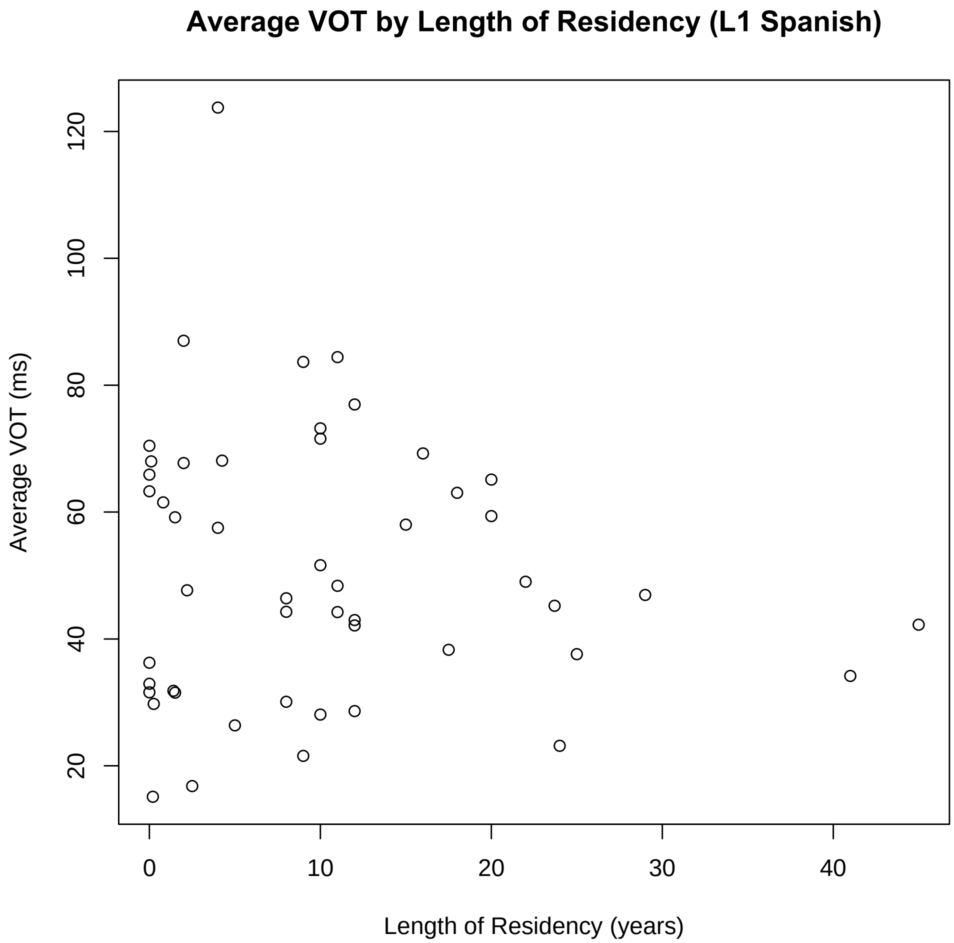


FIGURE 17. Length of residency and production of VOT by L1 Spanish speakers.

3.3. L1 French, L2 English

3.3.1 Participants and Materials

Data from all 51 native French speakers on the Speech Accent Archive were collected and processed using the methodology described in Sections 3.1.1 and 3.2.1. The same L1 English data used in previous sections was also used for comparison.

3.3.2 Procedure

The procedure used here is identical to that described in Section 3.2.2, targeting the initial stops in the words *call*, *peas*, *toy*, and *kids*.

3.3.3 Results and Discussion

Figure 18 and Table 4 give the basic results for L1 French speakers. VOT values for the initial voiceless English stop in each word differ in exactly the way we would expect if French speakers associate aspirated (high-VOT) English stops with their native unaspirated (low-VOT) voiceless stops: the values, in each case, are larger for L1 English speakers than L1 French speakers. This difference is highly significant for all individual words except *kids* and for the data set as a whole. Data for all four words will be collapsed in the analysis that follows.

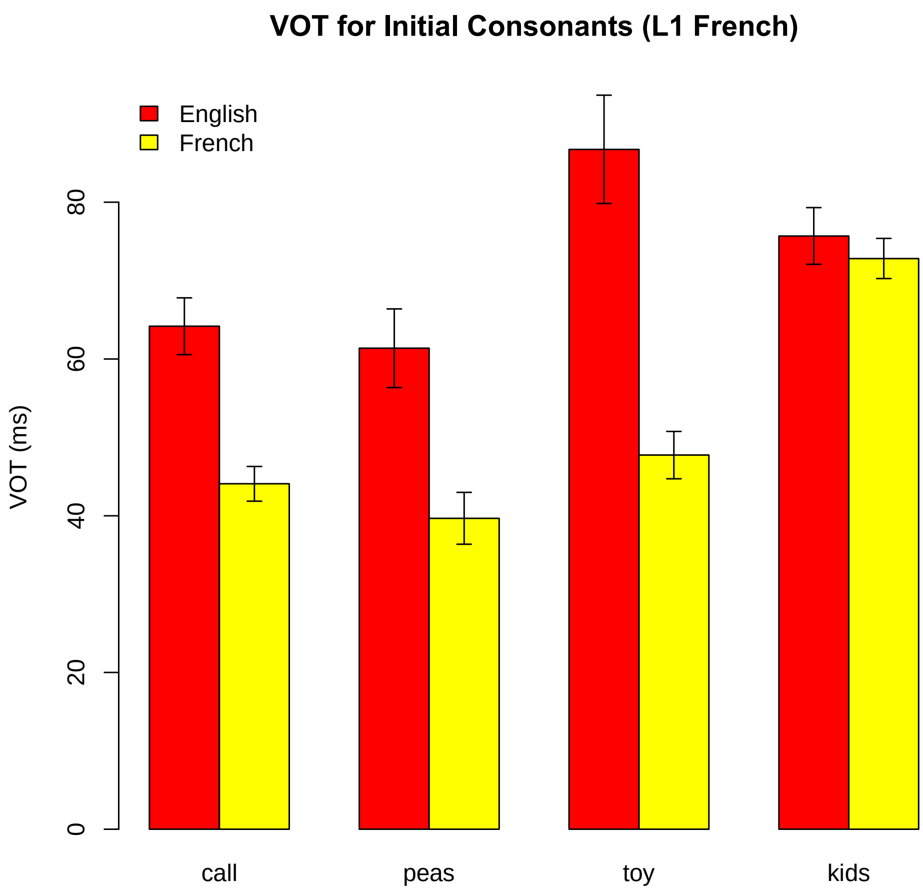


FIGURE 18. Average VOT for initial voiceless consonants produced by L1 French and L1 English speakers.

Word	French		English		Welch's t-test		
	Mean VOT (ms)	Std. Dev.	Mean VOT (ms)	Std. Dev.	t	d.f.	p
call	44.07	15.84	64.18	16.20	4.73	34.09	< 0.001
peas	39.67	23.61	61.37	22.46	3.61	36.42	< 0.001
toy	47.74	21.58	86.74	30.93	5.17	26.58	< 0.001
kids	72.81	18.29	75.69	13.14	0.74	48.25	0.46 (n.s.)
ALL WORDS	**51.07**	**23.72**	**72.00**	**23.64**	**-4.17**	**206.22**	**< 0.001**

TABLE 4. Descriptive and comparative statistics for initial voiceless stop data.

The effect of age of onset on production of voiced stops is shown in Figure 19. Perhaps surprisingly, no strong tendency is apparent: participants who began to acquire English earlier do not seem to produce more native-like (i.e. high-VOT) initial voiceless stops. A likelihood ratio comparison between a "null" model (with only random effects of speaker and word) and a model with an added fixed effect of age of onset confirms the impression that age of onset is not a strong predictor of VOT for L1 French speakers; the effect is only marginally significant ($\chi^2 = 3.52$, p = 0.06).

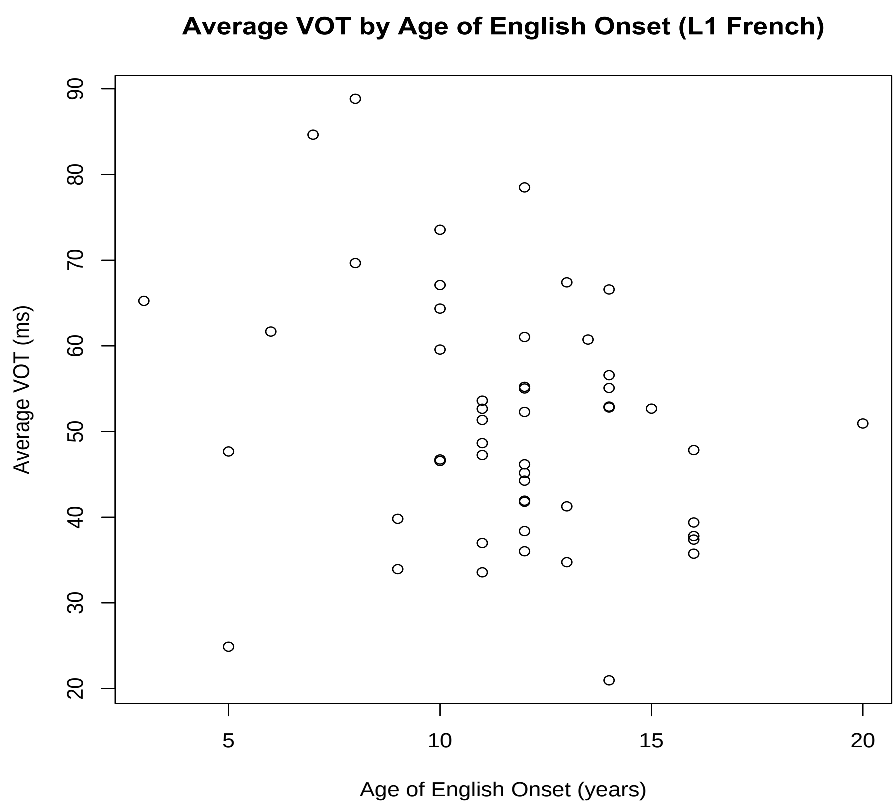


FIGURE 19. Age of onset and production of VOT by L1 French speakers.

Figure 20 shows the relationship between mean VOT and length of residency. A slight upward trend seems discernible here, indicating that speakers with longer LOR produce more English-like stops. Since the effect of age of onset proved to be marginal, I employ two additional mixed effects models to confirm the contribution of length of residency. The first model includes both age of onset and LOR as fixed effects; the likelihood ratio test confirms that this model significantly improves on the previous model that included only age of onset ($\chi^2 = 5.61$, $p < 0.05$). Furthermore, a model with only LOR as a fixed effect significantly improves on a model with only random effects ($\chi^2 = 6.41$, $p < 0.05$).

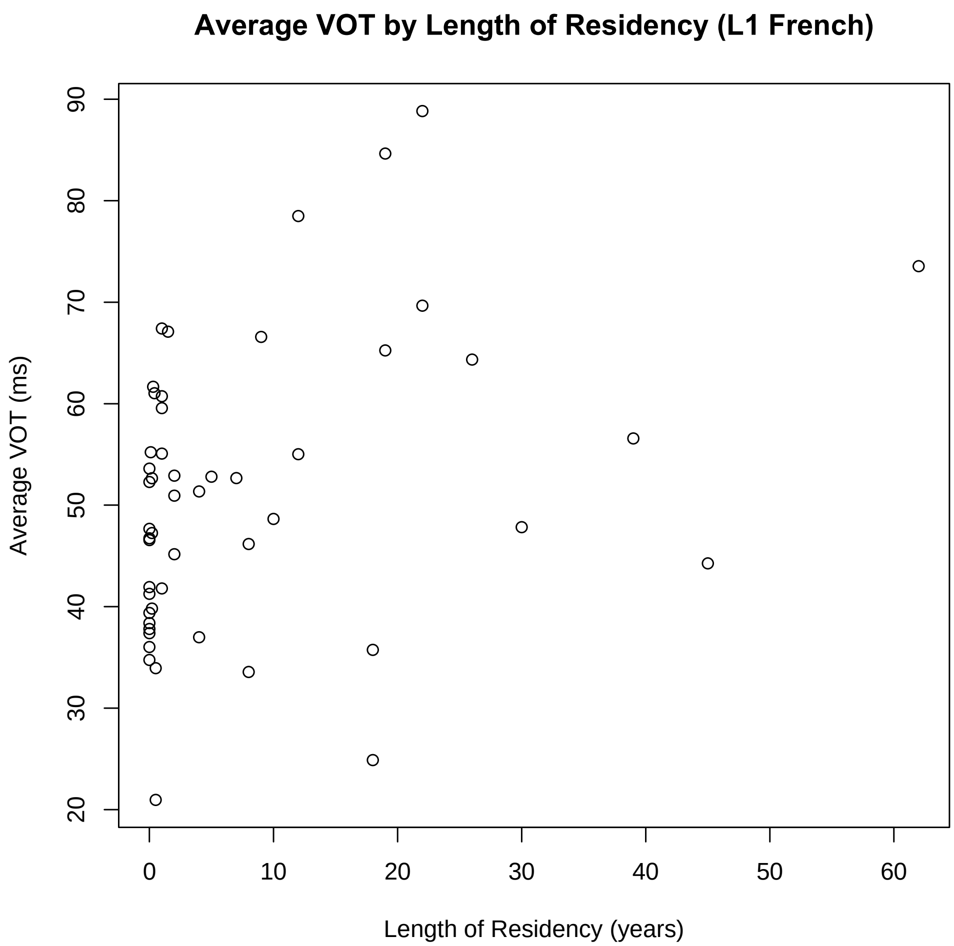


FIGURE 20. Length of residency and production of VOT by L1 French speakers.

4. General Discussion

It is not surprising that age of English onset emerged as a strong predictive factor for both Mandarin and Spanish speakers. Of all the linguistic domains, phonology is the most obviously susceptible to critical period effects (Lenneberg, 1967), so it is very much expected that early learners will generally manage to achieve native English-like VOT values while later learners will largely fail to do so. That is exactly what we found here for these two languages (see Figures 13 and 16). There is, however, one surprise in the results regarding age of onset: the null (or, at best, marginal) result obtained for the L1 French speakers. Linear mixed effects models led us to conclude that age of onset was not a significant predictor of VOT for these speakers. It may be the case that this factor has too narrow a range to reveal any significant correlation with VOT in this speaker group, since all participants had begun to learn English by the age of 20. A larger data set might reveal the expected result that younger learners should produce more English-like stops.

Turning to the main research question of this paper – whether L2 English learners can produce stops with more English-like VOT values with increased length of residency – the results are mixed, but point in a distinctly positive direction. LOR emerged as a very strong independent predictor of performance for the L1 Mandarin and French speakers, but had no effect on the performance of the L1 Spanish group. Whatever factor that accounts for this discrepancy is not likely to be found within phonological theory itself, since the relevant properties of Spanish and French stops are very similar. Both languages feature a contrast between fully voiced stops and voiceless, unaspirated stops, yet the pattern of acquisition of English stops differed markedly for the two groups. Since the effect of length of residency was observed for two very different languages, Mandarin and French, the preponderance of the evidence certainly indicates that this effect is real – that adult learners indeed can improve their pronunciation with experience, producing more native like values for VOT when producing English stops. The challenge is then to explain what is blocking this effect for the Spanish speakers. Sociolinguistic factors might plausibly come into play here. Perhaps Spanish speakers living in an English-speaking country tend to use their native language more often than similarly situated Mandarin or French speakers, causing them to retain L1 Spanish-like stops even with significant exposure to English. Or perhaps Spanish speakers, relative to Mandarin or French speakers, are more interested in maintaining their own linguistic identity than they are in assimilating to the dominant language community and therefore value the retention of certain aspects of their accent when speaking English. More detailed study is needed to provide a definitive answer to this problem.

To bring this discussion to a close, I would like to consider once again each of the models of phonological learning discussed in Section 2 of this paper, evaluating them in light of the results of the corpus study. The Speech Learning Model (Flege, 1995) predicts that late learners of a second language should not be able to form new categories for L2 sounds very similar to those found in their L1 due to deficiencies in linguistic perceptual ability. Only the results for the L1 Spanish speakers are consistent with this theory: early learners produced English-like stops, while later learners produced more Spanish-like stops, and length of residency had no effect on this pattern. The Speech Learning Model would have much more difficulty accounting for the Mandarin and French results, however. This model does not have any mechanism that could account for late learners improving over time, producing more native-like VOT values with longer LOR, as was robustly found in the present study. On the contrary, the SLM predicts fossilization at L1-like values when L1 and L2 sounds are as close together as they are in this study. Overall, the results discussed in this paper present a difficult challenge to proponents of the Speech Learning Model.

Precisely the opposite is true for the Full Transfer/Full Access Model (Escudero & Boersma, 2004; Schwartz & Sprouse, 1996), which predicts that, while beginning learners might produce sounds closer to their L1 than to their L2 target, they should be able to converge on L2 norms with the right kind and the right quantity of experience. The Mandarin and French data conform closely to this prediction; the right kind of experience, in this case, is residency in an L2-speaking country, and the right quantity appears to be twenty years or more of this kind of immersion. Since the Spanish data do not conform to this pattern, we cannot declare unequivocal victory for Full Transfer/Full Access; as discussed above, however, the evidence certainly seems to tip the scale firmly in that direction.

The Statistical Learning Model of McMurray et al. (2009) makes predictions similar to, but slightly different from those of Full Transfer/Full Access. On the analysis considered in Section 2, the Statistical Learning Model predicts that learners should initially produce L1-like stops, gradually expanding their existing category so that average VOT slowly approaches L2 norms, until finally forming an authentic L2 category with target-like production. The main observable difference should be a gap in the progress of VOT values: at some point, when a new category is formed by the learner, average values should jump more or less directly from some intermediate value to a target-like distribution. For reasons already amply discussed, the Spanish data will not provide support for any model that, like Statistical Learning, predicts progressive development, but it is worth taking a closer look at what the Mandarin and French data have to say. Figures 14 and 20 showed the mean VOT produced by each of the L1 Mandarin and French speakers, respectively, as a function of their length of residency. There is no clearly discernible gap in average VOT values of the kind that would be predicted by the Statistical Learning Model. It must be admitted that the data become somewhat sparse for larger values of LOR, so these data may well be insufficient to discern between the very similar predictions of Full Transfer/Full Access and the Statistical Learning Model.

5. Conclusion

In this paper, I presented the results of a corpus study using data from the George Mason University Speech Accent Archive. I looked at voice onset time (VOT) in the production of word-initial stops among three groups, native Mandarin, Spanish, and French speakers, all of whom were speakers of English as a second language. These three L2 English groups differed in expected ways from a control group of native English speakers, often producing stops with VOT closer to their L1 values or intermediate between L1 and target-like English values. Age of English onset emerged as a strong predictor of behavior among Mandarin and Spanish groups, in accord with the existence of a critical period effect for phonological acquisition. No such effect was observed for the French group, but this may have been due to the narrower range for age of onset in this group. More interestingly, length of residency in an English-speaking region emerged as a strong predictor of VOT production for the L1 Mandarin and French groups, providing strong support for the Full Transfer/Full Access model of acquisition, and a strong challenge for Flege's Speech Learning Model. This was not the case for the L1 Spanish speakers, who showed no indication of producing more authentic English stops as they gained linguistic experience. Since Spanish and French are very similar in the relevant phonological respects, it was speculated that sociolinguistic factors were at play in this null result for Spanish speakers. Further research is needed to confirm this speculation, and to more carefully evaluate these two models of phonological learning, as well as the Statistical Learning Model.

Appendix A: Corpus Study Data

VOT and demographic data for native English-speaking controls, L1 Mandarin speakers (Section 3.1), L1 Spanish speakers (Section 3.2), and L1 French speakers (Section 3.3) are presented in the tables below. Demographic data not considered in this paper are not included here, but are available from the George Mason University Speech Accent Archive (http://accent.gmu.edu) or from the author.

Speaker	Age	Sex	Word	VOT /b/ags (ms)	VOT /b/ig (ms)	VOT /b/ob (ms)	VOT /g/o (ms)	VOT /k/all (ms)	VOT /p/eas (ms)	VOT /t/oy (ms)	VOT /k/ids (ms)
english021	37	F	bags	16.236	16.719	-18.564	28.536	52.028	54.517	186.745	81.157
english023	43	M	bags	-119.096	-111.67	-100.662	-88.422	53.788	61.671	64.517	70.928
english044	63	F	bags	0	10.59	13.674	23.711	86.435	71.639	121.92	80.495
english068	52	M	bags	9.852	-57.656	-47.859	32.015	33.531	9.145	69.175	70.429
english075	32	M	bags	9.295	0	17.612	23.223	61.961	42.608	131.353	74.303
english133	36	F	bags	12.085	-41.47	0	10.98	52.855	69.256	85.302	69.581
english134	22	F	bags	0	9.913	-98.243	22.365	76.317	102.948	99.87	98.939
english135	22	M	bags	0	10.265	0	7.633	75.589	54.911	86.919	86.315
english137	26	M	bags	-87.744	-58.976	-50.565	14.189	45.894	83.066	57.234	69.804
english158	18	F	bags	10.792	7.853	-13.83	27.94	36.996	56.18	58.906	75.483
english159	18	M	bags	0	-12.113	15.995	24.581	81.758	73.422	100.253	81.335
english221	45	M	bags	21.834	9.282	12.804	31.794	64.042	26.068	78.713	78.256
english242	77	F	bags	0	10.399	17.984	46.095	90.917	42.027	84.82	91.703
english256	23	M	bags	-4.088	-19.271	0		85.38	54.32	78.845	93.326
english292	38	M	bags	-143.407	-104.351	-80.508	21.72	52.277	77.736	66.603	77.628
english325	32	M	bags	0	18.464	0	21.741	71.241	35.793	69.536	53.745
english332	20	F	bags	0	-9.808	-6.019		65.17	89.018	81.525	41.039
english369	27	M	bags	7.849	11.713	13.824	16.03	56.213	72.606	86.427	82.514
english412	31	F	bags	13.052	-62.617	-79.79	40.999	69.663	72.644	72.012	66.799
english441	19	F	bags	10.732	8.088	10.17	35.6	71.563	77.802	54.04	70.03

TABLE A1. Data for L1 English controls.

Speaker	Age	Sex	Onset	Method	LOR	VOT /b/ags (ms)	VOT /b/ig (ms)	VOT /b/ob (ms)	VOT /g/o (ms)
mandarin01	26	F	13	Academic	2	9.742	6.558	5.706	10.732
mandarin02	38	F	14	Academic	0.8	12.209	15.759	13.439	12.38
mandarin03	43	M	10	Academic	14	-24.069	-23.01	0	23.957
mandarin04	24	F	6	Academic	1	4.046	12.673	13.472	36.082
mandarin05	31	F	12	Academic	2	11.335	12.258	15.406	30.715
mandarin06	28	F	12	Academic	5	16.799	11.775	11.627	36.801
mandarin07	22	M	5	Naturalistic	22	-22.293	-19.251	-31.368	-11.79
mandarin08	29	M	12	Academic	5	-14.184	-34.379	-40.182	27.772
mandarin09	38	M	12	Academic	2	0	16.07	15.835	30.947
mandarin10	19	M	3	Academic	3.75	13.163	13.551	12.147	22.275
mandarin11	53	F	13	Academic	33	-32.249	-33.874	-10.768	18.996
mandarin12	23	M	1	Naturalistic	23	-19.799	-62.959	-41.224	14.114
mandarin13	29	M	13	Academic	0	11.17	7.146	7.926	18.123
mandarin14	49	M	20	Academic	21	20.333	13.963	11.172	23.905
mandarin15	28	F	11	Academic	4	-11.267	0	-19.984	24.737
mandarin16	32	M	10	Academic	1.35	10.521	-15.043	-12.838	23.499
mandarin17	26	M	13	Academic	2	-17.009	19.553	26.611	26.559
mandarin18	29	F	13	Academic	0	18.961	20.025	16.871	38.49
mandarin19	27	M	13	Academic	0.33	2.724	9.289	7.443	5.956
mandarin20	40	F	12	Academic	5	13.636	-83.527	-17.58	20.678
mandarin21	38	F	9	Academic	13	16.155	12.305	16.126	23.948
mandarin22	39	F	11	Academic	3	-57.214	-80.416	-85.928	29.729
mandarin23	46	F	13	Academic	22	-11.176	7.799	-33.972	21.994
mandarin24	21	F	10	Academic	0	12.044	8.142	13.664	21.07
mandarin25	28	F	13	Academic	0.8	11.763	12.44	8.606	28.962
mandarin26	31	F	12	Academic	3	17.622	17.237	25.629	18.583
mandarin27	18	M	15	Academic	3	21.515	14.73	-60.357	24.752
mandarin28	45	M	15	Academic	20	11.358	-130.403	13.723	20.395
mandarin29	24	M	4	Academic	0	-15.182	-11.19	-37.475	17.416
mandarin30	27	M	13	Academic	1	11.289	0	10.527	22.336
mandarin31	26	F	13	Academic	2.5	-4.698	12.246	-12.76	39.815
mandarin32	23	F	12	Academic	2	-14.242	-9.334	-11.677	26.895
mandarin33	25	M	6	Naturalistic	20	26.221	-98.457	22.872	-81.896
mandarin34	31	F	13	Academic	0	0	11.302	0	18.411
mandarin35	27	F	12	Academic	0	11.375	10.83	13.402	41.979
mandarin36	32	F	10	Academic	0	11.237	12.361	15.27	25.844
mandarin37	32	F	12	Academic	0	11.579	11.952	22.963	48.386
mandarin38	33	F	12	Academic	0.1	-11.267	25.431	49.484	29.997
mandarin39	24	F	12	Academic	0	11.171	-47.925	20.387	23.429
mandarin40	28	M	12	Academic	9	9.377	-63.113	16.322	29.776
mandarin41	34	F	13	Academic	0	10.765	13.639	14.385	24.07
mandarin42	47	F	16	Academic	0	9.42	9.157	9.021	16.239
mandarin43	24	F	15	Academic	0	12.215	12.769	19.604	22.621
mandarin44	21	M	16	Academic	1.5	8.602	9.362	11.03	25.508
mandarin45	42	F	12	Academic	0.1	7.827	7.552	16.074	20.075
mandarin46	43	F	9	Academic	0	0	10.82	16.969	16.327
mandarin47	28	F	8	Academic	0	9.568	7.097	-93.633	21.24
mandarin48	37	M	12	Academic	0	-76.162	10.019	10.44	23.143

TABLE A2. Data for L1 Mandarin speakers.

Speaker	Age	Sex	Onset	Method	LOR	VOT /k/all (ms)	VOT /p/eas (ms)	VOT /t/oy (ms)	VOT /k/ids (ms)
spanish065	32	M	12	Naturalistic	20	54.415	32.246	70.724	80.081
spanish066	30	M	12	Naturalistic	18	38.769	39.942	80.206	93.189
spanish067	48	F	14	Academic	4	70.222	184.319	73.62	166.856
spanish068	27	M	14	Academic	0.8	59.217	95.655	27.786	63.428
spanish069	20	F	7	Naturalistic	10	83.315	59.695	82.835	66.984
spanish070	30	F	10	Academic	1.5	34.04	0	25.631	66.475
spanish071	24	M	6	Academic	0	51.194	67.333	81.368	81.927
spanish072	20	F	5	Naturalistic	20	86.193	49.351	81.13	43.798
spanish073	19	M	15	Academic	0.1	38.752	76.528	51.241	105.443
spanish074	44	M	14	Academic	0	31.925	84.632	57.125	89.927
spanish075	28	F	5	Academic	8	31.359	25.066	62.302	58.411
spanish076	27	M	10	Academic	4	47.032	47.492	56.92	78.644
spanish077	45	M	13	Academic	2	66.845	41.32	70.234	92.552
spanish078	55	F	5	Academic	16	17.943	52.399	112.98	93.654
spanish079	70	F	22	Academic	41	38.907	21.359	21.512	54.883
spanish080	77	F	75	Naturalistic	12	52.26	12.985	15.363	33.889
spanish081	63	F	7	Academic	2.5	24.985	0	14.362	27.872
spanish082	49	M	14	Academic	12	42.398	30.012	38.303	61.175
spanish083	48	M	14	Academic	1.4	43.448	16.44	20.686	46.737
spanish084	19	F	4	Academic	0.2	29.931	-73.212	56.258	47.556
spanish085	44	F	6	Academic	0	34.265	-9.091	79.471	40.35
spanish086	23	M	4	Academic	0	27.881	14.974	39.523	49.335
spanish087	55	M	9	Academic	2.2	54.18	51.567	23.119	61.82
spanish088	53	M	25	Naturalistic	29	62.481	33.031	35.747	56.514
spanish089	22	M	13	Academic	9	41.061	84.4	107.741	101.442
spanish090	23	F	3	Naturalistic	2	51.092	124.663	73.645	98.638
spanish091	20	F	7	Naturalistic	10	80.414	55.705	86.815	63.357
spanish092	54	M	9	Naturalistic	45	46.796	32.513	46.431	43.237
spanish093	29	F	18	Naturalistic	11	36.902	47.135	39.364	53.506
spanish094	19	M	6	Academic	11	20.642	42.738	56.516	73.618
spanish095	31	F	21	Naturalistic	10	44.719	14.933	12.608	40.12
spanish096	29	F	12	Academic	4.25	32.441	47.725	78.028	114.255
spanish097	52	F	19	Academic	17.5	35.555	28.809	44.15	44.658
spanish098	34	M	29	Naturalistic	5	27.943	21.913	18.715	36.935
spanish099	20	F	4	Academic	1.5	74.937	21.734	91.542	48.497
spanish100	30	F	26	Academic	8	38.965	0	70.809	75.858
spanish101	39	F	19	Naturalistic	24	0	0	13.16	79.468
spanish102	31	M	26	Academic	9	26.933	12.942	13.181	33.198
spanish103	46	M	15	Naturalistic	25	33.829	29.789	19.498	67.33
spanish104	28	F	20	Naturalistic	8	35.148	8.086	32.67	44.436
spanish105	41	F	10	Academic	12	30.802	88.857	83.079	105.084
spanish106	21	M	5	Academic	10	42.12	50.727	55.44	58.189
spanish107	36	M	14	Naturalistic	22	24.823	69.382	52.792	49.101
spanish108	36	M	24	Academic	12	42.173	29.514	33.842	62.944
spanish109	21	M	16	Academic	0	31.474	17.199	24.756	52.923
spanish110	20	M	18	Academic	0	104.685	36.381	45.529	66.54
spanish111	55	M	33	Naturalistic	23.7	47.31	46.347	23.687	63.54
spanish112	25	M	15	Academic	15	46.973	38.122	94.403	52.53
spanish113	19	M	8	Academic	11	75.983	80.291	99.736	81.685
spanish114	22	F	8	Naturalistic	0.25	-6.355	34.094	57.048	34.273

TABLE A3. Data for L1 Spanish speakers.

Speaker	Age	Sex	Onset	Method	LOR	VOT /k/all (ms)	VOT /p/eas (ms)	VOT /t/oy (ms)	VOT /k/ids (ms)
french1	20	F	12	Academic	0.4	50.908	72.180	38.869	82.184
french2	19	M	14	Academic	0.5	29.588	-13.459	27.951	39.751
french3	22	F	11	Academic	0.2	43.053	29.583	58.736	79.217
french4	31	F	14	Academic	9	46.257	24.514	64.880	130.654
french5	36	F	11	Academic	4	30.860	41.743	26.847	48.497
french6	26	F	13	Academic	1	38.239	108.743	40.671	81.987
french7	18	M	5	Academic	18	24.347	14.812	19.118	41.237
french8	66	M	16	Academic	0	48.863	19.030	41.335	41.951
french9	21	M	11	Academic	0	67.608	38.246	58.987	49.542
french10	31	M	10	Academic	1	45.911	36.406	71.447	84.468
french11	31	M	11	Academic	8	22.362	27.510	19.977	64.378
french12	19	F	9	Academic	0.2	41.944	27.348	23.946	65.981
french13	19	M	12	Academic	0	31.455	17.067	31.389	64.183
french14	23	F	14	Academic	1	41.120	65.087	44.307	69.817
french15	32	M	12	Academic	0	27.841	36.574	28.223	60.838
french16	19	F	7	Academic	19	54.405	91.314	89.150	103.689
french17	39	M	14	Academic	5	48.865	37.200	40.721	84.420
french18	22	M	6	Academic	0.3	37.618	69.648	60.185	79.219
french19	39	M	12	Academic	12	64.505	33.742	59.691	62.167
french20	23	M	12	Academic	2	39.911	49.004	29.985	61.777
french21	20	M	9	Academic	0.5	31.218	18.685	27.966	57.891
french22	78	F	16	Academic	18	35.386	20.377	27.528	59.678
french23	76	F	16	Academic	0	30.734	45.619	27.745	53.442
french24	47	M	13	Academic	0	29.838	11.306	15.955	81.879
french25	20	M	12	Academic	0.1	25.423	67.046	61.941	66.468
french26	27	F	11	Academic	4	38.084	19.996	59.145	88.149
french27	38	F	12	Academic	8	46.696	46.031	36.098	55.847
french28	35	F	11	Naturalistic	10	47.760	25.425	52.347	69.042
french29	54	F	10	Academic	26	52.305	34.410	61.211	109.435
french30	37	M	12	Academic	12	117.219	56.513	48.400	91.850
french31	28	M	15	Academic	7	55.118	21.453	55.190	78.932
french32	60	M	16	Academic	0	50.787	18.056	19.164	61.496
french33	62	M	10	Academic	62	38.937	53.477	116.716	85.035
french34	56	M	16	Academic	30	72.404	13.036	20.144	85.709
french35	27	M	10	Academic	0	37.526	37.232	49.534	62.664
french36	32	F	10	Academic	1.5	47.575	61.665	75.937	83.196
french37	42	M	14	Naturalistic	2	30.641	60.586	67.291	53.120
french38	22	M	8	Academic	22	75.585	95.181	96.581	87.982
french39	28	M	13	Academic	0	35.424	27.228	24.003	78.350
french40	44	M	13.5	Academic	1	39.791	66.918	65.730	70.515
french41	24	M	12	Academic	0	38.011	22.842	47.815	59.028
french42	22	F	20	Academic	2	32.112	36.364	53.086	82.166
french43	22	M	11	Academic	0.2	53.512	29.783	41.035	64.658
french44	20	F	12	Academic	0	44.978	32.881	62.011	69.254
french45	22	M	8	Naturalistic	22	54.433	63.446	77.695	83.076
french46	22	M	10	Academic	0	40.328	19.824	45.129	80.962
french47	66	F	12	Academic	45	47.424	35.897	31.928	61.775
french48	19	F	3	Naturalistic	19	56.267	30.833	64.222	109.668
french49	39	M	12	Academic	1	41.735	30.169	36.038	59.251
french50	67	F	14	Academic	39	31.292	69.818	30.725	94.460
french51	18	M	5	Academic	0	33.459	24.772	59.840	72.580

TABLE A4. Data for L1 French speakers.

Appendix B: Linear Mixed Effects Model Results

Detailed results for each linear mixed effects model discussed in Section 3 are given below.

```
Linear mixed model fit by REML
Formula: M$VOT ~ (1 | M$Speaker) + (1 | M$Word)
  AIC  BIC logLik deviance REMLdev
 1801 1814 -896.5     1799    1793
Random effects:
 Groups    Name        Variance Std.Dev.
 M$Speaker (Intercept) 129.49   11.379
 M$Word    (Intercept) 137.33   11.719
 Residual              555.76   23.575
Number of obs: 192, groups: M$Speaker, 48; M$Word, 4

Fixed effects:
            Estimate Std. Error t value
(Intercept)    4.001      6.317   0.633
```

TABLE B1. L1 Mandarin, model with only random effects of speaker and word.

```
Linear mixed model fit by REML
Formula: M$VOT ~ M$EngOnset + (1 | M$Speaker) + (1 | M$Word)
  AIC  BIC logLik deviance REMLdev
 1796 1812 -892.9     1792    1786
Random effects:
 Groups    Name        Variance Std.Dev.
 M$Speaker (Intercept) 101.17   10.058
 M$Word    (Intercept) 137.33   11.719
 Residual              555.76   23.575
Number of obs: 192, groups: M$Speaker, 48; M$Word, 4

Fixed effects:
            Estimate Std. Error t value
(Intercept) -15.0947     9.7487  -1.548
M$EngOnset    1.6726     0.6538   2.558

Correlation of Fixed Effects:
           (Intr)
M$EngOnset -0.766
```

TABLE B2. L1 Mandarin, model with added fixed effect of age of onset.

```
Linear mixed model fit by REML
Formula: M$VOT ~ M$EngOnset + M$LOR + (1 | M$Speaker) + (1 | M$Word)
  AIC  BIC logLik deviance REMLdev
 1787 1806 -887.3     1780    1775
Random effects:
 Groups    Name        Variance Std.Dev.
 M$Speaker (Intercept)  49.409   7.0291
 M$Word    (Intercept) 137.332  11.7189
 Residual              555.757  23.5745
Number of obs: 192, groups: M$Speaker, 48; M$Word, 4

Fixed effects:
            Estimate Std. Error t value
(Intercept)  -6.8505     9.3236  -0.735
M$EngOnset    1.3635     0.5851   2.330
M$LOR        -0.9118     0.2469  -3.693

Correlation of Fixed Effects:
           (Intr) M$EngO
M$EngOnset -0.736
M$LOR      -0.239  0.143
```

TABLE B3. L1 Mandarin, model with added fixed effects of age of onset and length of residency.

```
Linear mixed model fit by REML
Formula: S$VOT ~ (1 | S$Speaker) + (1 | S$Word)
  AIC  BIC logLik deviance REMLdev
 1889 1902 -940.6     1887    1881
Random effects:
 Groups    Name        Variance Std.Dev.
 S$Speaker (Intercept) 327.34   18.092
 S$Word    (Intercept) 122.62   11.073
 Residual              511.45   22.615
Number of obs: 200, groups: S$Speaker, 50; S$Word, 4

Fixed effects:
            Estimate Std. Error t value
(Intercept)   50.839      6.304   8.064
```

TABLE B4. L1 Spanish, model with only random effects of speaker and word.

```
Linear mixed model fit by REML
Formula: S$VOT ~ S$EngOnset + (1 | S$Speaker) + (1 | S$Word)
  AIC  BIC logLik deviance REMLdev
 1888 1904 -938.8     1882    1878
Random effects:
 Groups    Name        Variance Std.Dev.
 S$Speaker (Intercept) 297.48   17.248
 S$Word    (Intercept) 122.62   11.074
 Residual              511.45   22.615
Number of obs: 200, groups: S$Speaker, 50; S$Word, 4

Fixed effects:
            Estimate Std. Error t value
(Intercept)  58.7141     7.2882   8.056
S$EngOnset   -0.5469     0.2595  -2.107

Correlation of Fixed Effects:
           (Intr)
S$EngOnset -0.513
```

TABLE B5. L1 Spanish, model with added fixed effect of age of onset.

```
Linear mixed model fit by REML
Formula: S$VOT ~ S$EngOnset + S$LOR + (1 | S$Speaker) + (1 | S$Word)
  AIC  BIC logLik deviance REMLdev
 1890 1910 -939.1     1882     1878
Random effects:
 Groups    Name        Variance Std.Dev.
 S$Speaker (Intercept) 305.50   17.479
 S$Word    (Intercept) 122.62   11.074
 Residual              511.45   22.615
Number of obs: 200, groups: S$Speaker, 50; S$Word, 4

Fixed effects:
            Estimate Std. Error t value
(Intercept) 59.44911    7.63905   7.782
S$EngOnset  -0.52832    0.26777  -1.973
S$LOR       -0.09807    0.29285  -0.335

Correlation of Fixed Effects:
           (Intr) S$EngO
S$EngOnset -0.424
S$LOR      -0.287 -0.207
```

TABLE B6. L1 Spanish, model with added fixed effects of age of onset of length of residency.

```
Linear mixed model fit by REML
Formula: VOT ~ (1 | Subject) + (1 | Word)
   AIC   BIC logLik deviance REMLdev
 724.8 734.5 -358.4    722.2   716.8
Random effects:
 Groups   Name        Variance Std.Dev.
 Subject  (Intercept) 100.49   10.024
 Word     (Intercept) 107.69   10.377
 Residual             253.77   15.930
Number of obs: 83, groups: Subject, 21; Word, 4

Fixed effects:
            Estimate Std. Error t value
(Intercept)    9.270      5.896   1.572
```

TABLE B7. L1 French, model with only random effects of speaker and word.

```
Linear mixed model fit by REML
Formula: VOT ~ Onset + (1 | Subject) + (1 | Word)
   AIC   BIC logLik deviance REMLdev
 725.3 737.4 -357.6    721.8   715.3
Random effects:
 Groups   Name        Variance Std.Dev.
 Subject  (Intercept) 105.82   10.287
 Word     (Intercept) 107.24   10.356
 Residual             253.83   15.932
Number of obs: 83, groups: Subject, 21; Word, 4

Fixed effects:
            Estimate Std. Error t value
(Intercept)   6.6472     7.3160   0.909
Onset         0.4320     0.7108   0.608

Correlation of Fixed Effects:
      (Intr)
Onset -0.590
```

TABLE B8. L1 French, model with added fixed effect of age of onset.

```
Linear mixed model fit by REML
Formula: VOT ~ LOR + (1 | Subject) + (1 | Word)
  AIC  BIC logLik deviance REMLdev
 1787 1804 -888.7     1781    1777
Random effects:
 Groups   Name        Variance Std.Dev.
 Subject  (Intercept) 115.44   10.744
 Word     (Intercept) 215.74   14.688
 Residual             265.94   16.308
Number of obs: 204, groups: Subject, 51; Word, 4

Fixed effects:
            Estimate Std. Error t value
(Intercept)  48.0296     7.6718   6.261
LOR           0.3750     0.1449   2.589

Correlation of Fixed Effects:
    (Intr)
LOR -0.153
```

TABLE B9. L1 French, model with added fixed effect of length of residency.

```
Linear mixed model fit by REML
Formula: VOT ~ Onset + LOR + (1 | Subject) + (1 | Word)
  AIC  BIC logLik deviance REMLdev
 1786 1806   -887     1779    1774
Random effects:
 Groups   Name        Variance Std.Dev.
 Subject  (Intercept) 109.33   10.456
 Word     (Intercept) 215.74   14.688
 Residual             265.94   16.308
Number of obs: 204, groups: Subject, 51; Word, 4

Fixed effects:
            Estimate Std. Error t value
(Intercept)  59.7657    10.4697   5.708
Onset        -0.9882     0.6009  -1.645
LOR           0.3435     0.1437   2.390

Correlation of Fixed Effects:
      (Intr) Onset
Onset -0.682
LOR   -0.200  0.133
```

TABLE B10. L1 French, model with added fixed effects of age of onset of length of residency.

Bibliography

Bates, D., Maechler, M., & Bolker, B. (2012). Linear mixed-effects models using S4 classes (Version 0.999999-0). Retrieved from http://lme4.r-forge.r-project.org/

Bialystok, E., Craik, F., & Ryan, J. (2006). Executive control in a modified antisaccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(6).

Boersma, P., & Weenink, D. (2012). Praat: doing phonetics by computer (Version 5.3.21). Retrieved from http://www.praat.org/

Chao, K.-Y., & Chen, L.-M. (2008). A Cross-Linguistic Study of Voice Onset Time in Stop Consonant Productions. *Computational Linguistics and Chinese Language Processing*, *13*(2), 215–232.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*(4), 551–585.

Flege, J. E. (1979). *Phonetic interference in second language acquisition* (Unpublished Ph.D. Thesis). Indiana University.

Flege, J. E. (1980). Phonetic approximation in second language acquisition. *Language Learning*, *30*, 117–134.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustic Socity of America*, *89*(1), 395–411.

Flege, J. E. (1995). Second-language speech learning: Theory, Findings, and Problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–278). Baltimore, MD: York Press.

Flege, J. E., & Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, *6*, 185–202.

Gandour, J., Petty, S. H., Dardarananda, R., Dechongkit, S., & Mukngoen, S. (1986). The acquisition of the voicing contrast in Thai: a study of voice onset time in word-initial stop consonants*. *Journal of Child Language*, *13*(3), 561–572.

Goldrick, M. (2004). Phonological features and phonotactic constraints in speech production. *Journal of Memory and Language*, *51*, 586–603.

Gonzalez Lopez, V. (2012). Spanish and English word-initial voiceless stop production in code-switched vs. monolingual structures. *Second Language Research*, *28*(2), 243–263.

Halicki, S. (2010). *Learner knowledge of target phonotactics: Judgments of French word transformations* (Ph.D. dissertation). Indiana University.

Kirov, C., & Wilson, C. (2012). The Specificity of Online Variation in Speech Production. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 3 4th Annual Conference of the Cognitive Science Society* (pp. 587–592). Austin TX: Cognitive Science Society.

Lenneberg, E. (1967). *Biological Foundations of Language*. New York: Wiley.

Lisker, L., & Abramson, A. (1964). Voice Onset Time in the Production and Perception of English Stops. Presented at the LSA General Meeting, New York.

Major, R. C. (1987). English voiceless stop production by speakers of Brazilian Portuguese. *Journal of Phonetics*, *15*, 197–202.

Maye, J., & Gerken, L. (2000). Learning phonemes without minimal pairs. In S. C. Howell, S. A. Fish, & T. Keith-Lucas (Eds.), *Proceedings of the 24th Annual Boston University Conference on Language Development* (pp. 522–533). Somerville MA: Cascadilla Proceedings Project.

Maye, J., Werker, J., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), 101–111.

McMurray, B., Aslin, R., & Toscano, J. (2009). (2009). Statistical learning of phonetic categories: Insights from a computational approach. *Developmental Science*, *12*(3), 369–378.

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, *39*(2), 132–142.

Pavlenko, A. (2000). L2 influence on L1 in late bilingualism. *Issues in Applied Linguistics*, *11*(2), 175–205.

R Development Core Team. (2011). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Riney, T., & Takagi, N. (1999). Global Foreign Accent and Voice Onset Time Among Japanese EFL Speakers. *Language Learning*, *49*(2), 275–302.

Rodriguez-Fornells, A., Balaguer, R., & Münte, T. (2006). Executive control in bilingual language processin. *Language Learning*, *56*(1), 133–190.

Schwartz, B., & Sprouse, R. (1996). L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research*, *12*, 40–72.

Scovel, T. (1995). Differentiation, Recognition, and Identification in the Discrimination of Foreign Accents. In J. Archibald (Ed.), *Phonological Acquisition and Phonological Theory*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Tees, R., & Werker, J. (1984). Perceptual flexibility: Maintenance or recovery of the ability to discriminate non-native speech sounds. *Canadian Journal of Psychology*, *38*(4), 579–590.

Weinberger, S. (2012). Speech Accent Archive. Retrieved from http://accent.gmu.edu

Williams, L. (1977). The voicing contrast in Spanish. *Journal of Phonetics*, *5*, 169–184.

Matrix Relationship Theory

I. Introduction

This paper explores sentences with two clauses and attempts to better understand how embedded clauses are structurally related to higher clauses. I conclude that clauses within the same sentence are either in a matrix relationship with each other or they are independent of each other. Clauses that are in a matrix relationship have a predictable structure between the embedded clause's CP and the higher clause's verb. If the structure does not exist, a matrix relationship does not exist, and the clauses are assumed to be independent of each other.

Section II, the theoretical framework section of this paper, outlines how clausal structure is broadly interpreted throughout the text, identifies why clauses are either open or closed, and introduces a referencing system to identify specific clauses in sentences with multiple clauses.

Section III, the analysis section of this paper, identifies the predictable structure that exists between clauses in a matrix relationship. This paper uses the word 'exists' because I make no attempt to claim how a matrix relationship structure surfaces from a cognitive process; I only attempt to show that the structure is required.

Section IV, the consequences section of this paper, posits that matrix relationships constitute the necessary conditions for filler extraction in complex noun phrases. The filler cannot be grammatically extracted in sentence 1 because the clauses lack a matrix relationship. The filler can be grammatically extracted in sentence 2 because the clauses are in a matrix relationship. If a matrix relationship does not exist, a filler cannot be extracted.

1. *What did Simon spread the rumor that they started ___?

2. What did the captain give the command to start __?

(Hofmeister and Sag, 2010, p. 371)

Section V, the discussion section of this paper, compares the Matrix Relationship Theory (MRT) to Chomsky's Barriers Theory (*Barriers)* and Hawkins's Filler-Gap-Complexity Hypothesis (FGCH). I conclude that MRT is the best proposal for predicting filler extraction; it explains complexity and works out-of-the-box. We will return to these thoughts later in the paper.

II. Theoretical framework

I will broadly interpret clausal structure as [CP [IP YP]. This broad interpretation supports more narrow interpretations of clausal structure:

- Adger: [C>T>(Neg)>(Perf)>(Prog)>(Pass)>v>V] (Adger, 2003)
- Chomsky: [XP[(EA) H YP] (Chomsky, 2000, p. 103)

In "Minimalist Inquiries," Chomsky (2000) posits that a clausal structure with finite T "is a closed system with regard to Case/agreement properties, determined internally without outside effect" (p. 103). Therefore, if the head of IP is finite T, I will annotate the clausal structure as [CP-closed [IP YP]]. If the head of IP is defective T, I will annotate the clausal structure as [CP-open [IP YP]].

When referencing clauses, I will always refer to the most embedded clause as n=1. The next highest clause will always be n+1. Therefore, the clauses in sentence 3 can be referenced as:

3. [CP **[**IP Bill claims [CP that [IP Sarah believes [CP that [IP Susan thinks [CP that [IP John bought a ring]]]]]]]].

References:

[CP that [IP-closed John bought a ring]]]]]]]	= Clause 1
[CP that [IP-closed Susan thinks	= Clause 2
[CP that [IP-closed Sarah believes	= Clause 3
[CP **[**IP-closed Bill claims	= Clause 4

Using the theoretical framework above, we can revise sentences 1, 2, and 3 as sentences 4, 5, and 6, respectively. Sentence 6 will not be revised with four clauses, as shown in sentence 3. I will revise sentence 6 with two clauses because this paper only examines structural relationships in declarative sentences with two clauses.

4. [CP-closed [IP Simon spread [$_{NP}$ the rumor [$_{CP\text{-closed}}$ that [$_{IP}$ they started a fight]]].

5. [CP-closed [IP The captain gave [$_{NP}$ the command [$_{CP\text{-open}}$ [$_{IP}$ to start a fight]]] (Hofmeister & Sag, 2010, p. 371, 20a and 20c").

6. [CP-closed [IP Susan thinks [CP-closed that [IP John bought a ring]]]] (Sprouse, 2012, pp. 82-83).

III. Analysis

I conclude that a matrix relationship does not exist between clause 1 and clause 2 of sentence 4, but does exist between clause 1 and clause 2 of sentences 5 and 6. The structural requirements for a matrix relationship are outlined below:

Matrix Relationship Requirements

Structural requirement between two closed clauses

For a matrix relationship to exist between a closed embedded clause and a closed higher clause, the CP of the embedded clause must be the sister of the higher clause's verb. There can be no intervening phrase, such as an NP, between the verb of the matrix clause and the CP of the embedded clause. This rule explains why sentence 4 does not have a matrix relationship, but sentence 6 does: sentence 6 maintains a sisterhood relationship.

Structural requirement between an embedded open clause and a higher closed clause

For a matrix relationship to exist between an embedded open clause and a higher closed clause, the embedded clause must be c-commanded by the higher clause's verb. The intervening NP in sentence 5 does not interfere with this loose c-command requirement and a matrix relationship is established. The complex NP only interferes with a strict sisterhood relationship between two closed clauses.

Here, we have a better understanding for why NPs are complex in filler-gap-dependences (FGDs); NPs have the ability to interfere with a structural relationship. It is not the case that NPs are inherently complex. Sentence 5b-d show that grammatical judgments do not change as the NP becomes more or less definite. Here, NPs are only complex when intervening in the structure of matrix relationships.

5b. [CP-closed [IP The captain gave [$_{NP}$ the command [$_{CP\text{-}open}$ [$_{IP}$ to start a fight]]].

5c. [CP-closed [IP The captain gave [$_{NP}$ a command [$_{CP\text{-}open}$ [$_{IP}$ to start a fight]]].

5d. [CP-closed [IP The captain gave [$_{NP}$ some command [$_{CP\text{-}open}$ [$_{IP}$ to start a fight]]].

This paper proposes that matrix relationships have a predictable structure. Further research will look to see if structural requirements change in sentences with interrogative CPs. For now, we can assume that the structural requirement is determined by the finiteness of T. Below are some general principles that I expect to remain constant in future work:

- In the context of a sentence, embedded non-finite clauses are always in a matrix relationship.
- Embedded finite clauses have a stricter structural requirement than non-finite clauses.

IV. Consequences

One can test to see if a matrix relationship exists between two clauses by employing a **filler-gap-dependency (FGD) test**. Sentences 7, 8, and 9 are the FGD counterparts to sentence 4, 5, and 6, respectively. The filler cannot be extracted grammatically in sentence 7, because clause 1 and clause 2 lack a matrix relationship. The filler can be extracted in sentence 8 and 9, because the clauses maintain a matrix relationship.

7. *[$_{\text{CP-closed}}$ What did [$_{\text{CP-closed}}$ [$_{\text{IP}}$ Simon spread [$_{\text{NP}}$ the rumor [$_{\text{CP-closed}}$ that [$_{\text{IP}}$ they started __]]]]]]]?

8. [$_{\text{CP-closed}}$ What did $_{\text{CP-closed}}$ [$_{\text{IP}}$ the captain give [$_{\text{NP}}$ the command [$_{\text{CP-open}}$ [$_{\text{IP}}$ to start __]]]]]]]?

9. [$_{\text{CP-closed}}$ What did [$_{\text{CP-closed}}$ [$_{\text{IP}}$ Susan think [$_{\text{CP-closed}}$ that [$_{\text{IP}}$ John bought ___]]]]]].

V. Discussion

There are two alternate proposals in the primary literature that identify the conditions necessary for filler extraction: Chomsky's *Barriers Theory* and Hawkins's *Filler-Gap-Complexity Hypothesis (FGCH)*.

Chomsky's *Barriers* claims that "XPs act as barriers to movement or extraction, specifically XPs that are not theta-governed" (Hofmeister & Sag, 2010, p. 371):

> Dependencies [fillers] that cross zero barriers are thus 0-subjacent and should sound perfectly acceptable, '1-subjacency' should translate to marginal acceptability, but anything higher [than 1-subjacency] 'should yield a considerable decrement in acceptability.' (Hofmeister & Sag, 2010, p. 371)

Barriers claims that extraction cannot occur grammatically in sentence 7 because the filler has to cross two barriers: the inherent CP barrier and a transformed IP barrier. Here, Chomsky claims that the tensed island changes the IP blocking category to a barrier.

Processing advocates criticize this approach because the grammarian does not have to prove why the tensed island in sentence 7 changes a blocking category to a barrier. Homeister & Sag claim:

> [*Barriers* exhibits] a pattern of asserting without argument that counterexamples [such as a tensed island] are 'special' or 'exceptional" and therefore "Perhaps nothing beyond considerations of elegance and good taste stand in the way of a grammarian. (Hofmeister & Sag, 2010, p. 370, p. 402)

The problem with *Barriers* is that the theory does not work 'out-of-the-box.' The theory creates a framework for subjacency with an inherent CP barrier, but it does not provide further insight on the conditions that change blocking categories to barriers. Here, the theory cannot predict anything more than 1-subjacency. Out of the box, it predicts every sentence to be marginally grammatical and only after the grammarian can change the blocking category to a barrier. The theory lacks transparency.

Hawkins posits the Filler Gap Complexity Hypothesis to make the simple claim that languages prefer less complex structures.

> **Filler-Gap Complexity Hypothesis:** If an FGD of complexity n on a complexity hierarchy H is grammatical, then FGDs for all less complex variables on H (n-1) will also be grammatical (Hawkins, 1999, p. 252).

Here, acceptability decreases as complexity increases. In this hierarchical approach, Hawkins claims that filler extraction is dependent on the environment of the FGD: "infinitival phrases are most hospitable to gaps, that finite subordinate clauses are more resistant, and complex NP environments are most resistant of all." (Hawkins, 252).

Hawkins also proposes that acceptability decreases as the number of nodes within an FGD increase. An increase in nodes equals an increase in complexity.

> **Minimize FGDs:** The human processor prefers FGDs to be as small as possible (251).

Based on the proposals above, Hawkins (1999) posits that sentence 11 is less acceptable than sentence 10, because sentence 11 has more complex structures.

10. Who did you hope that you would see ___? (Hawkins, 1999, 34a,)

11. *Who did you know the professor that taught ___? (Hawkins, 1999, 35a)

(Hawkins, 1999, p. 264)

Following the logic above, I assume that FGCH would prefer sentence 8 to sentence 7, because sentence 8 has more preferred structures. It includes a hospitable infinitival and a fewer count of syntactic nodes.

The Matrix Relationship Theory (MRT) proposes that extraction cannot occur between two clauses that lack a matrix relationship. If a matrix relationship does not exist, extraction cannot occur. Extraction cannot occur in sentence 7 because the clauses lack a matrix relationship.

MRT differs from FGCH because it explains why the complex NP environment is a resistant structure. MRT proposes that a complex NP is 'complex' when it intervenes with a strict structural requirement: sisterhood. The complex NP is not 'complex' in sentence 8 because it does not intervene with the loose structural requirement: c-command.

While *Barriers* posits that the tensed island in sentence 7 changes the IP blocking category to a barrier, MRT simply accounts for the tensed island as a free feature of the closed clause.

MRT is similar to *Barriers*, but it tries to provide more transparency into the factors that create barriers. At the very least, I hope this paper introduces the idea that structural relationships, and the lack thereof, play a critical role in predicting the conditions necessary for filler extraction.

References

Chomsky, Noam. 2000. Minimalist inquiries. In R. Martin, D. Michaels & J. Uriagereka (eds.), Step by step, 89–155. Cambridge, Mass.: MIT Press.

Hawkins. (1999). Processing complexity and filler gap dependencies across grammar. *Language*, *75*(2), 244-285.

Hofmeister. , & Sag, (2010). Cognitive constraints and island effects. *Language*, *86*(2), 366-415.

Sprouse. (2012). A test of the relation between working memory capacity and syntactic island effects. *Language*, *88*(1), 82-123.