An Optimality Theory Account of Phonological Variation in Kimberley Kriol

A Senior Thesis by Hannah LaPalombara
Advised by Assoc. Prof. Claire Bowern and Asst. Prof. Gaja Jarosz

Submitted to the faculty of the Department of Linguistics in partial fulfillment of the requirements for the degree of Bachelor of Arts

Yale University
May 2013

Special thanks to Ingrid Ningarmara, Glennis Galbat-Newry, Dwayne Newry, Jimmy Paddy, KJ Olawsky, Frances Kofod, the Mirima Council, and Mirima Dawang Woorlab-gerring Language and Culture Centre

Abstract: The kind of variation typically exhibited by speakers of creole languages—which is somewhat unique in that it is structured with respect to other natural languages, namely the superstrate and substrate languages—has long represented a challenge to traditional grammatical theory (Rickford 1980). In this paper, I argue that by employing the same methods used to describe phonological variation in other natural languages, we can account for phonological variation in creoles. This study presents a Stochastic Optimality Theory analysis of free variation among stops and fricatives in Kimberley Kriol, a creole language spoken in the Kimberley region of Australia. In the paper, I compare three different models of phonological variation in Kimberley Kriol that I ran through Praat's Stochastic OT learning algorithm; these three models use individual lexical items (Trial 1), individual phonemes (Trial 2), and natural classes (Trial 3) as their respective inputs. The frequencies of specific phonological variants predicted by Praat's learned grammar more closely resembled the actual frequencies observed in the data as the inputs became more abstract across the trials. These results suggest that the most abstract model, Trial 3, is the most accurate and effective way to represent phonological variation in Kimberley Kriol.

1 Introduction and Hypothesis

This paper is a study on the phonology of Kimberley Kriol, a creole language spoken in the Kimberley region of Australia. I address the question of how to account for free variation among stops and fricatives in the upper mesolect of Kimberley Kriol by using the framework of Stochastic Optimality Theory, and I contextualize the results within broader theories of the creole continuum and modeling variation in language. The Stochastic OT analysis in this paper compares three different models run through Praat's OT Learning Program that use individual lexical items, individual phonemes, and natural classes as their respective inputs; the results of these models suggest that the most abstracted version of the data—the natural classes model—is the most accurate and effective way to represent phonological variation in Kimberley Kriol.

The contribution of this paper is relevant in numerous subfields. For one, creole studies are valuable for the linguistics community at large, particularly in relation to the theory of Universal Grammar. By approaching Kimberley Kriol from an OT viewpoint, I am automatically placing it within the framework of a universal set of constraints on linguistic outputs. In the topic of variation, a Stochastic OT account of free variation in the phonology of Kimberley Kriol will add to the growing number of studies on free variation in language and the grammar's obligation to account for it as part of language competence. Furthermore, Kimberley Kriol itself has next to no literature about it, so my paper serves as a contribution in that area as well, both linguistically and socially.

The following subsections provide background information about creole studies (Section 1.1), Kimberley Kriol (Section 1.2), and Optimality Theory (Section 1.3), while the final subsection of Section 1 presents my hypothesis (Section 1.4). Section 2 describes the methods and materials used in this study, and Section 3 provides the Optimality Theory analysis and the results from the Stochastic OT models run in Praat (Boersma and Weenink 1992-2010). In the final section (Section 4), I discuss the data depicted in Section 3 and draw conclusions from the results.

1.1 Creole Linguistics

There is a great deal about the study of pidgins and creoles that is not agreed upon within the linguistic community. The origin and development of creoles, the idea of creoles as more 'simple' than other natural languages, the argument for creolization as evidence for Universal Grammar, the extent of variation in creoles and how to account for it, the idea of 'decreolization,' and even the very definitions of 'pidgins' and 'creoles' are constantly debated by various experts in the field. In the following sections, I discuss a number of these topics.

1.1.1 Pidgins and Creole Formation

Although there are many different theories about the specifics of the 'life-cycle' of pidgin and creole languages, it is undisputed that pidgins are contact languages that arise as a method of communication among groups of people in the same area that do not speak the same language (Siegel 2008:1). Typical locations of pidgin languages include port cities and work environments with laborers from a variety of locations. Over time, a pidgin can expand its grammar and lexicon and become a primary language, which generally results in creolization (Siegel 2008).[1] Unlike pidgins, a creole is a fully-fledged language with a complete lexicon, complex grammatical rules, the full range of descriptive capabilities, and a community that speaks it as a first language (Siegel 2008, Mühlhäusler 1980). Like their pidgin predecessors, creoles are formed from a socially, politically, and/or economically dominant superstrate language that provides the bulk of the lexicon and one or more substrate languages that

---

[1] This is not *always* the case; some creoles develop independently from pidgins in areas that are not typical environments for pidgin formation (Mühlhäusler 1980: 32).

contribute to other aspects of the language (Siegel 2008: 1).[2] After its formation, a creole often develops into a number of different varieties; these varieties are described as exiting on a continuum, which is discussed below.

1.1.2 Decreolization and the (Post)-Creole Continuum

One of the most extensively discussed and arguably most controversial aspects of pidgin and creole studies is the idea of the 'post-creole continuum' or the 'creole continuum.' The theory of the post-creole continuum was pioneered by David DeCamp, who described it as the stage after the formation of a creole whereby the creole grammar is gradually restructured in order to more resemble and eventually unite with the grammar of the superstrate language (1971). DeCamp clarifies that not every creole goes through this process; he says that in order for decreolization to occur, the superstrate language of the creole must be the "dominant official language" and "…there must be sufficient social mobility to motivate large numbers of creole speakers to modify their speech in the direction of the standard" (1971: 306-7).

The creole continuum is taken to be a continuous, overlapping series of slightly different varieties or lects with "lexical, phonological, and grammatical features ranging from those closest to a standard form of the creole's lexifier language…to those furthest from the lexifier language, and therefore the most 'creole-like'" (Siegel 2008: 235). The lect closest to the substrate is referred to as the *basilect*, the lect closest to the superstrate is called the *acrolect*, and the lects that span the middle of the spectrum are collectively called the *mesolect*, which is sometimes split into *lower*, *upper*, and *mid-mesolect* (Sandefur 1986: 5). Speakers typically use more than one lect, and alternation between these lects is attributed to social context in many cases (although not always). Additionally, the lects are not thought of as entirely separate units, but rather grammars that bleed into each other on a continuous scale, making "non-arbitrary division" impossible (Sandefur 1986: 5).

DeCamp's 'post-creole continuum' has been highly influential in the field, but it has also raised a number of protests. Many linguists take issue with the idea of a 'post' creole continuum because it implies that the creole in question no longer exists or has become "unrecognizable" (Bickerton 1980: 110). Bickerton specifies that decreolization takes place when a creole

---

[2] For example, Haitian Creole has French as its superstrate and a number of African languages as substrates.

language remains in "direct contact" with its superstrate language (1980: 109). Other linguists object to the idea that change and variation occur unidirectionally along the continuum towards the superstrate (Rickford 1980: 176). Rickford points out that the 'post-creole continuum' viewpoint ignores a whole range of social variables that might prefer the substrate over the superstrate, such as ties to ethnic, nationalistic, or familial identity (1980: 177). Regardless of the specifics of its definition, the creole continuum is a useful tool for analyzing variation in creoles. For the purposes of this paper, I do not operating under the assumption that the creole continuum represents a gradual change towards the superstrate language. I simply talk about the creole continuum as a series of slightly different 'lects' spanning from the substrate language or languages to the superstrate language, with the lects at the substrate end displaying more features of the substrate languages and the lects at the superstrate end showing more features of the superstrate language.[3]

1.1.3 Creole Phonology

While the lexicon of creole languages draws primarily from the superstrate language, other aspects of the creole language often exhibit influence from the substrate languages. Although substrate influence has received substantial attention within the realm of morphosyntax and the lexicon, work on substrate influences in phonology—and creole phonology in general—has been comparatively sparse (Migge and Smith 2007; Klein 2006). Notwithstanding, there are a number of proposed theories to account for substrate influence in creole phonology at the time of creolization. These include comparisons to the mapping of unfamiliar phonemes in the secondary language onto existing phonological categories in the primary language during Second Language Acquisition (SLA) and loanword adaptation (Plag 2009, Winford 2008).[4]

However, although it is useful to conceptualize the factors involved in the formation of creole phonology in order to understand how certain phonological patterns in the lexicon might have arisen, this study's discussion of Kimberley Kriol phonology focuses on speech occurring long past the time of creolization. Specifically, I am focusing on current phonological variation

---

[3] See Section 1.2 for a discussion of the relevant superstrate and substrate languages in Kimberley Kriol.

[4] Other linguists have opposed the theory of the formation of creole phonology as analogous to SLA and loanword adaptation (LA), arguing that neither theory accounts for "the [different] nature of the linguistic input to each process and the means and manners by which participants come into contact with the source" (Webb 2010: 265)

within the upper mesolect of Kimberley Kriol. Compared to the small pool of research on creole phonological formation, there are even fewer accounts of variation within creole phonology. The following section discusses variation in creoles, and Section 1.3.4 presents previous Optimality Theory accounts of phonological variation in creoles.

1.1.4 Variation in Creoles

Many linguists have noted the impressive amount of phonological, lexical, and grammatical variation in creoles, both within speech communities and individuals (DeCamp 1971; Rickford 1980; Sankoff 1980). Although there is certainly variation in other types of natural languages, variation in creoles is unique in that it is structured with respect to other natural languages, namely the superstrate and substrates. The amount and type of variation within creoles also depends on social context and the extent of a speaker's exposure to the superstrate language. All in all, variation in creoles is complex and pervasive in a way that is different from variation in other natural languages.[5] This extensive variation "has posed a challenge to current grammatical theory" (Sankoff in Rickford 1980: 166). Linguists focusing in creole studies often advocate for incorporating the variation in creoles into grammars as an integral part of language competence. As Mervyn Alleyne noted, "…creolists have been in the forefront of the battle against the monolithic assumptions of traditional linguistics, recognizing variability as a more than marginal property of language systems and speech communities" (1980: 11). The concept of the creole continuum is a way to describe variation across creole speaking communities.

In the creole continuum, different grammatical, lexical, and phonological variants theoretically correspond to different lects along a continuum (DeCamp 1971; Rickford 1980). DeCamp (1971) analyzed the various lects within the Jamaican Creole English continuum by a technique of "implicational scaling"; DeCamp created a scale by coding how "acrolectal" a number of grammatical, lexical, and phonological variants were, and then he mapped individuals' speech patterns onto this scale, thereby presenting an analysis of "a series of lects, which, while different from each other in at least one respect, were not discrete, but related as part of an implicational series…" (in Rickford 1980: 167). Through this method, DeCamp was able to roughly describe the grammars of specific lects along the continuum, while allowing for

---

[5] In the case of Kimberley Kriol, speakers of Standard Australian English do not show the kind of variation across phoneme categories that speakers of Kimberley Kriol do. See Section 1.2.3.

these grammars to bleed into each other instead of standing as individual units. Other linguists (e.g. Bickerton 1973, Rickford 1973) have expanded upon DeCamp's implicational scaling by narrowing the range of variants to one paradigm or system (such as pronouns), allowing for more than a binary set of variants within that subcategory (such as three variants for the third person singular masculine pronoun), and allowing for "variation at the level of the individual lect" (Rickford 1980: 167-8). These changes to DeCamp's system assist in the creation of 'panlectal grids,' "in which all of the possible isolects between the basilect and acrolect are specified, each isolect differing from the one next to it in just one respect" (Rickford 1980: 166). Although I do not analyze my data via the method of implicational scaling, it is important to mention because of its attention to individual lects and its capability to describe variation within individual lects. Moreover, the majority of these implicational scaling models have focused on morphosyntactic rather than phonological variants, which is not the area that I examine in this paper.

It is worthwhile to note the relation between linguistic variation and linguistic change. This topic is also debated among creolists, some of which assert that linguistic variation along the continuum is the "synchronic aspect of linguistic change" (Bickerton 1975: 16 in Rickford 1980: 175), and others of which claim that synchronic variation is not necessarily or always indicative of long-term linguistic change in progress (Rickford 1980: 176). In the case of creoles, variation as an indication of linguistic change would—in many cases—support the idea of decreolization and the *post*-creole continuum. However Rickford's argument against the unidirectionality of the creole spectrum (presented in Section 1.1.2) also applies to the directionality of linguistic change; he suggests that "while there are pressures in creole continua encouraging movement in the direction of the acrolect or standard language, there are also pressures favouring the basilect or creole language" (1980: 177). In my own research for this paper, I address synchronic variation in the phonology of Kimberley Kriol; I do not attempt to make any claims about diachronic linguistic changes occurring in Kimberley Kriol, but that does not mean that I am not open to an analysis of the synchronic variation as indicative of long-term change.[6]

---

[6] Sandefur (1986) asserts that Kriol is not undergoing decreolization in most areas in which it is spoken. However, I will not address the debate of variation vs. change in this study.

1.2 Kimberley Kriol

In my discussion of creole phonology, I use data from Kimberley Kriol, a dialect of Australian Kriol spoken in the Kimberley region of Australia. Very little linguistic work has been done on Kriol, and even less of the existing literature about Kriol deals with the Kimberley dialect. Most of the published information that I have access to refers to dialects of Kriol that are geographically close to Kimberley region, but they are not specifically about Kimberley Kriol. The superstrate of Kriol is Standard Australian English, with various Aboriginal Australian languages as the substrates. Kriol is now the first language of about 20,000 Aboriginal Australians, and it is spoken mostly in the northern part of the country (Schultze-Berndt et. al 2012: 241).

1.2.1 Historical Background

Like most creoles, Australian Kriol began its life cycle as a pidgin language. Soon after the arrival of the British at Port Jackson (modern-day Sydney), a pidgin began to form from the interaction between English and Australian Aboriginal language speakers (Sandefur 1986: 18). As more Europeans settled around the continent, pidgin spread accordingly such that "by the early part of the 20th century pidgin had gained wide usage as a lingua franca throughout most of outback Australia" (Sandefur 1986: 18). Typical historical analyses of Australian Pidgin English state that it was brought from its origin in Port Jackson, New South Wales to Queensland, the Northern Territory, and Western Australia by the movements of European pastoralists (Sandefur 1986: 19). There is no consensus on the exact location or nature of the creolization of Australian Pidgin English, but most scholars believe that creolization first took place at a Christian mission at Roper River in the early twentieth century (Schultze-Berndt et. al 2012: 242).

In the Kimberley region specifically, there were two distinct pidgins prior to creolization—an eastern pidgin and a western pidgin. According to Sandefur, the eastern pidgin was directly related to Roper River Kriol and underwent creolization in the 1940s; Kriol was then brought to towns in the western Kimberley region (1986: 21). Kununurra, the town in the Kimberley region where I collected my data, was founded by white Australians in the 1960s as an agricultural hub, but it attracted a large population of Aboriginal Australians from surrounding areas—including some former cattle stations—as well as an influx of white Australian workers. Kriol is the primary language of the majority of the Aboriginal community in Kununurra, with only a few fluent speakers of the indigenous language Miriwoong and no remaining fluent speakers of its close relative, Gajirrabeng.

1.2.2 Aboriginal English

It is important to mention that Kriol is distinct from Aboriginal English (AE), which is a non-standard dialect of Standard Australian English spoken by the Australian Aboriginal community. In fact, "some varieties of AE are linguistically very close to or are identical with 'white [non-standard] Australian English (Kaldor and Malcolm 1982, Eagleson 1982b)'" (Sandefur 1986: 26). Sandefur points out that most dialects of Aboriginal English generally have entirely different grammatical structures than those that appear in Kriol (1986: 28). Additionally, in many cases, Kriol is not historically connected to Aboriginal English, although there are some cases of regional dialects of Aboriginal English occurring "as a result of decreolization" (Sandefur 1986: 27-8).

1.2.3 Phonological Variation

Similarly to other creoles, there is a fair amount of variation in Kriol (Sandefur 1986). Apart from geographic variation and variation along the continuum within the community due partially to social context (Sandefur 1986: 49), there is individual variation even within one utterance (Sandefur 1986: 51). One of the most noticeable areas of variation is the phonology of Kriol. Sandefur says, "With a few exceptions, every stream of Kriol speech will contain some words with heavy [basilectal] pronunciations and some with light [acrolectal] pronunciations. Within the same conversation and even within the same sentence, it is not uncommon for Kriol speakers to use more than one of the pronunciation alternatives" (1986: 51).[7] Sandefur also says,

> …most Kriol speakers control virtually all pronunciations in their active everyday speech. No Kriol speaker speaks with a consistently light pronunciation. There are, however, some Kriol speakers who tend to have consistently heavy pronunciation in Kriol. These are mostly mother tongue speakers of a traditional language who speak Kriol as a second language and who speak no (Aboriginal) English. (Sandefur 1986: 51).

---

[7] Sandefur makes the argument that the 'light' end of the Kriol spectrum is not as close to the superstrate language as acrolects in other creole continua are and thus avoids the terms 'basilect,' 'mesolect,' and 'acrolect' (1986: 50); however, I continue to use these terms.

This kind of variation within the speech of individuals is what I focus on in this paper.

1.2.4 Phonology of Kimberley Kriol

Although the lexicon of Kriol is English-based, the phonology draws in part from the substrate indigenous Australian languages. Across the continent, indigenous languages—including Miriwoong and Gajirrabeng, the languages spoken in Kununurra—exhibit some general phonological trends whose effects are apparent in the phonology of Kriol.[8] These trends include a lack of voicing contrast in stops, a lack of fricatives, limited consonant clusters, and more. The effects of these phonological patterns are most evident in the basilect. Sandefur states:

> The extreme heavy phonological subsystem is virtually identical with that of traditional Aboriginal languages. Typically this means, for example, no affricates, no fricatives, no contrastive voicing with stops, no consonant clusters within a syllable, but five points of articulation for stops and nasals. The extreme light subsystem, in contrast, includes virtually all the contrasts which occur in English. (1986: 50-1).

As Sandefur's observation suggests, the phonologies of the lects within the mesolect exhibit a combination of phonological patterns of Aboriginal languages and English.

The lect my data falls into seems to be within the mesolect but on the acrolectal side of the spectrum; therefore, I refer to it as the upper mesolect.[9] Based on the data from my speakers, I have determined that the upper mesolect they represent has fricatives, as seen in such underlying minimal pairs as /si/ ('see') and /ti/ ('tea').[10] There is no evidence of phonologically contrastive voicing in fricatives, so I posit that fricatives in Kimberley Kriol are underlyingly

---

[8] There are clearly some exceptions to these generalizations, but they hold for the indigenous languages spoken in Kununurra. Also, these generalizations are apparent in Kimberley Kriol and Kriol as a whole.

[9] Given the social context of my recordings (see Section 2.2.2), it is also possible that the lect of my participants is, in fact, the acrolect. However, it is difficult to delineate lects perfectly or to determine the end point of the spectrum without more data, so the safest assumption to make is that my speakers do not represent the absolute end of the continuum but rather some part of the mesolect. I remain agnostic as to whether or not what I am referring to as the 'upper mesolect' of Kimberley Kriol is one individual lect or a chunk of lects on the creole continuum. As mentioned before, it is difficult to delineate between individual lects, and the variation I observed could be within one lect or across multiple lects. Either case is entirely compatible with my analysis, especially because the ultimate goal is to account for variation across the entire continuum. For convenience's sake, I simply refer to my data as belonging to the 'upper mesolect.'

[10] See below for a discussion of the coronal fricative /s/ and its variants.

voiceless: /f/, /s/,and /θ/.[11] Additionally, the upper mesolect has a voicing contrast in stops—as seen in such minimal and near minimal pairs as /bak/ ('bark') vs. /pak/ ('park'), /daʊg/ ('dog') vs. /taʊk/ ('talk'), and /gɝl/ ('girl') vs. /kɪd/ ('kid').[12] There were numerous lexical items that showed consistent voicing or lack of voicing across all of their tokens, which also lends support to the conclusion that voicing is contrastive in stops. Therefore, I propose that the stop inventory is /p/, /b/, /t/, /d/, /k/, and /g/.[13]

There is a fair amount of phonological variation within the upper mesolect. Although fricatives do not appear to be contrastive in voicing, voiced fricatives do appear in some contexts. However, they only appear immediately after a sonorant; in this position, voiceless and voiced fricatives are in free variation. In sum, voicing in fricatives is not contrastive, but voiced fricatives are allophonically conditioned (although variable) after sonorant segments. Because voicing only occurs in a specific phonological position, I posit that the underlying forms of all fricatives are voiceless, and they have both voiceless and voiced fricatives as allophones.[14] For example, the labial continuant /f/ can appear as voiceless /f/ anywhere and voiced /v/ only after a sonorant segment. This is the same for the coronal continuant /s/—which can surface as /s/ anywhere and /z/ only after a sonorant segment—and the coronal anterior fricative /θ/, which can

---

[11] This is not necessarily a complete list of the fricatives in Kriol, but it is the ones for which I have word-initial and intervocalic tokens and thus the ones I analyze. For example, /ʃ/ is possibly a phoneme in the language. However, I do not have enough tokens to determine its status as such.

[12] The words in parentheses are the English translations; they mean the same thing in Kriol, but this is not Kriol orthography.

[13] Like the fricatives, this stop inventory does not necessarily represent the complete list of the stops in Kriol. Rather, it the list of stops for which I have word-initial and intervocalic tokens and thus the ones I analyze in this study. For example, /c/ is possibly a phoneme in the language. However, I do not have enough tokens to determine its status as such in the word-initial or intervocalic positions, which is what this paper focuses on.

[14] I considered representing the underlying fricatives as underspecified for voicing. However, underspecification generally poses a problem for Optimality Theory (although there have been some successful OT analyses using underspecified segments: Ito et al. 1995, Inkelas 1994, Inkelas et al. 1997). Given that voiced fricatives occur in predictable environments and voiceless fricatives occur elsewhere, I decided to represent the underlying forms as voiceless fricatives rather than underspecified for voicing. In terms of voicing in stops, I followed Artstein's argument that in the case of unpredictable variation between a marked and unmarked form, the marked form should be the underlying representation because otherwise "this forces the grammar to have structural constraints that favor marked elements" (1998: 13). Therefore, I represented any form with variation in voicing of the word-initial stop as underlyingly voiced.

surface as /θ/ anywhere and /ð/ only after a sonorant segment.[15] The fricative /h/ behaves differently, preferring deletion over changing to the glottal stop.[16]

Apart from variation in fricative voicing after sonorant segments, underlying fricatives sometimes appear as voiceless stops (and voiceless affricates in the case of /s/), and underlying voiced stops sometimes appear as their voiceless counterparts across multiple tokens of the same lexical items.[17] These alternations are in free variation. For underlying fricatives, voiced stops only appear as variants intervocalically due to variable intervocalic voicing and the marked nature of fricative segments;[18] in this position, the voiced stop variants are in free variation with voiceless stops, voiceless fricatives, and voiced fricatives. Additionally, the coronal fricative /s/ can surface as the coronal affricate /tç/, although I have no examples of variants with intervocalic /s/ surfacing as the voiceless stop /c/, the voiceless affricate /tç/, or the voiced affricate /dʝ/.[19]

One thing to note is that all of the phonemes I have described have multiple non-contrastive phonetic realizations. However, I have attributed this to pure phonetics and thus have not included them in the phonological analysis. The chart below lists the phonetic realizations for the sake of thoroughness.

Additionally, there is a process of intervocalic voicing of fricatives and frication and voicing of stops that applies variably across tokens of the same lexical items. I have chosen not to address intervocalic /t/ and /d/ because it seems that there is a separate phonological process of

---

[15] In the case of /θ/, the only tokens where it surfaces as /ð/ are intervocalic. I have no tokens where it surfaced as /ð/ after a sonorant across a word boundary. See Section 2.2.3 for a discussion of this.

[16] The deletion of the fricative /h/ sometimes results in a pre-glottalization of the resulting word-initial vowel. However, the data indicates that it is not phonologically contrastive and therefore not a phonemic distinction, but rather a phonetic effect. Additionally, the fact that it appears to be more like pre-glottalization rather than a full glottal stop supports the claim that the grammar is deleting the /h/ rather than changing it to a glottal stop. Therefore, I say that the alternation is between /h/ and deletion of that /h/, although the minute phonetics of the deleted /h/ may involved pre-glottalization of the vowel in some cases.

[17] While examining my tokens, I did find two lexical items with an underlying initial *stop* that had one variant each that surfaced with a fricative: /pɪtçə/ → [fɪtçə] (‘picture’) and /baɪtəm/ → [βaɪtəm] (‘bite’). However, these two instances were both uttered by the same speaker in the same phonetic context, and they were one out of a total of two and three tokens, respectively, for the lexical items. Given the scarcity of the tokens for these two lexical items and the possibility of an idiolectal phonological process, I did not include these forms in my analysis and thus did not have to account for underlying initial stops erroneously surfacing as fricatives.

[18] See the *FRIC constraint used in Section 3.

[19] See Section 2.2.3 for a discussion of accidental gaps in the data.

tapping of the consonant and/or glottalization of the following vowel.[20] From the observations presented above, there are some general tendencies that the natural classes display. Voiceless stops are invariant word-initially, but they vary with voiced stops, voiced fricatives, and/or voiceless fricatives intervocalically. Voiced stops vary with voiceless stops word initially and with voiced fricatives intervocalically. Finally, voiceless fricatives vary with voiceless stops word-initially after a non-sonorant; voiceless stops and voiced fricatives word-initially after a sonorant; and voiceless stops, voiced stops, and voiced fricatives intervocalically.

Below is a comprehensive table of the fricatives and stops of the upper mesolect. The first column presents the underlying phonemes; the second column gives the allophones of each underlying phoneme, which are not necessarily separate phonemes but represent salient phonological distinctions; and the third column presents the possible phonetic realizations of the allophones, which are phonetically different but are neither separate phonemes nor phonologically salient.[21]

---

[20] I have determined that the tapping process is separate from the intervocalic processes affecting other oral stops because it occurs in a more specific context and does not result in frication. The process seems to be identical to that of English, whereby intervocalic coronal stops [t] and [d] lenite to the alveolar tap in onset position of unstressed syllables in natural speech.

[21] The third column presents the recorded phonetic variants of the allophones. However, these variants are not salient to speakers, nor are they phonologically contrastive. This dichotomy is present in all languages. For example, because there is no phonological distinction between dental and alveolar stops in English, the sounds /t/ and /t̪/ would both be interpreted as belonging to the phonological category /t/, even though they are phonetically distinct. In other words, although there is a phonetic difference between the two stops, that difference is not salient to the speakers and therefore does not indicate an allophonic or phonemic difference.

Phoneme	Allophonic Variants[22]	Possible Phonetic Realizations[23]	In Free Variation With…
/f/	[f]	f, ɸ	[p] word-initially; [v] and [p] after [+son] across a word boundary; [v], [p], and [b] intervocalically (when /f/ is underlying)
	[v]	v, β, ʋ[24]	[p] and [f] after [+son] across a word boundary; [b], [p], and [f] intervocalically (when /f/ is underlying)
	[p]	p	[f] word-initially; [f] and [v] after [+son] across a word boundary; [f], [v], and [b] intervocalically (when /f/ is underlying)
	[b]	b	[p], [f], and [v] intervocalically (when /f/ is underlying)
/s/[25]	[s]	s, ç, ʃ, θ	[c] and [tç] word-initially; [z], [c], and [tç] after [+son] across a word boundary; [ɟ] intervocalically (when /s/ is underlying)
	[z]	z, ʝ, ʒ	[s], [c], and [tç] after [+son] across a word boundary (when /s/ is underlying)
	[tç]	tç, tʃ, ts	[s] and [c] word-initially; [z], [c], and [s] after [+son] across a word boundary (when /s/ is underlying)
	[c]	c	[s] and [tç] word-initially; [z], [s], and [tç] after [+son] across a word boundary (when /s/ is underlying)
	[ɟ]	ɟ	[s] intervocalically (when /s/ is underlying)
/θ/	[θ]	θ	[t] word-initially; [ð], [d], and [t] intervocalically (when /θ/ is underlying)
	[ð]	ð	[d], [t], and [θ] intervocalically (when /θ/ is underlying)
	[t]	t, t̪[26]	[θ] word-initially; [ð], [d], and [θ] intervocalically (when

---

[22] There are some phones missing in the list of allophonic variants that should theoretically be possible based on general observations across natural classes. See Section 2.2.3 for a discussion of accidental gaps in the data.

[23] None of these phonetic variants appeared to be phonologically contrastive and were faithful to the general place of articulation, so I did not encode them in the phonological representations.

[24] The occurrence of the labiodental approximant /ʋ/ is why I referred to the phoneme /f/ as a labial voiceless continuant. The approximant matches /f/ in terms of continuancy and place of articulation, but not stridency. I do not address the specifics of these phonetics.

[25] /s/ and /θ/ have a slight overlap in their phonetic realizations, namely /θ/. This is due to two tokens with /s/ surfacing as /θ/ from the same speaker. It is possible that this is something specific to the speaker's idiolect and is thus not representative of variation in the upper mesolect as a whole. In that case, the phonetic realizations of /s/ would not include /θ/.

			/θ/ is underlying)
	[d]	d, d̪	[ð], [t], and [θ] intervocalically (when /θ/ is underlying)
/h/	[h]	h	Ø anywhere when /h/ is underlying
	Ø	Ø	[h] anywhere when /h/ is underlying
/p/	[p]	p	invariant word-initially; [v] intervocalically (when /p/ is underlying); [b] word-initially (when /b/ is underlying)
	[v]	v, β, ʋ	[p] intervocalically (when /p/ is underlying)
/b/	[b]	b	[p] word-initially; [v] intervocalically (when /b/ is underlying)
	[v]	v, β, ʋ	[b] intervocalically (when /b/ is underlying)
/t/	[t]	t, t̪	invariant (when /t/ is underlying); [d] word-initially (when /d/ is underlying)
/d/	[d]	d, d̪	[t] word-initially (when /d/ is underlying)
/k/	[k]	k	[g] word-initially (when /g/ is underlying); [g], [x], and [ɣ] intervocalically (when /k/ is underlying); invariant word-initially (when /k/ is underlying)
	[x]	x	[k], [g], and [ɣ] intervocalically (when /k/ is underlying)
	[ɣ]	ɣ	[k], [x], and [g] intervocalically (when /k/ is underlying)
/g/	[g]	g	[k] word-initially (when /g/ is underlying); [k], [x], and [ɣ] intervocalically (when /k/ is underlying); [ɣ] intervocalically (when /g/ is underlying)
	[ɣ]	ɣ	[g] intervocalically (when /g/ is underlying)

There is some evidence for this analysis of the consonant inventory of the upper mesolect of Kriol from such sources as Sandefur's *An Introduction to Conversational Kriol* (1982) and a provisional Kriol-English dictionary compiled by the Western Australia Department of Education and Training and the Kimberley Language Resource Center (2006), both of which seem to emphasize the voicing distinction in stops but not in fricatives in their written representations of lexical items. Sandefur also includes commentary about free variation between stops and fricatives and voicing in stops in his manual (2006). However, both of these sources use orthography rather than phonetic transcriptions; Kriol orthography is not

---

[26] Although there is a distinction between dental and alveolar stops in the substrate languages, Miriwoong and Gajirrabeng, the distinction does not appear to be contrastive in Kimberley Kriol. Therefore, /t/ can phonetically be /t/ or /t̪/, while /d/ can phonetically be /d/ or /d̪/.

standardized, so while these sources suggest an analysis similar to my own, it is difficult to determine for sure if the interpretations align.

1.3 Optimality Theory

Optimality Theory (OT), first introduced by Prince and Smolensky (1993), stands in contrast to rule-based theories of phonology. Optimality Theory dictates that an optimal output form is selected from a set of candidates based on a ranking of violable well-formedness constraints; the candidate that minimally violates the constraints in the given ranking is selected as the optimal candidate and thus appears as the surface form. These candidates are evaluated in parallel instead of subject to a series of ordered rules, as a in a rule-based theory. Additionally, the set of constraints in OT is proposed to be universal, and the phonologies of languages theoretically differ solely in the ranking order of the constraints.[27] OT is able to account for a range of phonological phenomena that rule-based theories cannot explain; crucially for this paper, OT provides an arguably more realistic explanation of free variation.[28]

1.3.1 Optimality Theory and Variation

Standard versions of OT require irreflexivity ("no constraint can be ranked above or below itself"), asymmetry ("if *x* is ranked above *y*, it cannot be ranked below *y*"), transitivity ("if *x* is ranked above *y* and *y* is ranked above *z*, then *x* is ranked above *z*"), and connectedness ("every constraint is ranked with respect to every other constraint" (Anttila and Cho 1998: 36). However, in order to account for variation, some of these principles must be violated, namely asymmetry and connectedness. The following versions of OT provide alternate ways of representing phonological variation.

1.3.2 Partially Ordered Constraints

One method of accounting for variation, as proposed by Anttila (1997 *et seq.*), is the theory of Partially Ordered Constraints (POC). This theory suggests that rather than having a fully-determined ranking of constraints in one grammar, the speaker has multiple grammars that each has a different, set ranking of the constraints relevant to the variation. These specific constraints

---

[27] This description of Optimality Theory is paraphrased from (Ito et al. 1995).

[28] Here an elsewhere, I am using "free variation" to mean stylistic free variation that is not dependent on phonological context.

that account for the different variants are unranked in the speaker's grammar. At the moment of evaluation of a candidate, one of the possible grammars with a set order of constraints is randomly selected: "When some of these total orders pick different candidates as optimal, variation results" (Coetzee and Pater 2009: 7). Additionally, the POC theory provides a theoretical statistical calculation of the probability that a specific candidate will win: "if a candidate wins in *n* tableaux [where each tableaux represents one of the possible total orders given by the partially ranked constraints] and *t* is the total number of tableaux, then the candidate's probability of occurrence is *n*/*t*" (Anttila and Cho 1998: 39). The POC theory accounts for variation and provides some measure of the probability that certain variants will arise.

However, there are a number of drawbacks to the POC version of Optimality Theory. For one, the theory's predictions about the probability of certain candidates have been argued by many to be too strong because they determine a rigid probability for a given candidate based on the fraction of possible fixed rankings that would result in that candidate (Coetzee and Pater 2009: 12). Boersma and Hayes (2001), for example, suggest that the POC would not be able to account for "cases that involve large disparities among output frequencies" (28). Additionally, it is unclear how a speaker could possibly rerank constraints during acquisition in order to learn a POC grammar. Stochastic OT, on the other hand, accounts for non-absolute probabilities of output forms and has an associated learning algorithm that is presented below in Section 1.3.3.

1.3.3 Stochastic Optimality Theory and the Gradual Learning Algorithm

In contrast to the rigid probabilities of the POC theory, the stochastic version of OT first proposed by Boersma (1997) and further developed by Boersma and Hayes (2001) presents a theory of an unbounded, "linear scale of constraint strictness, in which higher values correspond to higher-ranked constraints" (2001: 3).[29] Each constraint has a ranking value that is conceptualized as a probability distribution with a normal (Gaussian) distribution whose standard deviation is identical to the other constraints' (Boersma and Hayes 2001: 4-5).[30] The following

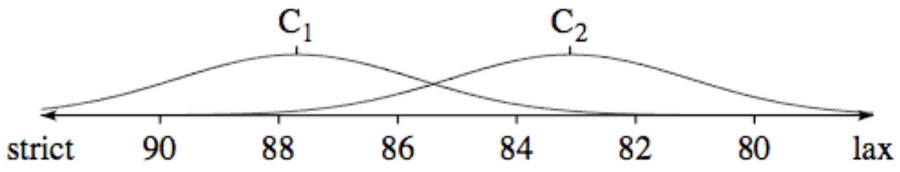
---

[29] The units of the constraint rankings are arbitrary, and when the GLA is given data to learn, it should converge on the same relative ranking and amount of overlap regardless of the initial ranking values of the constraints (Boersma and Hayes 2001: 6). Boersma and Hayes give every constraint the ranking value of 100 in the initial state of the GLA (2001: 6).

[30] The standard deviation is also an arbitrary number. Boersma and Hayes use 2.0 (2001: 5).

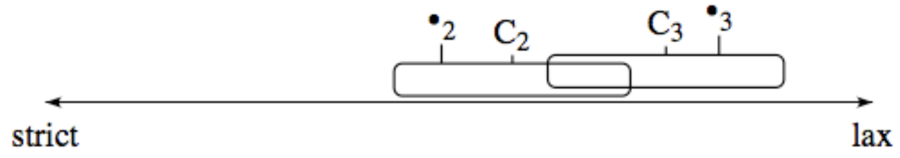
diagram from Boersma and Hayes (2001) presents a visual interpretation of constraints with overlapping probability distributions:

Overlapping ranking distributions

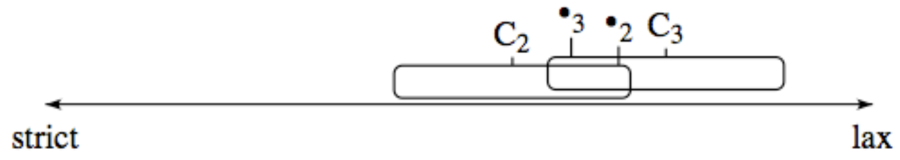


The continuous nature of the scale of constraints allows the grammar to “maintain any degree of optionality” (Boersma 1997: 2). Crucially, “at every evaluation of the candidate set, a small noise component[31] is temporarily added to the ranking value of each constraint [in order to generate a random selection point within the probability distribution], so that the grammar can produce variable outputs if some constraint rankings are close to each other” (Boersma and Hayes 2001: 1-2). A visual representation of this variation, also from Boersma and Hayes (2001), is presented below:

a. *Common result*: $C_2 >> C_3$



b. *Rare result*: $C_3 >> C_2$



In short, the amount of overlap between two constraints corresponds to the probability that one constraint will be ranked over the other in the moment of evaluation, which thus corresponds to the predicted frequency of one candidate over another (Boersma and Hayes 2001).[32] If those two candidates are different, the grammar will exhibit free variation (Boersma and Hayes 2001: 4). Thus, Stochastic OT “predicts that final-state grammars can be variable,” meaning that there is

---

[31] This noise value is within what Boersma and Hayes refer to as the ‘evaluation noise,’ which is equal to the standard deviation of the distribution of each constraint ranking (2001: 5).

[32] This is true for cases where one constraint prefers a *different* candidate than the other constraint.

free variation within those adult grammars (Jarosz 2011: 7-8). However, Stochastic OT still allows for relatively strict rankings; if two constraints are sufficiently far apart such that their overlap is negligible, they are understood to have the same effect on candidates as if they were ranked in a non-stochastic version of OT (Boersma and Hayes 2001).

Boersma and Hayes' Stochastic OT is accompanied by a language-learning model called the Gradual Learning Algorithm (GLA). The GLA is an error-driven algorithm; in other words, "it alters rankings only when the input data conflict with its current ranking hypothesis" (Boersma and Hayes 2001: 1). At the time of evaluation of a candidate, the added random noise of a negative or positive value temporarily changes each constraint's ranking value, generating a random selection point for each constraint within its standard deviation. From these selection points, the GLA creates a specific constraint ranking at that evaluation time which it uses to select the most optimal candidate. If that candidate does not match the correct (underlying) form[33], the GLA promotes the constraints that prefer the correct form instead of the selected candidate and demotes the constraints that prefer the selected candidate over the correct form (Boersma and Hayes 2001).

The amount by which the constraints are adjusted is called the *plasticity*; a large plasticity results in drastic changes in the ranking values of relevant constraints after evaluation of a candidate, whereas a small plasticity adjusts constraints at a much finer-grained level. The plasticity that Boersma and Hayes use varies (2001). In discussing plasticity, they say:

> A small plasticity value does a better job of matching learning data frequencies in the end, but a large plasticity value nears its goal faster. The virtues of the two approaches can be combined by adopting a learning schedule that decreases the plasticity as learning proceeds. This seems in principle realistic: in humans, grammar apparently stabilizes in adulthood, as nonlexical learning slows or halts. (Boersma and Hayes 2001: 35).

After exposure to all of the data, the GLA's series of small promotions and demotions of constraints' ranking values will have theoretically converged on the correct grammar (Boersma and Hayes 2001: 8).[34] In the case of free variation, the prediction of the GLA is as follows:

---

[33] The input underlying form is presumed to be an adult surface form that the learner hears and uses as the underlying form, i.e. it is entirely faithful (Boersma and Hayes 2001: 7).

[34] This is usually the case; however, there are rare cases where the GLA does not converge on the correct grammar (e.g. Pater 2008).

"with sufficient data, the algorithm will produce a grammar that *mimics the relative frequency of free variants* in the learning set" (Boersma and Hayes 2001: 9). Praat's Optimality Theory Learning Program, which is discussed in depth in Section 2.2.2, is a computational model that implements the GLA.

1.3.4 Optimality Theory and Creole Phonology

There are very few Optimality Theory accounts of creole phonology in the available literature, and even fewer OT accounts that attempt to tackle the kind of phonological variation in creole that is uniquely structured with respect to other natural languages (the superstrate and substrates). Hume and Tserdanelis (2002) use an OT framework to discuss asymmetric assimilation of nasals with different places of articulation in Sri Lankan Portuguese Creole, but she does not discuss any sort of variation and she works within Classic OT. Similarly, Klein (n.d.) presents and an account of a predictable morphophonological pattern in French-lexified Antillean Creoles, but he does not touch variation and uses an OT model called Lexical Representation as Pure Markedness, which I am not considering for my analysis. Ito and Mester (2001) discuss lexical variants within Jamaican Creole, but they attribute the different variants to individual, non-stochastic OT grammars; no single grammar that they posit allows free variation between more than one of the variants (8-9).

There are also studies that use OT to explain the phonological shape of specific lexical items in creoles adopted from the superstrate at the time of creolization, describing the pattern as "[lying] somewhere between the full range of structures found in the lexifier language and the patterns of the substrate" (Lipski 1999: 173). For example, both Abler and Plag (2001) and Lipski (1999) use OT to address variable methods of complex syllable structure repair—such as epenthesis, deletion, and faithfulness—in Sranan and Afro-Iberian creoles, respectively. However, these studies do not deal with free variation within individual lexical items in the current state of any creole language, but rather they describe how the interaction between the superstrate and substrate grammars affected *specific* lexical items *during* the time of creolization. Thus, my Stochastic OT account of free variation within the upper mesolect of Kimberley Kriol has seemingly no precedent to draw from.

1.4 Hypothesis

The following study is a Stochastic OT analysis of free variation among stops and fricatives in the upper mesolect of Kimberley Kriol. I will address the following questions: can we account for this somewhat unique free variation within the upper mesolect by using Stochastic OT, a framework that is used to describe variation in other natural languages? How does this analysis fit into theories of the creole continuum and variation in creoles? In order to explore these questions of how to represent variation in Kimberley Kriol, I run three different Stochastic OT models of the phonological grammar of Kimberley Kriol through Praat's OT Learning Program (Boersma and Weenink; 1992-2010). The first uses individual lexical items as its input, the second uses individual phonemes as its input, and the third uses general natural classes as its input.[35] I predict that the most abstracted version of the data—the natural classes trial—will be the most accurate model of the phonological variation observed in Kimberley Kriol.

2 Materials and Methods

In order to address the question of whether or not free phonological variation in the upper mesolect of Kimberley Kriol can be addressed in Stochastic Optimality Theory, I first examine tokens that I extracted from my recordings of Kriol speakers in order to record the nature and percentages of variants for each underlying form. I then infer the relevant constraint rankings that account for each attested variant, which in turn predict and inform the Stochastic OT models that I run in Praat (Boersma and Weenink; 1992-2010).

2.1 Materials

In the summer of 2012, I spent eight weeks working at Mirima Dawang Woorlab-gerring Language and Culture Centre (MDWg) in Kununurra, Western Australia in the Kimberley region of Australia. Although MDWg primarily works on the local indigenous languages, Miriwoong and Gajirrabeng, I was able to make three audio recordings of Kriol from which I pulled tokens to serve as my data. I used a variety of materials and methods to elicit natural speech in Kriol that will be described briefly below.

---

[35] See Section 2.2.2 for a more in-depth explanation of these three models.

2.1.1 Subjects

The subjects of my recordings are four Aboriginal Australians in their 30s and 40s—two men and two women—who work at MDWg. They are all native speakers of Kriol and use it as their primary language, although they are all also fluent in Aboriginal English (a non-standard variety of Australian English). None of the subjects speak an indigenous Australian language fluently, but they are all currently second-language learners of Miriwoong.

2.1.2 Methods of Elicitation

I recorded the subjects in groups of two—separated by gender[36]—in an isolated, soundproof room at MDWg using a digital recorder with a built-in omnidirectional microphone. I was the only other person present. The elicitation materials were wordlists in English that the subjects translated into Kriol, sentences in Kriol, a map task in which one person verbally directed the other to a specific location on the map, prompts to start natural conversations, and a wordless picture book from the Frog Story series that the subjects had to verbally describe. Although I was able to record two sessions with the women, I was only able to record once with the men. As a result, I do not have recordings of the men reading the Kriol sentences or describing the Frog Story.

2.1.3 Speech Data Analysis

From my data, I extracted 424 tokens representing variants of 87 underlying representations.[37] The tokens that I used were exclusively lexical words and included no grammatical words because I wanted to avoid any sort of phonological processes that might apply uniquely or fail to apply to grammatical words. The tokens fell under nine conditions:[38]

---

[36] I recorded the subjects in gender groups because of social and cultural reasons. There does not appear to be an effect of gender on my data, although it is possible that I do not have enough data from the men to rule it out entirely.

[37] Within the 87 items, word-initial fricatives after sonorants are treated separately from word-initial fricatives after non-sonorants, even if they are technically the same lexical item. For example, /faɪndəm/ ('find') is treated as a different item than /[+ son]faɪndəm/. This was necessary in order to generate the proper tableaus and constraint interactions.

[38] Some conditions had a few variants that did not arise; these are possibly accidental gaps based on scarcity of some of my tokens rather than impossible occurrences. In Condition 4, it is possible that word-initial /s/ also varies with voiced affricates after sonorants. In Condition 7, it is possible that

1) Word-initial underlying voiced stops (varying with voiceless stops)

2) Word-initial underlying voiceless stops (invariant)

3) Word-initial underlying voiceless fricatives after non-sonorants (varying with voiceless stops in all cases and also with a voiceless affricate in the case of /s/)

4) Word-initial underlying voiceless fricatives after sonorants (varying with voiced fricatives and voiceless stops in all cases, and also with voiceless affricates in the case of /s/ after sonorants)[39]

5) Word initial underlying /h/ after non-sonorants (varying with Ø)

6) Word-initial underlying /h/ after sonorants (varying with Ø)

7) Intervocalic underlying voiced stops (varying with voiced fricatives)

8) Intervocalic voiceless stops (varying with voiced stops, voiced fricatives, and voiceless fricatives)

9) Intervocalic voiceless fricatives (varying with voiced fricatives, voiceless stops, and voiced stops)[40]

Some of the tokens I extracted fulfill more than one of these conditions. I transcribed each token, confirming my transcriptions with spectrogram analysis in Praat. For every underlying form that had a voiced and voiceless stop as the only word-initial variants, I posited the voiced stop as the underlying representation and therefore placed it in Condition 1. Similarly, for every underlying form that had a fricative as one of its word-initial variants, I posited that the voiceless fricative was underlying (Conditions 3-6).[41]

---

intervocalic voiced stops also vary with voiceless stops and voiceless fricatives. My data doesn't reflect this, but is possible that I just didn't have any tokens for these variants.

[39] There was only one lexical item where an initial fricative after a sonorant surfaced as a voiced stop, and it was for one out of the two tokens for the lexical item /fɪʃ/ ('fish') after a sonorant. This lexical item also surfaced with a voiced stop for three out of the four tokens where the word did not occur after a sonorant. This was the only example of a voiceless fricative surfacing as a voiced stop in any word-initial context. Given the scarcity of tokens for the word and the potential that the word is somehow treated differently, I judged it to be anomalous and therefore did not account for it in my analysis.

[40] I only have one token where an intervocalic fricative surfaced as a voiceless stop: /lafɪn/ → /lapin/ ('laughing'). The analysis accounts for it, but it may actually be anomalous.

[41] Although all of the lexical items for which I posited underlying fricatives have fricatives in their English counterparts, the reverse is not true. Not all of the lexical items in the token list whose English counterparts have fricatives exhibited underlying fricatives in their Kriol forms.

For intervocalic segments, determining the underlying representation was more complicated. In cases where the consonant was intervocalic because it was previously at the end of a word that then received a suffix (e.g. /kʌk/ vs. /kʌkəm/, ‘cook’), I based my analysis of the underlying segment on a comparison with the unaffixed Kriol form. Sometimes—although this was not ideal—I had to rely on a comparison with the English counterpart of the Kriol word if no related Kriol form was available. After looking at my data, I decided not to address intervocalic /t/ and /d/ because it appears that there is a separate phonological process of tapping at play in this context. See footnote 20 in Section 1.2.4 for a more detailed explanation.

Additionally, there is undoubtedly variation in both vowel quality and stops and fricatives word-finally and in clusters (word-initial, -medial, and -final). However, because I am not addressing variation in these contexts in this study, I ignored that variation in my transcription of underlying forms and represented varying segments as consistently faithful to the underlying form if they were not singleton word-initial or intervocalic stops or fricatives. Likewise, I did not address the possibility of variation within grammatical suffixes or across the morpheme boundary of compound words—even when their affixation resulted in an intervocalic singleton stop or fricative—because of the potential for phonological processes to affect these forms differently.[42] In this case, I also transcribed the tokens as faithful to the underlying form for all segments apart from other non-initial and intervocalic singleton stops and fricatives. I also noted phonological environment for each token, which allowed me to confirm that the alternations of underlying fricative vs. stop, underlying voiced stop vs. voiceless stop, and underlying /h/ vs. Ø are, in fact, in free variation word initially.

Intervocalically, there are variable phonological processes of frication and voicing. These result in underlying voiced stops varying with voiced fricatives; underlying voiceless fricatives varying with voiced fricatives, voiced stops, and voiceless stops; and underlying voiceless stops varying with voiced stops, voiceless fricatives, and voiced fricatives. Because of these processes, there were a handful of underlying representations for lexical items that were difficult to determine. In some cases, it was hard to say when an intervocalic phoneme was underlyingly a fricative or a voiced stop. When there were tokens for a lexical item with either a

---

[42] For example, I did not consider voicing or lenition processes for the /b/ in the grammatical affix /-abat/ (continuous aspect), nor did I consider those processes for the /d/ in the second half of the compound /pʌpədaʊg/ (‘puppy’).

voiced or voiceless fricative intervocalically and the word corresponded to an English word that would have a fricative, I ended up positing the fricative underlyingly. For example, I posited that the underlying form /nɛfə/ ('never') had a fricative, but it is possible that it is actually /nɛbə/. Although I did my best not to approach the English-lexified Kriol data from an English viewpoint, this was admittedly an added confound that in cases like /nɛfə/ helped guide my analysis.

2.1.4 Datasets

After analyzing the data and deciding upon the phoneme inventory and underlying representations of each form, I created three datasets that correspond to the three trials in Section 2.2.2. Although I tried to extract equal numbers of tokens for each lexical item, underlying phoneme in each position, and natural class in each position, I was unable to do so with the recordings that I had. As a result, there are certain words, phonemes, and natural classes that have more tokens than others. This potentially affects my trials in Praat, which I discuss in Sections 3 and 4. The datasets are presented below:

Dataset 1:

This dataset is comprised of 87 underlying representations with a total of 424 tokens, where each underlying representation is an individual lexical item and each token is a single utterance of a lexical item by one of the speakers.[43] Out of the 87 underlying representations, twelve had initial /f/, eleven had initial /s/, two had initial /θ/, eight had initial /h/, thirteen had initial /b/, two had initial /p/, three had initial /d/, eight had initial /t/, six had intervocalic /f/, one had intervocalic /s/, four had intervocalic /θ/, two had intervocalic /b/, three had intervocalic /p/, five had

---

[43] This is counting words with initial fricatives as different underlying representations when they follow sonorants vs. when they do not. This does not mean that the underlying initial phoneme in "/faɪndəm/" and "[+son]|faɪndəm," for example, is different. It simply means that in order to account for the different behavior of initial fricatives following a sonorant vs. a non-sonorant, they must be entered as separate forms in separate tableaus in the codes written for the Praat program because they have different violations for some constraints.

intervocalic /g/, and ten had intervocalic /k/.[44] This dataset corresponds to Trial 1 (see Section 2.2.2).

Dataset 2:

This dataset is comprised of 21 underlying representations with a total of 481 tokens, where each underlying representation is an individual stop or fricative in word-initial or intervocalic position (as well as word-initial after a sonorant for fricatives) and each token is a single utterance of a stop or fricative.[45] These are separated into 58 tokens of word-initial /b/, 17 tokens of word-initial /d/, 46 tokens of word-initial /g/, 6 tokens of word-initial /p/, 29 tokens of word-initial /t/, 35 tokens of word-initial /k/, 27 tokens of word-initial /f/ after a non-sonorant, 14 tokens of word-initial /s/ after a non-sonorant, 4 tokens of word-initial /θ/ after a non-sonorant, 43 tokens of word-initial /f/ after a sonorant, 24 tokens of word-initial /s/ after a sonorant, 3 tokens of word-initial /θ/ after a sonorant, 9 tokens of word-initial /h/ after a non-sonorant, 9 tokens of word-initial /h/ after a sonorant, 13 tokens of intervocalic /b/, 19 tokens of intervocalic /g/, 9 tokens of intervocalic /p/, 50 tokens of intervocalic /k/, 37 tokens of intervocalic /f/, 3 tokens of intervocalic /s/, and 26 tokens of intervocalic /θ/.[46] This corresponds to Trial 2 (see Section 2.2.2).

Dataset 3:

This dataset is comprised of 9 underlying representations with a total of 444 tokens, where each underlying representation is a natural classes of stops or fricatives in word-initial or intervocalic position (separated into word-initial post-sonorant and post-non-sonorant for fricatives) and each token is an individual utterance of a segment in one of those natural classes. The results of this regrouping are 121 tokens of word-initial voiced stops, 70 tokens of word-initial voiceless stops, 31 tokens of word-initial voiceless fricatives after non-sonorants, 46 tokens of word-initial fricatives after sonorants, 9 tokens of word-initial /h/ after non-sonorants, 9 tokens of word-

---

[44] There is obviously some overlap in forms with intervocalic and word-initial segments of interest.

[45] There are more tokens here than in Dataset 1 because some of the individual lexical items have more than one phoneme of interest in them, but they became separate tokens in this dataset. Additionally, the phoneme tokens corresponding few words have been removed from Trial 2 due to lexical exception; see Section 3.2.1.

[46] These are the number of tokens for the underlying representations, not necessarily the surface variants.

initial /h/ after sonorants, 33 tokens of intervocalic voiced stops, 60 tokens of intervocalic voiceless stops, and 66 tokens of intervocalic voiceless fricatives.[47] I did not include tokens for word-initial /s/ when I abstracted to underlying representations of natural classes because of the phoneme's variance with the affricate /tç/ in that position. This eliminated 38 tokens from the 481 tokens in the list of underlying representations of individual phonemes, which resulted in the 444 tokens.[48] I did, however, include the tokens for intervocalic /s/ in the intervocalic fricative class because my data did not show variance with affricates in that context. I include word-initial /h/ as separate category in the natural classes delineation because it behave entirely differently from the other fricatives. This corresponds to Trial 3 (see Section 2.2.2). Below is a comprehensive table describing the three trials:

Trial	**Underlying Representations**	**Tokens**
1	87	424
2	21	481
3	9	444

2.1.5 Speaker Conflation

Before I move on to an explanation of my methods, it is crucial to address the issue of speaker conflation. Given the scarcity of tokens for some of my lexical items and the small amount of data from the male subjects in comparison to the female subjects, I have conflated the data as if all of the tokens came from one speaker instead of four. Thus, the three datasets represent tokens from the four speakers as a whole. Before conflating the data, I endeavored to make sure that each speaker exhibited the *general* variations I am examining, namely voicing variation in underlying voiced stops, variation between fricatives and stops, and variable deletion of /h/. Each speaker shows at least on example of these variations except for one of the male speakers, who did not have any instances of variation in stop voicing or /h/ deletion. However, the token data from that speaker in those categories is very sparse. Additionally, not every speaker

---

[47] Including /h/ as a separate category means that this dataset is not technically "natural classes," but rather a combination of natural classes and individual phonemes. While it would be ideal to include /h/ in the fricatives natural class, its unique behavior would throw off the phonological generalizations about the other fricatives. Therefore, I keep it separate from the fricatives natural class. For the sake of convenience, I refer to this dataset and trial as the "natural classes" trial throughout the paper.

[48] The same lexical items whose tokens were excluded from Trial 2 are also excluded from Trial 3. See footnote 45 and Section 3.2.1.

exhibited every potential variant for all of the nine conditions or all of the variants for each relevant phoneme. However, I propose that these missing variants and the lack of variation in two areas for one of the male subjects are due to scarcity of data rather than actual gaps in the subjects' speech patterns. Because every speaker exhibits the patterns of variation that I am examining to some extent, I maintain that it is reasonable to conflate the tokens to represent one "speaker." Although conflation is not ideal, it is the most efficient way for me to construct a grammar for the upper mesolect that encompasses all four speakers, especially given that I do not have enough data from each subject to construct separate grammars that model individual variation.

2.2 Methods

2.2.1 Analysis by Hand

Using the token data that I describe in Section 2.1, I construct an OT analysis that accounts for each type of variant within the nine conditions listed in Section 2.1.3. This analysis does not provide a stochastic ranking, but it does present the constraint rankings and tableaus for each type of variant that the stochastic analysis must account for. Practically speaking, this analysis gives rankings that represent the particular constraint rankings at the time of evaluation of a specific candidate in Stochastic OT. This is useful because it identifies clusters of constraints that interact in ways such that they necessarily overlap in Stochastic OT.

2.2.2 Analysis by Praat

In order to present a Stochastic OT analysis of the data from Kimberley Kriol, I use Praat's OT Learning Program, which implements the Gradual Learning Algorithm in Stochastic OT (Boersma and Weenink 1992-2010).[49] The learning program requires two files: a grammar file and a pair distribution file. The grammar file contains a list of the constraints, their initial ranking values, and a list of all of the input tableaus with each candidate's violations for the constraints. The tableaus that I input are based on the ones that I generate in the first part of my analysis. I set the initial ranking value of all of my constraints at the same number, 100, and I do not specify any fixed rankings between the constraints.[50] This allows the GLA to learn the

---

[49] I use version 5.1.43 of Praat for this study.

[50] These are the default settings in Praat (Boersma and Weenink 1992-2010).

grammar from the most neutral, unbiased starting point. The pair distribution file contains each input-output pair and the proportion of times that each output surfaces from any given input (i.e. the frequency that that variant occurs in adult speech). I calculate these frequencies from the corpus of token data described in Section 2.1 in order to write them into the pair distribution file.

From the pair distribution file, Praat's GLA algorithm generates two files of strings of the input and output forms, which it uses in conjunction with the grammar file to learn the grammar and generate final rankings values for the constraints. The GLA computational model then chooses the optimal candidate from a set based on the grammar it generates at each individual evaluation and adjusts constraints accordingly. I use Praat's default settings, which sets the evaluation noise (which is equal to the standard deviation of each constraint) at 2.0 and the plasticity at .1, and runs each input through the grammar 100,000 times.[51] The output grammar presents a constraint ranking and generates frequencies for each variant that the final constraint ranking predicts.[52]

This study presents three Stochastic Optimality Theory models that I run through Praat's learning algorithm.[53] Trial 1 is comprised of individual lexical items and the frequencies of each variant for each lexical item. However, because Praat's algorithm deals with overall percentages rather than absolute numbers of each token, each lexical item ends up being weighted the same no matter how many tokens it has. This gives undue weight to lexical items that have very few tokens, which would skew the results and cause the algorithm to output a grammar that predicts incorrect frequencies for specific variants. For those lexical items with many possible variants, the amount of token data is particularly crucial; for example, a lexical item with ten tokens but only two possible variants most likely gives a more accurate reflection of that item's variation than a lexical item with ten tokens but four possible variants does of its respective variation.

---

[51] Changing these settings in various ways could have a slight effect on the output grammar from Praat. For example, increasing the standard deviation would allow for more variation because it would increase the range of possible selection points for each constraint's ranking at evaluation time. As mentioned in Section 1.3.2, a large plasticity helps the GLA adjust constraints more drastically and arrive at the correct ranking sooner, while a small plasticity allows the GLA to adjust the constraint ranking at a finer level. I did not change the standard deviation from the automatic setting of 2.0 because this is the value that is widely used in the Stochastic OT and GLA literature. However, I did run trials with different plasticities and numbers of runs of the input forms. See Section 3.2 for a description of these trials.

[52] See Section 1.3.3 for a description of the GLA.

[53] See Section 2.1.4 for the datasets that correspond to the three trials.

In order to combat this potential obstacle of uneven or insufficient token data for some lexical items, Trial 2 abstracts away from the data in Trial 1. Trial 2 is not comprised of individual lexical items, but rather presents the frequency of each variant of every individual *phoneme* in each position of interest, resulting in 21 ‘underlying representations’ under which the 424 tokens will be assigned. Trial 3 is a further abstraction of the individual phonemes into overall natural classes—voiced stops, voiceless stops, and fricatives—in each position of interest, which results in nine ‘underlying representations.’ The abstractions in Trials 2-3 potentially gloss over a lack of variation in specific lexical items or any effect that is specific to one or more lexical items; however, for many of the lexical items, I do not have enough tokens to determine the actual nature or amount of variation. Therefore, I assume that their tokens represent variants in the overall pattern rather than anomalies.[54] An example of an input for Trial 1 is the lexical item /daʊg/ with its variants possible [daʊg] and [taʊg]; an example of an input for Trial 2 is the phoneme /d/ in word-initial position with its possible variants [d] and [t]; and an example of an input for Trial 3 is the class /[+voice, -continuant]/ in word-initial position with its possible variants [+voice, -continuant] and [-voice, -continuant].

In all of the three models, I endeavor to describe the general phonological tendencies of the grammar; the variable that changes from trial to trial is the type of input that the algorithm receives. The three models utilize the same constraints, which refer to general phonological features instead of processes specific to certain lexical items or phonemes. The model of the grammar should ideally be able to account for individual lexical items that abide by overarching phonological processes; however, the abstractions in Trials 2 and 3 are the most effective way to capture the phonological generalizations about variation in word-initial and intervocalic variation in stops and fricatives given the scarcity of some of my data.[55]

I compare the output frequencies of all three trials with the frequencies that I calculated based on the empirical data from the corpus data. If the frequencies match, then the learning algorithm arrived at a grammar that correctly predicts the amount and kind of variation observed in the corpus. If there is a large amount of error between the empirically observed frequencies and those generated by the GLA’s output grammar, this means that the program did not

---

[54] See Section 3.2.1 for a more in-depth discussion of the effect of small token amounts for specific lexical items.

[55] See Section 3.2 for specific examples that justify the abstractions in Trials 2 and 3.

effectively learn a grammar that accounts for the variation observed. It should also be pointed out here that assessment of the effectiveness of an output grammar from the GLA is largely subjective in the literature and is often measured in comparison with the performance of other language acquisition models (Boersma and Hayes 2001). For a baseline comparison, the output frequencies by the final grammars of Boersma and Hayes' GLA trials differed from the input frequencies of variants by an average of .39-1.09% (2001).[56] There are many possible explanations for a large average error margin that I discuss in Section 3.2 and Section 4.

3 Analysis

This section presents an Optimality Theory account of the phonological free variation in Kimberley Kriol data described in Sections 1-2. In Section 3.1, I determine the constraint orders that prefer each type of variant, thus identifying constraints that should overlap in a Stochastic OT approach. Section 3.2 provides three different models of a Stochastic OT analysis through use of Praat's OT Learning Program.

3.1 Optimality Theory Account

The markedness constraints that I will be using in this OT analysis of phonological variation in Kimberley Kriol are *FRIC, *GLOTTAL, *[+strid], *VOICEOBS, AGREE(voice), AGREE(cont), and *[+son]||[-voice]. These constraints account for fricatives variably arising as stops, the tendency of [h] to delete rather than change to a glottal stop, the tendency for coronal fricatives to arise as stops to avoid stridency, devoicing of voiced obstruents, voicing and frication of intervocalic segments, and voicing of word-initial fricatives after sonorants. The faithfulness constraints that I will be using are MAX, IDENT(cont), IDENT(voice), and IDENT(strid). These constraints account for retention of [h], fricative and stop variants that are faithful in continuancy and voicing, and the tendency for coronal fricatives to arise as affricates in order to maintain stridency. Together, these constraints account for variation within Kimberley Kriol phonology; they will be explained fully in the following subsections, which are separated by the conditions listed in Section 2.1.3.

---

[56] It is unclear in the Boersma and Hayes paper if the percentages input into the GLA reflect actual frequencies of variants from corpus data, or if the frequencies are merely theoretical. The GLA was given 21,000 or 388,000 pairs of underlying and surface representations, depending on the trial (2001). However, it is unclear if these inputs reflect actual token amounts from a corpus of data or tokens that the authors predicted would occur (Boersma and Hayes 2001).

The subsections below suggest that certain constraints will overlap in a Stochastic OT analysis. In particular, there should be close, overlapping clumps of constraints dealing with continuancy (*FRIC, IDENT(cont), AGREE(cont)), stridency (*[+strid], IDENT(strid)), voicing (IDENT(voice), AGREE(voice), *VOICEOBS, *[+son]||[-voice]), and deletion of glottal segments (MAX, *GLOTTAL). There are also constraints that should overlap across these clumps, which will likely result in most of the constraints having ranking values relatively close together. However, the results suggest that constraints that do not interact, such as AGREE(voice) and *[+son]||[-voice], should only overlap to small extent if they overlap at all.

3.1.1 Conditions 1-2

Empirical Generalization 1**:** Word-initial underlyingly voiced stops in Kimberley Kriol are in free variation with their voiceless counterparts, although word-initial voiceless stops do not vary. This is caused by the interaction of faithfulness and markedness constraints related to voicing.

In order to account for this generalization, I use the following constraints:

***VOICEOBS:** Assign one violation mark for every voiced obstruent in the output form.[57]

IDENT(voice): Assign one violation mark for every output segment whose voicing is not the same as that of its input correspondent.[58]

The following tableaus illustrate the interaction between IDENT(voice) and *VOICEOBS for word-initial underlyingly voiced and voiceless stops.

---

[57] This constraint will assign violations for voiced obstruents in positions other than word-initial or intervocalic. Although there is variation in positions other than these two, I do not analyze this variation in this paper. Therefore, although I mark all of the violations in the tableaus for each candidate, because I do not vary the voicing of obstruents in other positions, the selected output does not change. It would have had the same effect if I chose not to mark any violations for voiced obstruents in other positions.

[58] I have not include the constraint IDENT(place) because there are no winning candidates whose place of articulation differ from the input underlying forms. However, if I wanted to account for richness of the base, IDENT(place) would eliminate any potential candidates that change place of articulation.

Tableau 1:[59, 60]

daʊg	IDENT(voice)	*VOICEOBS
☞ daʊg		**
taʊg	*!	*

For this candidate, IDENT(voice) >> *VOICEOBS

Tableau 2:

daʊg	*VOICEOBS	IDENT(voice)
daʊg	**!	
☞ taʊg	*	*

For this candidate, *VOICEOBS >> IDENT(voice).

Tableau 3:[61]

kʌm	*VOICEOBS	IDENT(voice)
☞ kʌm		
gʌm	*!	*

For this candidate, the order of *VOICEOBS and IDENT(voice) does not matter.

3.1.2 Condition 3

Empirical Generalization 2: Word-initial voiceless fricatives are in free variation with voiceless stops in the same place of articulation. This results from an interaction between faithfulness constraints and markedness constraints against fricatives.

---

[59] Kriol 'dorg'; English 'dog' (Western Australia Department of Education and Training and the Kimberley Language Resource Center (2006)).

[60] The example words that I use to illustrate constraint interactions do not necessarily exhibit all of the variants themselves; however, every variant that they are employed to explain occurs as a variant of some word within the relevant condition. This applies to all of the example words throughout Section 3.1.

[61] Kriol 'kum'; English 'come' (Western Australia Department of Education and Training and the Kimberley Language Resource Center (2006)).

To account for this generalization, I use the following constraints:

***FRIC:** Assign one violation mark for every segment that is a fricative.

IDENT(cont): Assign one violation mark for every output segment whose continuancy is not the same as that of its input correspondent.

The following tableaus illustrates the interaction between IDENT(cont) and *FRIC for voiceless initial fricatives.

Tableau 4:[62]

faındəm	IDENT(cont)	*FRIC
☞ faındəm		*
paındəm	*!	

For this candidate, IDENT(cont) >> *FRIC.

Tableau 5:

faındəm	*FRIC	IDENT(cont)
faındəm	*!	
☞ paındəm		*

For this candidate, *FRIC >> IDENT(cont).

Tableau 6:

kʌm	*FRIC	IDENT(cont)
☞ kʌm		
xʌm	*!	*

Thus, for this candidate, the order of *FRIC and IDENT(cont) does not matter.

---

[62] Kriol ‘fiendim’; English ‘find’ (Western Australia Department of Education and Training and the Kimberley Language Resource Center (2006)).

***Empirical Generalization 3*:** Word-initial voiceless coronal fricatives occurring after non-sonorants vary freely with both stops and affricates.[63] This results from an interaction between faithfulness and markedness constraints dealing with stridency and frication, which cover the distinctions between stops, fricatives, and affricates.

In order to account for this variation, I use the following constraints in addition to *FRIC and IDENT(cont):

***[+strid]:** Assign one violation mark for every strident segment.

IDENT(strid): Assign one violation mark for every output segment whose stridency is not the same as that of its input correspondent.

The following tableaus shows the interactions between IDENT(cont), *[+strid], *FRIC, and IDENT(strid) for word-initial underlyingly coronal fricatives.

Tableau 7:[64]

sıtın	IDENT(cont)	IDENT(strid)	*[+strid]	*FRIC
☞ sıtın			*	*
cıtın	*!	*		
tçıtın	*!		*	

For this candidate, IDENT(cont) >> {*FRIC, *[+strid]}; IDENT(strid) unranked.

---

[63] The lack of affricate variants for labial and coronal anterior fricatives could possibly be accidental gaps in my data. Alternatively, their absence could be accounted for by positing markedness constraints for affricates in those specific places of articulation that would be ranked high. However, I do not directly address this in my analysis.

[64] Kriol ‘sidin’; English ‘sitting.’

Tableau 8:

sıtın	*[+strid]	IDENT(cont)	*FRIC	IDENT(strid)
sıtın	*!		*	
☞ cıtın		*		*
tçıtın	*!	*		

For this candidate, *[+strid] >> {IDENT(cont), IDENT(strid)}; *FRIC unranked.

Tableau 9:

sıtın	IDENT(strid)	*FRIC	*[+strid]	IDENT(cont)
sıtın		*!	*	
cıtın	*!			*
☞ tçıtın			*	*

For this candidate, {IDENT(strid), *FRIC} >> {*[+strid], IDENT(cont)}.

3.1.1 Condition 4

***Empirical Generalization 4*:** Word-initial voiceless fricatives occurring after sonorants vary freely with voiced fricatives and voiceless stops. This is a result of the interaction between faithfulness and markedness constraints related to frication and voicing.

In order to account for this variation, I use *VOICEOBS, IDENT(voice), *FRIC, IDENT(cont) in addition to the following constraint:

***[+son]||[-voice]:** Assign one violation mark for every voiceless fricative that immediately follows a sonorant segment across a word boundary.[65]

---

[65] I have made this constraint specific to fricatives as opposed to all voiceless obstruents because in my data, word initial voiceless stops never surface as their voiced counterparts, even when following a sonorant segment. It is possible that this is an accidental gap in my data and that there are instances of this variation in my subjects' speech patterns. However, given the reality of my data, broadening *[+son]||[-voice] to include all obstruents would most likely negatively impact the GLA's ability to account for the data because the constraint would predict that the voiced stop variants of word-initial post-sonorant voiceless stops would sometimes occur. Likewise, eliminating this constraint and simply using AGREE constraints that would voice obstruents between sonorants both within words and across word boundaries would predict voiced stop variants of word-initial voiceless stops. Both of these solutions

The following tableaus illustrate the interaction between IDENT(voice), *VOICEOBS, *FRIC and *[+son]‖[-voice] for word-initial underlying voiceless fricatives occurring after sonorant segments.

Tableau 10:

[+son]faɪndəm	IDENT(cont)	IDENT(voice)	*VOICEOBS	*[+son]‖[-voice]	*FRIC
☞ faɪndəm				*	*
paɪndəm	*!				
vaɪndəm		*!	*		*
baɪndəm	*!	*	*		

For this candidate, {IDENT(cont), IDENT(voice)} >> {*FRIC , *[+son]‖[-voice]}; *VOICEOBS could be unranked or replace IDENT(voice).

Tableau 11:

[+son]faɪndəm	IDENT(cont)	*[+son]‖[-voice]	*FRIC	IDENT(voice)	*VOICEOBS
faɪndəm		*!	*		
paɪndəm	*!				
☞ vaɪndəm			*	*	*
baɪndəm	*!			*	*

For this candidate, {IDENT(cont), *[+son]‖[-voice]} >> {*FRIC, IDENT(voice), *VOICEOBS}.

---

would make inaccurate predictions for my dataset. Therefore, although *[+son]‖[-voice] is perhaps less natural than one that would apply to all voiceless obstruents, it more accurately accounts for my data.

Tableau 12:

[+son]faɪndəm	*FRIC	IDENT(voice)	*VOICEOBS	*[+son]‖[-voice]	IDENT(cont)
faɪndəm	*!			*	
☞ paɪndəm					*
vaɪndəm	*!	*	*		
baɪndəm		*!	*		*

For this candidate, {*FRIC, IDENT(voice)} >> { *[+son]‖[-voice], IDENT(cont)}; *VOICEOBS could be unranked or replace IDENT(voice).

Empirical Generalization 5**:** Word-initial voiceless coronal fricatives occurring after sonorants vary freely with voiced fricatives, voiceless stops, and voiceless affricates.[66] This is a result of the interaction between faithfulness and markedness constraints related to frication, voicing, and stridency, which were introduced in the preceding subsections.

The following tableaus illustrate the interaction between IDENT(cont), *VOICEOBS, IDENT(voice), *FRIC, *[+son]‖[-voice], *[+strid], and IDENT(strid) for word-initial underlying voiceless coronal fricatives.

Tableau 13:

[+son]sɪtɪn	IDENT(cont)	IDENT(strid)	*VOICEOBS	IDENT(voice)	*[+strid]	*FRIC	*[+son]‖[-voice]
☞ sɪtɪn					*	*	*
cɪtɪn	*!	*					
tçɪtɪn	*!				*		
zɪtɪn			*!	*	*	*	
ɟɪtɪn	*!	*	*	*			
dʝɪtɪn	*!		*	*	*		

For this candidate, {IDENT(cont), *VOICEOBS } >> {*FRIC, *[+son]‖[-voice], *[+strid]}; IDENT(strid) unranked, IDENT(voice)could be unranked or replace *VOICEOBS.

---

[66] There is only one token of /z/ occurring after a sonorant across a word boundary. This is possibly an anomaly, or the absence of tokens with voiced affricates and stops could be an accidental gap in my data (see footnote 38).

Tableau 14:

[+son]sıtın	*[+strid]	*VOICEOBS	*FRIC	*[+son]‖[-voice]	IDENT(cont)	IDENT(voice)	IDENT(strid)
sıtın	*!		*	*			
☞ cıtın					*		*
tçıtın	*!				*		
zıtın	*!	*	*			*	
ɟıtın		*!			*	*	*
dʝıtın	*!	*			*	*	

For this candidate, {*[+strid], IDENT(voice)} >> { IDENT(cont), IDENT(strid)}; *FRIC and *[+son]‖[-voice] unranked, *VOICEOBS could be unranked or replace IDENT(voice).

Tableau 15:

[+son]sıtın	IDENT(cont)	IDENT(strid)	*[+son]‖[-voice]	*[+strid]	*FRIC	IDENT(voice)	*VOICEOBS
sıtın			*!	*	*		
cıtın	*!	*					
tçıtın	*!			*			
☞ zıtın				*	*	*	*
ɟıtın	*!	*				*	*
dʝıtın	*!			*		*	*

For this candidate, {IDENT(cont), *[+son]‖[-voice]} >> {*[+strid], *VOICEOBS, IDENT(voice), IDENT(strid), *FRIC}.

Tableau 16:

[+son]sıtın	IDENT(strid)	*FRIC	*VOICEOBS	*[+son]‖[-voice]	IDENT(voice)	*[+strid]	IDENT(cont)
sıtın		*!		*		*	
cıtın	*!						*
☞ tçıtın						*	*
zıtın		*!	*		*	*	
ɟıtın	*!		*		*		*
dʝıtın			*!		*	*	*

For this candidate, {IDENT(strid), *FRIC, *VOICEOBS } >> {IDENT(cont), *[+son]‖[-voice]}; *[+strid] unranked, IDENT(voice) could be unranked or replace *VOICEOBS.

3.1.4 Conditions 5-6

***Empirical Generalization 6*:** Word-initial /h/ is variably deleted. This is accomplished by an interaction of faithfulness constraints and markedness constraints against glottal segments.[67]

In addition to *VOICEOBS, *FRIC, IDENT(voice), and IDENT(cont), I use the following constraints:
MAX: Assign one violation mark for every input segment that does not have an output correspondent.
***GLOTTAL:** Assign one violation mark for every glottal segment.

The following tableaus illustrate the interactions between MAX, *VOICEOBS, IDENT(voice), *FRIC, *GLOTTAL, and IDENT(cont) for word-initial underlying /h/ after either a non-sonorant or a sonorant segment.

Tableau 17:[68]

hʌcbɪnd	MAX	*VOICEOBS	*FRIC	IDENT(voice)	*GLOTTAL	IDENT(cont)
☞ hʌcbɪnd			*		*	
ʔʌcbɪnd		*!		*	*	*
ʌcbɪnd	*!					

For this candidate, {MAX, IDENT(voice)} >> {*FRIC, *GLOTTAL}; IDENT(cont) unranked, *VOICEOBS could be unranked or replace IDENT(voice).

---

[67] The fact that other fricatives do not variably delete word-initially could easily be accounted for by positing that markedness constraints against the other places of articulation or perhaps against those specific fricatives are ranked below MAX. However, I do not directly address this issue in my analysis.
[68] Kriol 'husbind'; English 'husband.'

Tableau 18:

hʌcbɪnd	IDENT(cont)	*FRIC	*GLOTTAL	MAX	IDENT(voice)	*VOICEOBS
hʌcbɪnd		*!	*			
ʔʌcbɪnd	*!		*		*	*
☞ ʌcbɪnd				*		

For this candidate, {IDENT(cont), *FRIC} >> MAX; IDENT(voice) unranked, *GLOTTAL and could be unranked or replace IDENT(cont).

Tableau 19:

[+son]hʌcbɪnd	*VOICEOBS	MAX	IDENT(cont)	IDENT(voice)	*GLOTTAL	*FRIC	*[+son]||[-voice]
☞ hʌcbɪnd					*	*	*
ʔʌcbɪnd	*!		*	*	*		
ʌcbɪnd		*!					
ɦʌcbɪnd	*!			*	*	*	

For this candidate, {MAX, *VOICEOBS} >> {*[+son]||[-voice], *FRIC, *GLOTTAL}; IDENT(cont) unranked; IDENT(voice) could be unranked or replace *VOICEOBS.

Tableau 20:

[+son]hʌcbɪnd	*GLOTTAL	*VOICEOBS	*FRIC	IDENT(voice)	*[+son]||[-voice]	MAX	IDENT(cont)
hʌcbɪnd	*!		*		*		
ʔʌcbɪnd	*!	*		*			*
☞ ʌcbɪnd						*	
ɦʌcbɪnd	*!	*	*	*			

For this candidate, *GLOTTAL >> MAX; *[+son]||[-voice] and either IDENT(voice) or *VOICEOBS could be unranked or replace *GLOTTAL; IDENT(cont) and *FRIC could be unranked or replace *GLOTTAL.

3.1.5 Condition 7

***Empirical Generalization 7*:** Word-medial intervocalic underlying voiced stops vary freely with voiced fricatives. This is caused by a complex interaction between faithfulness constraints and phonological pressure to voice and fricate intervocalic obstruents.

In addition to *VOICEOBS, *FRIC, IDENT(voice), and IDENT(cont), I use the following constraints:

AGREE(voice): Assign one violation mark for sequence of vowel + obstruent + vowel that does not agree in voicing.

AGREE(cont): Assign one violation mark for sequence of vowel + obstruent + vowel that does not agree in continuancy.

The following tableaus illustrate the interaction between IDENT(voice), IDENT(cont), *FRIC, AGREE(cont), *VOICEOBS, and AGREE(voice) for word-medial intervocalic voiced stops.

Tableau 21:[69], [70]

dɪgəm	*FRIC	IDENT(voice)	IDENT(cont)	AGREE(voice)	*VOICEOBS	AGREE(cont)
☞ dɪgəm					**	*
dɪɣəm	*!		*		**	
dɪkəm		*!		*	*	*
dɪxəm	*!	*	*	*	*	
tɪgəm		*!			*	*
tɪɣəm	*!	*	*		*	
tɪkəm		*!*		*		*
tɪxəm	*!	**	*	*		

For this candidate, {IDENT(voice), *FRIC} >> {AGREE(cont), *VOICEOBS}; AGREE(voice) unranked; IDENT(cont) could be unranked or replace.

---

[69] Kriol 'digim'; English 'dig.'

[70] I do not present tableaus that illustrate every combination of word-initial segments of interest and medial intervocalic segments (such as a word with both an initial post-non-sonorant voiceless fricative and a medial intervocalic voiced stop, a word with an initial post-sonorant voiceless fricative and a medial intervocalic stop, etc.). Firstly, there are no tokens for some of the combinations. Secondly, I have shown the relevant constraint interactions for each segment in each position; this is sufficient to understand the predicted nature and amount of constraint overlap. The data that I input into Praat is coded for variation in both positions, so the GLA adjusts the constraints accordingly. The only reason why I used an item that has variation in both locations for this case (/dɪgəm/) is because there were no tokens with intervocalic voiced stops that had an initial segment that my analysis does not address.

Tableau 22:

dɪgəm	AGREE(cont)	AGREE(voice)	IDENT(voice)	*VOICEOBS	*FRIC	IDENT(cont)
dɪgəm	*!			**		
☞ dɪɣəm				**	*	*
dɪkəm	*!	*	*	*		
dɪxəm		*!	*	*	*	*
tɪgəm	*!		*	*		
tɪɣəm			*!	*	*	*
tɪkəm	*!	*	**			
tɪxəm		*!	**		*	*

For this candidate, {IDENT(voice), AGREE(cont)} >> {IDENT(cont), *VOICEOBS, *FRIC}; AGREE(voice) unranked.

Tableau 23:

dɪgəm	AGREE(voice)	*VOICEOBS	IDENT(cont)	*FRIC	IDENT(voice)	AGREE(cont)
dɪgəm		**!				*
dɪɣəm		**!	*	*		
dɪkəm	*!	*			*	*
dɪxəm	*!	*	*	*	*	
☞ tɪgəm		*			*	*
tɪɣəm		*	*!	*	*	
tɪkəm	*!				**	*
tɪxəm	*!		*	*	**	

For this candidate, AGREE(voice) >> {IDENT(cont), *VOICEOBS} >> {IDENT(voice), AGREE(cont)}; *FRIC unranked.

Tableau 24:

dɪgəm	AGREE(voice)	*VOICEOBS	AGREE(cont)	IDENT(voice)	*FRIC	IDENT(cont)
dɪgəm		**!	*			
dɪɣəm		**!			*	*
dɪkəm	*!	*	*	*		
dɪxəm	*!	*		*	*	*
tɪgəm		*	*!	*		
☞ tɪɣəm		*		*	*	*
tɪkəm	*!		*	**		
tɪxəm	*!			**	*	*

For this candidate, AGREE(voice) >> {AGREE(cont), *VOICEOBS} >> **{**IDENT(voice), IDENT(cont), *FRIC}.

3.1.6 Condition 8

Empirical Generalization 8**:** Word-medial intervocalic underlying voiceless stops vary freely with voiced stops, voiceless fricatives, and voiced fricatives. This is caused by a complex interaction between faithfulness constraints and phonological pressure to voice and fricate intervocalic obstruents.

The following tableaus illustrate the interaction between IDENT(voice), IDENT(cont)**,** *FRIC, AGREE(cont), *VOICEOBS, and AGREE(voice) for word-medial intervocalic voiceless stops.

Tableau 25:[71]

lʊkɪn	IDENT(cont)	*VOICEOBS	*FRIC	IDENT(voice)	AGREE(voice)	AGREE(cont)
☞ lʊkɪn					*	*
lʊgɪn		*!		*		*
lʊxɪn	*!		*		*	
lʊɣɪn	*!	*	*	*		

For this candidate, {IDENT(cont), *VOICEOBS} >> {AGREE(cont), AGREE(voice)}; *FRIC and IDENT(voice) could replace IDENT(cont) and *VOICEOBS or be unranked.

---

[71] Kriol 'loogin'; English 'looking.'

Tableau 26:

lʊkɪn	AGREE(voice)	AGREE(cont)	IDENT(voice)	*VOICEOBS	IDENT(cont)	*FRIC
lʊkɪn	*!	*				
lʊgɪn		*!	*	*		
lʊxɪn	*!				*	*
☞ lʊɣɪn			*	*	*	*

For this candidate, {AGREE(cont), AGREE(voice)} >> {IDENT(cont), *FRIC, IDENT(voice), *VOICEOBS}.

Tableau 27:

lʊkɪn	AGREE(cont)	IDENT(voice)	*VOICEOBS	AGREE(voice)	*FRIC	IDENT(cont)
lʊkɪn	*!			*		
lʊgɪn	*!	*	*			
☞ lʊxɪn				*	*	*
lʊɣɪn		*!	*		*	*

For this candidate, {IDENT(voice), AGREE(cont)} >> {AGREE(voice), IDENT(cont), *FRIC}; *VOICEOBS could replace IDENT(voice) or be unranked.

Tableau 28:

lʊkɪn	*FRIC	AGREE(voice)	IDENT(cont)	*VOICEOBS	AGREE(cont)	IDENT(voice)
lʊkɪn		*!			*	
☞ lʊgɪn				*	*	*
lʊxɪn	*!	*	*			
lʊɣɪn	*!		*	*		*

For this candidate, {*FRIC, AGREE(voice)} >> {AGREE(cont), IDENT(voice), *VOICEOBS}; IDENT(cont) could replace *FRIC or be unranked.

3.1.7 Condition 9

***Empirical Generalization 9*:** Word-medial intervocalic underlying voiceless fricatives vary freely with voiced stops, voiceless stops, and voiced fricatives.[72] This is caused by a complex interaction between faithfulness constraints and phonological pressure to voice and fricate intervocalic obstruents.

The following tableaus show the interaction between IDENT(cont), AGREE(cont), IDENT(voice), *VOICEOBS, *FRIC, and AGREE(voice) for word-medial intervocalic underlyingly voiceless fricatives.

Tableau 29:[73]

nʌθɪŋ	IDENT(voice)	IDENT(cont)	*FRIC	*VOICEOBS	AGREE(voice)	AGREE(cont)
☞ nʌθɪŋ			*		*	
nʌðɪŋ	*!		*	*		
nʌtɪŋ		*!			*	*
nʌdɪŋ	*!	*		*		*

For this candidate, {IDENT(cont), IDENT(voice)} >> {*FRIC, AGREE(voice)}; *VOICEOBS and AGREE(cont) could replace IDENT(cont) and IDENT(voice) or be unranked.

---

[72] There is only one token where an intervocalic fricative surfaces as a voiceless stop (/f/ → /p/), and there are no tokens where an intervocalic /f/ surfaces as the voiceless fricative. There are also no tokens of the coronal fricative /s/ varying with the voiced or voiceless affricate intervocalically. These could be anomalies, accidental gaps in my token data, or accurately representative of variation in Kriol. I do not have the means to further address these gaps and others that have been mentioned previously. However, there is at least one token of *some* intervocalic fricative surfacing as each of the voiced and voiceless stop and fricative variants.

[73] Kriol 'nuthin'; English 'nothing.'

Tableau 30:

nʌθɪŋ	AGREE(voice)	AGREE(cont)	IDENT(cont)	IDENT(voice)	*FRIC	*VOICEOBS
nʌθɪŋ	*!				*	
☞ nʌðɪŋ				*	*	*
nʌtɪŋ	*!	*	*			
nʌdɪŋ		*!	*	*		*

For this candidate, {AGREE(cont), AGREE(voice)} >> {IDENT(voice), *VOICEOBS}; *FRIC unranked; IDENT(cont) could replace AGREE(cont) or be unranked.

Tableau 31:

nʌθɪŋ	*FRIC	IDENT(voice)	*VOICEOBS	AGREE(voice)	IDENT(cont)	AGREE(cont)
nʌθɪŋ	*!			*		
nʌðɪŋ	*!	*	*			
☞ nʌtɪŋ				*	*	*
nʌdɪŋ		*!	*		*	*

For this candidate, {*FRIC, *VOICEOBS} >> {AGREE(voice), AGREE(cont), IDENT(cont)}; IDENT(voice) could replace *VOICEOBS or be unranked.

Tableau 32:

nʌθɪŋ	AGREE(voice)	*FRIC	IDENT(voice)	AGREE(cont)	IDENT(cont)	*VOICEOBS
nʌθɪŋ	*!	*				
nʌðɪŋ		*!	*			*
nʌtɪŋ	*!			*	*	
☞ nʌdɪŋ			*	*	*	*

For this candidate, {*FRIC, AGREE(voice)} >> {IDENT(voice), *VOICEOBS, AGREE(cont), IDENT(cont)}.

3.1.8 Overall Observations

The sections above present a complicated picture of the Kriol grammar. Judging from the tableaus, most of the constraints are ranked close together, resulting in a significant amount of overlap and thus accounting for the observed variants. Groups of constraints that should be ranked close together to allow for overlap include continuancy constraints (*FRIC, IDENT(cont), AGREE(cont)), stridency constraints (*[+strid], IDENT(strid)), voicing constraints (IDENT(voice), AGREE(voice), *VOICEOBS, *[+son]|[-voice]), and constraints dealing with deletion or retention of /h/ (MAX, *GLOTTAL). Many of these processes interact, however, so I expect to see substantial overlap across these clumps of constraints as well.

3.2 Results from GLA

In this section of my analysis, I ran Trials 1-3 through Praat's OT Learning Program, which is a computational implementation of the Gradual Learning Algorithm for Stochastic Optimality Theory. I then computed a number of statistical tests on the final stochastic grammars' predicted frequencies for each output variant from the given input forms to see how they compared to the frequencies that I recorded from the corpus data. I calculated the individual standard deviations for the observed frequencies and those predicted by the learned grammars to see how similar they were. I then calculated the Pearson product-moment correlation coefficient of the two sets of frequencies to determine how well they correlated across each pair. Finally, I calculated the difference in percentage for each pair and averaged the resulting numbers to see how much the two sets of frequencies differed from each other on average. In addition to the main trials, I ran a subset of the data in each trial to more closely examine the constraint interactions and frequency predictions.

Decreasing the plasticity and increasing the number of times each input is run through Praat's Optimality Learning Program could possibly affect the results. In order to test the effect of these metrics, I ran each trial two additional times with altered values for plasticity and number of runs per input. The algorithm in the original trials adjusted constraints by a plasticity of .1 and learned each input 100,000 times.[74] In the first set of trials with altered metrics, I changed the plasticity to .001 and ran each input form through the grammar 1,000,000 times; in the second set of trials with altered metrics, the grammar learned the variation initially with a plasticity of .1 and then with a plasticity of .001, and each input was presented to the model 100,000 times. The first set of trials with changed metrics performed worse than the original models of the trials, while the second set of trials predicted frequencies for variants that were negligibly different than the original trials. Given the insignificant difference of these outcomes,

---

[74] These are Praat's default settings. See Section 2.2.2.

I only present the results for the original trials that use the default metrics in Praat's GLA program.

In the following sections, I begin by examining phonological patterns at the level of the individual lexical item (Trial 1), and then I collapse that data into individual phonemes (Trial 2) and overall natural classes (Trial 3). The following sections also provide a justification for the increasing abstractness across trials. I show that in Trial 1, the significant differences in percentages of observed variants of items in the same condition are generally due to small amounts of raw token data with insignificantly different distributions of variants rather than lexically-specific variation. Although the models are intended to capture general phonological patterns that apply at the level of the individual lexical items, I show that there are a handful of lexical items that do exhibit statistically different distributions of variants and thus suggest lexical exceptions or idiosyncratic variation. Removing these specific items and collapsing the remaining data into individual phonemes and overall natural classes decreases the statistical noise and thus allows the grammar to account for general phonological processes. I also show that in the most general model, Trial 3, the inclusion of a few additional constraints improves the accuracy of the output grammar's predictions, indicating that there are more phonological processes for which the grammar must account than the original analysis includes.

3.2.1 Trial 1: Lexical Items

In Trial 1, the learning algorithm generated a grammar in Stochastic OT based on a list of 87 underlying representations for lexical items and 424 tokens. I examined the frequencies for each variant that the output grammar predicted and compared those with the empirical frequencies in the corpus data. In order to determine how accurate the output constraint ranking's predicted variants were in comparison with the observed frequencies of variants, I calculated a series of statistical analyses that are presented in the following table:

Calculation	**Value**
Standard Deviation of Actual Frequencies	34.60
Standard Deviation of Praat's Frequencies	28.39
Pearson Coefficient	.72
Average Difference (in % Points)	16.41

The results show that the difference between the standard deviations of each set of frequencies was fairly large and the individual pairs of frequencies were only somewhat correlated. The following table presents a few of the 87 underlying representations with the observed and predicted frequencies for their variants:

Condition	Underlying Representation	Variant	Observed Number of Tokens	Observed Frequency	Praat Output Frequency	Difference
1	/bɪɡwʌn/	[bɪɡwʌn]	3	75%	83%	8
		[pɪɡwʌn]	1	25%	17%	8
1	/ɡʌntçɹi/	[ɡʌntçɹi]	1	20%	83%	63
		[kʌntçɹi]	4	80%	17%	63
1	/daʊɡ/	[daʊɡ]	5	50%	83%	33
		[taʊɡ]	5	50%	17%	33
2	/kʌm/	[kʌm]	17	100%	100%	0
		[ɡʌm]	0	0%	0%	0
3	/faɪndəm/	[faɪndəm]	4	57%	81%	24
		[paɪndəm]	3	43%	19%	24
3	/sɪtɪn/	[sɪtɪn]	4	100%	74%	26
		[cɪtɪn]	0	0%	15%	15
		[tçɪtɪn]	0	0%	11%	11
3	/θɪŋkɪnabat/	[θɪŋkɪnabat]	2	50%	93%	43
		[tɪŋkɪnabat]	2	50%	7%	43
4	/[+ son]faɪndəm/	[faɪndəm]	18	75%	76%	1
		[paɪndəm]	2	8%	20%	12
		[vaɪndəm]	4	17%	4%	13
		[baɪndəm]	0	0%	0%	0
4	/[+ son]sɪtɪn/	[sɪtɪn]	2	40%	67%	27
		[cɪtɪn]	2	40%	15%	25
		[tçɪtɪn]	1	20%	14%	6
		[zɪtɪn]	0	0%	4%	4
		[ʝɪtɪn]	0	0%	0%	0

		[dʝɪtɪn]	0	0%	0%	0
4	/[＋son]θɪŋkɪnabat/	[θɪŋkɪnabat]	0	0%	79%	79
		[tɪŋkɪnabat]	3	100%	16%	84
		[ðɪŋkɪnabat]	0	0%	4%	4
		[dɪŋkɪnabat]	0	0%	1%	1
5	/hʌcbɪnd/	[hʌcbɪnd]	1	17%	41%	24
		[ʌcbɪnd]	5	83%	52%	31
		[ʔʌcbɪnd]	0	0%	7%	7
6	/[＋son]hʌcbɪnd/	[hʌcbɪnd]	1	100%	37%	63
		[ʌcbɪnd]	0	0%	54%	54
		[ʔʌcbɪnd]	0	0%	7%	7
		[ɦʌcbɪnd]	0	0%	2%	2
7	/dɪgəm/	[dɪgəm]	4	58%	49%	9
		[dɪxəm]	0	0%	0%	0
		[dɪɣəm]	1	14%	34%	20
		[tɪkəm]	0	0%	2%	2
		[tɪxəm]	0	0%	1%	1
		[tɪgəm]	1	14%	8%	6
		[tɪɣəm]	1	14%	6%	8
		[dɪkəm]	0	0%	0%	0
8	/oʊkə/	[oʊkə]	2	29%	21%	8
		[oʊxə]	0	0%	9%	9
		[oʊgə]	2	29%	50%	21
		[oʊɣə]	3	42%	20%	22
8	/pʌpədaʊg/	[pʌpədaʊg]	1	100%	21%	79
		[pʌfədaʊg]	0	0%	9%	9
		[pʌbədaʊg]	0	0%	50%	50
		[pʌvədaʊg]	0	0%	20%	20
9	/nɛfə/	[nɛfə]	0	0%	26%	26
		[nɛpə]	0	0%	4%	4
		[nɛvə]	11	79%	62%	17
		[nɛbə]	3	21%	8%	13
9	/lɪsɪn/	[lɪsɪn]	1	33%	25%	8
		[lɪcɪn]	0	0%	3%	3

		[lɪtçɪn]	0	0%	1%	1
		[lɪzɪn]	0	0%	60%	60
		[lɪɟɪn]	2	67%	7%	60
		[lɪdʝɪn]	0	0%	4%	4

As the table above illustrates, the accuracy of the output grammar's predictions varied greatly across individual lexical items. The output grammar predicted that word-initial underlyingly voiced stops would surface as their voiced variants 83% of the time and as their voiceless variants 17% of the time; this prediction was more accurate for some lexical items, such as /bɪɡwʌn/, than others, like /ɡʌntçɹi/.[75] The output grammar always predicted the variants of words with initial underlyingly voiceless stops entirely accurately—as did the other two trials—because the underlying voiceless stop is undominated in the word-initial context and will thus always be the most optimal candidate. The output grammar for Trial 1 did not predict the variants of words with initial underlying fricatives very well, although the grammar was generally more accurate at predicting the variants of words with initial fricatives after non-sonorants than after sonorants (e.g. /[+son]faɪndəm/ vs. /faɪndəm/). Additionally, the output grammar predicted the frequency of variants of words with intervocalic underlyingly voiced stops fairly well—such as /dɪɡəm/—but it had mixed accuracy for lexical items with intervocalic underlyingly voiceless stops and fricatives, such as /oʊkə/, /pʌpədaʊɡ/, /nɛfə/, and /lɪsɪn/.

Taking a closer look at the data, a few patterns emerge. The output grammar poorly predicted frequencies for a few items with underlyingly voiced initial stops, in particular /bʊl/ ('pull'), /dʌɾəɡi/ ('turkey'), and /ɡʌntçɹi/ ('country), and /ɡʌntçɹimən/ ('countryman');[76] the

---

[75] The output grammar's percentages of predicted variants are not simply the average of all of the relevant inputs' frequencies of variants. For example, the percentage of voiced and voiceless variants predicted by the output grammar for words with underlyingly voiced word-initial stops are not averages of all of the empirically observed frequencies of variants for each word beginning with a voiced stop. This is because the constraints that determine these specific output variants—*VOICEOBS, IDENT(voice), etc.—are also affected by any non-word-initial voiced obstruents and any segmental voicing changes in variants of other words in the input. Therefore, the resulting percentages of variants in the output grammar are the result of complicated interactions between the constraints and the input forms.

[76] The English counterparts of these words are voiceless, but they appear to be voiced in Kriol. If I had determined that the underlying forms had word-initial *voiceless* stops, the presence of voiced variants for these underlyingly voiceless segments would require a different analysis and set of constraints from those

grammar predicted that there would be more voiced variants than I observed in the corpus data. The observed frequencies of variants of these items differed significantly from the observed frequencies of variants of other lexical items with word-initial voiced stops. In order to determine whether the grammar's inaccurate predictions for these forms were due to sparse token data or idiosyncratic variation within the specific lexical items, I performed a series of Fisher's exact tests. In these tests, I compared the number of raw tokens for variants of these items—as well as a few other items—with the number of tokens for variants of a lexical item with an initial underlyingly voiced stop and a large number of tokens, /gaɾə/. The table below presents a comparison of the tokens and observed frequencies for a handful of forms with an initial underlyingly voiced stop:

Lexical Item	Variants	Tokens	Frequency	*p*-value from Fisher's exact test comparing tokens with /gaɾə/[77]	*p*-value from Fisher's exact test comparing frequencies with /gaɾə/
/gaɾə/	[gaɾə]	9	69%	NA	NA
	[kaɾə]	4	31%		
/gʌntçɹi/	[gʌntçɹi]	1	20%	.1176	.0001
	[kʌntçɹi]	4	80%		
/dʌɾəgi/	[dʌɾəgi]	1	11%	.0115	.0001
	[tʌɾəgi]	8	89%		
/bʊl/	[bʊl]	3	30%	.0995	.0001
	[pʊl]	7	70%		
/daʊg/	[daʊg]	5	50%	.4173	.0093
	[taʊg]	5	50%		
/bɪgwʌn/	[bɪgwʌn]	3	75%	1.0000	.4312
	[pɪgwʌn]	1	25%		
/gʌntçɹimən/	[gʌntçɹimən]	1	25%	.2500	.0001
	[kʌntçɹimən]	3	75%		

---

presented in this paper, so I did not consider this explanation. Perhaps the frequency of the devoiced variants is due to analogy with the English forms.

[77] Where $p < .05$ is significant.

	[kat]	3	75%		

The series of Fisher's exact tests indicate that in comparison with the lexical item with the most tokens, /gaɾə/, only /dʌɾəgi/ had a significantly different distribution of variants for both the raw tokens *and* the resulting percentages.[78] These results suggest that apart from /dʌɾəgi/—which does not behave similarly to the rest of the items its specific condition—the drastically different observed percentages of variants for items in this group are most likely not due to idiosyncratic patterns of variation for individual lexical items. It is difficult to determine whether the difference in percentages of variants is due to a small amount of token data from some items or missed phonological generalizations, such as ones that might apply to a specific class of words differently than another. It is also possible that with more tokens, some of these items might reveal themselves to have lexically-specific variation similar to /dʌɾəgi/. However, with no direct evidence for lexically-specific variation from the observed data in any case except for /dʌɾəgi/, it is logical to omit this word as an exception to the general phonological trends and collapse the remaining data from Trial 1 into tokens for the more general categories of individual phonemes in Trial 2.

I also conducted a series of Fisher's exact tests comparing the token amounts for lexical items with intervocalic underlying fricatives that had significantly different percentages of variants.[79] Out of these items, the only form that differed significantly from the others in its distribution of raw tokens was /oʊfə/. On the other hand, for /lɪsɪn/, which had significantly different percentages of variants in comparison with other items with intervocalic fricatives, the distribution of raw tokens was not significantly different from these other items. These two cases suggest that the significant difference in percentage of variants of /oʊfə/—which had ten

---

[78] Only one word, /bɪgwʌn/, did not differ significantly from /gaɾə/ in terms of the resulting percentages of each variant. However, although the distributions were not significantly different, the individual percentages (69% vs. 75% and 31% vs. 25%) are still different enough so as to require slightly different amounts of overlap for constraints.

[79] Although I do not present the table for these Fisher's tests, they are of the same format as the ones in the table on p.53, except that the items are all compared to each other as opposed to each item being compared solely to one representative form.

total tokens whose distribution was significantly different from other intervocalic fricatives—is due to its status as a lexical exception, while the significant difference in percentage of variants of /lɪsɪn/—which had three total tokens whose distribution was not significantly different from other intervocalic fricatives —is likely not due to lexically-specific variation.[80] These results lend support to eliminating /oʊfə/ and collapsing the rest of the data, including the tokens of /lɪsɪn/, into individual phoneme inputs in Trial 2.

Two final items which merit discussion are /pʌpədaʊg/ ('puppy') and /[+son]hapi/ ('happy'). For both of these items, the output constraints predicted a significantly different percentage of variants in comparison with the rest of the lexical items in the intervocalic voiceless stop condition.[81] The output grammar predicted more variants with the voiced intervocalic stop for each of these items than the corpus data exhibited. However, although this resulted in significant differences in *percentages* of variants, the difference in the distribution of raw tokens for each of these items was not significantly different from other items in their condition.[82] This indicates that the difference in percentages of variants is not due to lexically-specific variation. Instead, it is likely due to either small amounts of tokens—each item had one token each—or a failure of the grammar to capture some phonological generalization. However, it is difficult to tease these factors apart at this level of detail in the grammar, where there is a large amount of overall noise resulting from uneven amounts of tokens for individual inputs. Therefore, the results from these Fisher's exact tests support collapsing the data into phonemes in Trial 2. All in all, the output constraint rankings from Trial 1 are presented below:

---

[80] See Section 2.1.3 for a detailed discussion of the difficulty of determining the underlying forms of intervocalic segments. I ran alternate models for all three trials that listed the underlying forms for 'never,' 'over,' and 'river' as /nɛbə/, /oʊbə/, and /ɹɪbə/ instead of /nɛfə/, /oʊfə/, and /ɹɪfə/, respectively. I only changed these three items because they were the most ambiguous in their underlying forms, whereas the other intervocalic fricatives were plausibly fricatives underlyingly. However, the overall differences between the predicted frequencies from the two sets of trials' output grammars were not significant. Therefore, it seems that changing the underlying representations of these forms does not improve the accuracy of the grammar.

[81] See below for a subset of the data from Trial 1 run without these and a few other lexical items.

[82] According to Fisher's exact tests like those presented in the table on p.53.

Constraint	**Ranking Value**
AGREE(voice)	102.074
IDENT(voice)	100.405
IDENT(cont)	100.243
*GLOTTAL	100.002
MAX	99.974
IDENT(strid)	99.905
AGREE(cont)	99.868
*[+strid]	98.395
*VOICEOBS	97.687
*FRIC	97.464
*[+son]‖[-voice]	96.347

This table shows that for the final stochastic grammar, violations of AGREE(voice), IDENT(voice), and IDENT(cont) are generally dis-preferred, while fricatives, stridents, voiced obstruents, and initial fricatives after words ending in sonorants are generally allowed. Additionally, glottal segments are often avoided. The ranking also shows clumps of highly interactive constraints whose constraint rankings are close together; because their standard deviation is 2.0, this results in a substantial amount of overlap. Although the results from Trial 1 are interesting and provide some useful information, it seems that the data is potentially confounded both by lexical items with few tokens and a small number of lexical exceptions that behave categorically differently than the other items within their condition. In order to help determine the faults of Trial 1, I ran a subset of the data using only lexical items with 10 or more tokens. The following table presents the observed and output frequencies of these items:

Condition	Underlying Representation	Variant	Observed Number of Tokens	Observed Frequency	Praat Output Frequency	Difference
1	/bʊl/	[bʊl]	3	30%	58%	28
		[pʊl]	7	70%	42%	28
1	/gaɾə/	[gaɾə]	9	69%	58%	11
		[kaɾə]	4	31%	42%	11
1	/daʊg/	[daʊg]	5	50%	59%	9
		[taʊg]	5	50%	41%	9

2	/kʌm/	[kʌm]	17	100%	100%	0
		[gʌm]	0	0%	0%	0
4	/[+ son]faɪndəm/	[faɪndəm]	18	75%	64%	11
		[paɪndəm]	2	8%	31%	23
		[vaɪndəm]	4	17%	5%	12
		[baɪndəm]	0	0%	0%	0
8	/lʊkɪn/	[lʊgɪn]	0	58%	48%	48
		[lʊxɪn]	0	0%	5%	5
		[lʊɣɪn]	8	67%	40%	27
		[lʊkɪn]	4	33%	7%	26
9	/oʊfə/	[oʊfə]	0	0%	11%	11
		[oʊpə]	0	0%	2%	2
		[oʊvə]	10	100%	76%	24
		[oʊbə]	0	0%	11%	11
9	/ɹɪfə/	[ɹɪfə]	0	0%	11%	11
		[ɹɪpə]	0	0%	2%	2
		[ɹɪvə]	6	54%	76%	22
		[ɹɪbə]	5	46%	11%	35
9	/nɛfə/	[nɛfə]	0	0%	11%	11
		[nɛpə]	0	0%	2%	2
		[nɛvə]	11	79%	76%	3
		[nɛbə]	3	21%	11%	10
9	/nʌθɪŋ/	[nʌθɪŋ]	1	9%	11%	2
		[nʌtɪŋ]	0	0%	2%	2
		[nʌðɪŋ]	7	58%	76%	18
		[nʌdɪŋ]	4	33%	11%	22

Based on this data, I calculated the same statistical tests that I ran on the trial with the complete dataset, which are presented below:

Calculation	**Value**
Standard Deviation of Actual Frequencies	32.66
Standard Deviation of Praat's Frequencies	29.68
Pearson Coefficient	.83
Average Difference (in % Points)	13.62

These two tables representing the subset of Trial 1 indicate that although increasing the tokens for the input items decreases noise and improves overall accuracy, the output grammar was still not able to accurately predict the frequencies of certain variants for some words that have a large number of tokens. However, the lack of significant differences in distribution of raw tokens[83]—apart from /oʊfə/—indicates that the resulting differences in percentages of variants within the same condition are not due to lexically-specific variation. This observation, which suggests that the differing percentages of variants are due to noise based on sparse data, provides further support for collapsing data in Trial 2. Perhaps more token data would decrease the differences in observed frequencies of variants for different lexical items within the same condition in Trial 1, or perhaps additional constraints would better capture phonological patterns. However, at this level of specificity, it is difficult to determine which of these factors is at play. In the more abstract trials to follow, it is easier to determine if there are grammatical generalizations that the constraint inventory is not taking into account.

The general lack of significant differences between raw token distributions among individual lexical items supports the collapse of the data in Trial 2 and Trial 3, although two individual lexical items—/oʊfə/ and /dʌɾəgi/—have been removed from these trials because they seem to represent lexical exceptions that differ slightly from the other words in their respective conditions. In Trial 2, these potentially idiosyncratic items are removed, while the rest of the data from the individual lexical items is collapsed into tokens for individual phonemes in word-initial and intervocalic position.

---

[83] Significance calculated with Fisher's exact tests.

3.2.2 Trial 2: Individual Phonemes

In Trial 2, the algorithm generated a grammar in Stochastic OT based on a list of phonemes in word-initial and intervocalic positions totaling in 21 underlying representations and 481 tokens. I examined the frequencies for each variant that the output grammar predicted and compared those with the empirically calculated frequencies from the corpus data. The statistics for Trial 2 are presented in the following table:

Calculation	**Value**
Standard Deviation of Actual Frequencies	31.42
Standard Deviation of Praat's Frequencies	27.57
Pearson Coefficient	.76
Average Difference (in % Points)	13.83

The results show that the difference between the standard deviations of each set of frequencies was smaller than in Trial 1, but still fairly large. The Pearson coefficient indicates that the individual pairs of frequencies were negligibly less closely correlated than in Trial 1, and the overall average difference in percentage points between the observed and predicted frequency of each variant decreased substantially from the values in Trial 1. The following table lists the observed and predicted frequencies for each of the phonemes in word-initial and intervocalic position:

Condition	Underlying Representation	Variant	Observed Number of Tokens	Observed Frequency	Praat Output Frequency	Difference
1	/#b/	[b]	42	72%	60%	12
		[p]	16	28%	40%	12
1	/#g/	[g]	31	67%	60%	7
		[k]	15	33%	40%	7
1	/#d/	[d]	10	59%	60%	1
		[t]	7	41%	40%	1
2	/#k/	[k]	35	100%	100%	0
		[g]	0	0%	0%	0

2	/#t/	[t]	29	100%	100%	0
		[d]	0	0%	0%	0
2	/#p/	[p]	6	100%	100%	0
		[b]	0	0	0%	0
3	/#f/	[f]	23	85%	56%	29
		[p]	4	15%	44%	29
3	/#s/	[s]	9	64%	35%	29
		[c]	1	7%	26%	19
		[tç]	4	29%	39%	10
3	/#θ/	[θ]	2	50%	42%	8
		[t]	2	50%	58%	8
4	/[+son]|f/	[f]	33	75%	40%	35
		[p]	4	9%	49%	40
		[v]	6	14%	11%	3
		[b]	0	0%	0%	0
4	/[+son]|s/	[s]	16	67%	23%	44
		[c]	4	17%	27%	10
		[z]	1	4%	6%	2
		[ɟ]	0	0%	0%	0
		[tç]	3	12%	44%	32
		[dʝ]	0	0%	0%	0
4	/[+son]|θ/	[θ]	0	0%	26%	26
		[t]	3	100%	67%	33
		[ð]	0	0%	7%	7
		[d]	0	0%	0%	0
5	/#h/	[h]	3	33%	21%	12
		[ʔ]	0	0%	6%	6
		Ø	6	67%	73%	6
6	/[+son]|h/	[h]	4	44%	15%	29
		[ʔ]	0	0%	7%	7
		Ø	5	56%	75%	19
		[ɦ]	0	0%	3%	3
7	/VbV/	[b]	9	69%	74%	5
		[p]	0	0%	4%	4

		[f]	0	0%	1%	1
		[v]	4	31%	21%	10
7	/VgV/	[g]	11	59%	67%	8
		[k]	0	0%	3%	3
		[x]	0	0%	2%	2
		[ɣ]	8	41%	28%	13
8	/VpV/	[b]	0	0%	68%	68
		[p]	4	44%	11%	33
		[f]	0	0%	3%	3
		[v]	5	56%	18%	38
8	/VkV/	[g]	8	16%	60%	44
		[k]	17	34%	10%	24
		[x]	4	8%	4%	4
		[ɣ]	21	42%	26%	16
9	/VfV/	[b]	11	30%	30%	0
		[p]	1	3%	5%	2
		[f]	0	0%	9%	9
		[v]	25	67%	56%	11
9	/VsV/	[ɟ]	2	67%	13%	54
		[c]	0	0%	2%	2
		[s]	1	33%	7%	26
		[z]	0	0%	38%	38
		[tç]	0	0%	2%	2
		[dʝ]	0	0%	38%	38
9	/VθV/	[d]	8	31%	33%	2
		[t]	0	0%	5%	5
		[θ]	1	4%	9%	5
		[ð]	17	65%	53%	12

As the table indicates, the final grammar for Trial 2 predicted the frequency of variants perfectly for initial voiceless stops. Within voiced stops, the grammar was extremely accurate in predicting frequency of variants for initial /#d/, less so for initial /#g/, and even worse for initial /#b/. Out of the set of word-initial fricatives after sonorants, the grammar's constraint ranking predicted the variants of /[+son]|h/ most accurately, followed by /[+son]|θ/, /[+son]|f/, and /[+son]|s/. For word-initial fricatives, the grammar's predictions were most accurate for /#θ/, /#h/, /#s/, and finally /#f/. For intervocalic segments, the output grammar was fairly accurate at predicting the variants of /VbV/, /VɡV/, /VfV/, and /VθV/, while it was comparatively less accurate at predicting the variants of /VkV/, /VpV/, and /VsV/. Generally speaking, the grammar's output frequencies for variants of word-initial fricatives were more accurate after sonorants than non-sonorants, and its frequencies for variants of intervocalic segments were more accurate for voiced stops than voiceless stops or fricatives. These tendencies are possibly due to the grammar's failure to capture some generalizations. However, they could also be due to a complex interaction between the existing constraints and uneven amounts of tokens for certain phonemes within that category; it is difficult to tease apart these factors. Below, I discuss a number of phonemes with small amounts of tokens that could have skewed the grammar's learning. I argue that collapsing the data into natural classes for Trial 3 is the most comprehensive way to narrow down the potential factors contributing to the grammar's failure to perform accurately in certain areas.

At a closer level of analysis, the frequencies from the final stochastic grammar for Trial 2 show a few trends. The two most inaccurate frequencies that arose from the final grammar's constraint ranking dealt with intervocalic segments; for intervocalic /VsV/, the grammar predicted fewer instances of /ɟ/, and for intervocalic /p/ the grammar predicted more occurrences of /b/. However, it is probable that the source of the significant difference in percentages of variants for intervocalic /VsV/ and /VpV/ in comparison with other intervocalic segments in their conditions (intervocalic fricatives and intervocalic voiceless stops, respectively) is not due to lexically-specific variation, but rather small token data. The table below shows a series of Fisher's exact tests comparing intervocalic fricatives to /VfV/ and intervocalic stops to /VkV/, the inputs in with the most tokens in their respective conditions:

Lexical Item	Variants	Tokens	Frequency	Average *p*-value from Fisher's exact test comparing tokens with /VfV/ or /VkV/[84]	Average *p*-value from Fisher's exact test comparing frequencies with /VfV/ or /VkV/
/VfV/	[b]	11	30%	NA	NA
	[p]	1	3%		
	[f]	0	0%		
	[v]	25	67%		
/VθV/	[d]	8	31%	.8114	.2713
	[t]	0	0%		
	[θ]	1	4%		
	[ð]	17	65%		
/VsV/	[ɟ]	2	67%	.4506	.0068[85]
	[c]	0	0%		
	[s]	1	33%		
	[z]	0	0%		
	[tç]	0	0%		
	[dʝ]	0	0%		
/VkV/	[ɡ]	8	16%	NA	NA
	[k]	17	34%		
	[x]	4	8%		
	[ɣ]	21	42%		
/VpV/	[b]	0	0%	.7721	.2008
	[p]	4	44%		
	[f]	0	0%		
	[v]	5	56%		

---

[84] Where $p < .05$ is significant.

[85] In all of these averages, I did not include tests where both of the values for variants of the same input were 0; this always resulted in a *p*-value of 1.0, which seems to be a default lack of significance that does not accurately reflect the data. If I included those values in the averages, the *p*-values for /VsV/ would be .5606 and .1723, and the *p*-values for /VpV/ would be .8101 and .3340 (from left to right across the columns).

This table shows that although the percentages of variants of /VsV/ are significantly different from the percentages of variants of /VfV/ and the percentages of variants of /VpV/ are very different from the percentages of variants of /VkV/,[86] the distributions of raw tokens are not significantly different in comparison with other phonemes in their conditions. Therefore, the significant difference in percentages is seemingly not due to phoneme-specific variation. Rather, it is either due to a small amount of tokens for /VsV/ and /VpV/ or some phonological patterns that the current constraints do not take into account. With a mere three tokens for /VsV/ and nine tokens for /VpV/, it is difficult to determine which of these factors is mainly at play here; it is quite possible that there are voiceless stop and voiced fricative variants of /VsV/ and voiced stop and voiceless fricative variants of /VpV/ that occur in the language but were simply not attested in the corpus data. It is possible that with more tokens, the observed frequencies of variants of these two inputs would have more closely resembled the frequencies of variants for other forms in their conditions. In any case, the overall lack of significant difference in distributions of raw tokens for the intervocalic phonemes supports the conflation of data into natural classes in Trial 3.

Another interesting point of discussion that emerged from Trial 2 was the pattern of accuracy within the grammar's predictions of the frequency of variants of different word-initial fricatives. The fricatives whose variants the grammar was most accurate at predicting in word-initial position were /#θ/ and /[+son]|θ/. The following chart presents a comparison between /#θ/ and /#f/ and /[+son]|θ/ and /[+son]|f/:

---

[86] Although the percentages of variants of /VpV/ are not *significantly* different from the percentages of variants of /VkV/, they are still different enough so as to require slightly different amounts of overlap of the relevant constraints. Therefore, the final grammar of Trial 2 cannot accurately account for the percentages of both forms' variants.

Lexical Item	Variants	Tokens	Frequency	Average p-value from Fisher's exact test comparing tokens with /VfV/ or /VkV/[87]	Average p-value from Fisher's exact test comparing frequencies with /VfV/ or /VkV/
/#f/	[f]	23	85%	NA	NA
	[p]	4	15%		
/#θ/	[θ]	2	50%	.1594	.0001
	[t]	2	50%		
/[+son]|f/	[f]	33	75%	NA	NA
	[p]	4	9%		
	[v]	6	14%		
	[b]	0	0%		
/[+son]|θ/	[θ]	0	0%	.3986	.3334
	[t]	3	100%		
	[ð]	0	0%		
	[d]	0	0%		

The above table shows that although the percentages of variants for /#θ/ are significantly different than the percentages of variants for /#f/, the distribution of raw tokens is not significantly different. Additionally, Although the percentages of variants of /[+son]|θ/ are not significantly different from the percentages of variants of /[+son]|f/, they are still different enough so as to require slightly different amounts of overlap of the relevant constraints. However, the distributions of raw tokens for these two items are not significantly different. For both /#θ/ and /[+son]|θ/, the percentages of variants are supported by very few tokens, all of which represent the same lexical item, /θɪŋkɪnabat/. It is difficult to determine if additional tokens for this lexical item would reveal it to have idiosyncratic variation or whether it would

---

[87] Where $p < .05$ is significant. In all of these averages, I did not include tests where both of the values for variants of the same input were 0; this always resulted in a p-value of 1.0, which seems to be a default lack of significance that does not accurately reflect the data. If I included those values in the averages, the p-values for /[+son]|θ/ would be .6993 and .6667 (from left to right across the columns).

ultimately pattern similarly to word-initial /#f/. However, the lack of significant differences in the distribution of raw tokens suggests that the difference in percentages for word-initial /#θ/, both after sonorants and non-sonorants—is not due to lexical or phoneme-specific variation. Therefore, it is logical to collapse the tokens for word-initial fricatives in Trial 3 in order to examine behavior of the natural class as a whole.

Additionally, the table on p.59 indicates that the out of all of the fricative phonemes in word-initial position, the grammar was most accurate at predicting the frequencies of variants for /#θ/. Given these results, it seems that weighing word-initial /#θ/ and /[+son]|θ/ equally with the other underlying forms in Trial 2 affected the overall results, likely skewing them in favor of the percentages of variants of /#θ/ and /[+son]|θ/. This is problematic because /#θ/ and /[+son]|θ/ are the fricatives with the fewest tokens in comparison with the other phonemes in their conditions; thus, they are less likely than the other fricative phonemes to reliably reflect the actual frequencies of word-initial fricatives in the language, simply because there are so few tokens in the corpus. Once again, this is an argument for combining the tokens of word-initial fricatives in Trial 3, given that their distributions of raw tokens do not differ significantly. Below are the rankings from the final stochastic grammar for Trial 2:

Constraint	**Ranking Value**
AGREE(voice)	102.155
MAX	100.244
IDENT(voice)	99.947
*GLOTTAL	99.756
AGREE(cont)	99.429
IDENT(strid)	98.870
*VOICEOBS	98.814
IDENT(cont)	98.757
*[+son]|[-voice]	98.143
*[+strid]	97.900
*FRIC	97.896

This table shows that the grammar for Trial 2 has a more refined constraint ranking than the grammar in Trial 1. Violations of AGREE(voice) are still highly dis-preferred, and fricatives, voiced obstruents, stridents, and initial fricatives after words ending in sonorants are often allowed. Most of the other constraints are very close together, although they have separated out slightly into clumps around 99, 98 and 97. With a standard deviation of 2.0, these constraints overlap to a large extent. In order to closer examine the patterns of individual phonemes in this trial, I ran a subset of the data that did not include /#h/ (both after sonorants and non-sonorants), initial /#d/, initial /#θ/ (both after sonorants and non-sonorants), intervocalic /VpV/, and intervocalic /VsV/ either because they had comparatively few tokens or because they seem to behave differently than the other phonemes in their natural classes.[88] The following table presents the results:

Condition	Underlying Representation	Variant	Observed Number of Tokens	Observed Frequency	Praat Output Frequency	Difference
1	/#b/	[b]	42	72%	57%	15
		[p]	16	28%	43%	15
1	/#g/	[g]	31	67%	57%	10
		[k]	15	33%	43%	10
2	/#k/	[k]	35	100%	100%	0
		[g]	0	0%	0%	0
3	/#f/	[f]	23	85%	85%	0
		[p]	4	15%	15%	0
3	/#s/	[s]	9	64%	70%	6
		[c]	1	7%	9%	2
		[tç]	4	29%	21%	8
4	/[+son]|f/	[f]	33	75%	78%	3
		[p]	4	9%	16%	7
		[v]	6	14%	6%	8
		[b]	0	0%	0%	0
4	/[+son]|s/	[s]	16	67%	61%	6
		[c]	4	17%	9%	8
		[z]	1	4%	4%	4

---

[88] I did not include word-initial /s/ because it is the only fricative that varies with affricates. The rest of the phonemes that I excluded were removed because they have comparatively fewer tokens than the other phonemes in their respective conditions. I also included only one word-initial voiceless stop because the grammar always perfectly predicts the variants of these forms because they are undominated and will always surface.

		[ɟ]	0	0%	0%	0
		[tç]	3	12%	26%	14
		[dʝ]	0	0%	0%	0
7	/VbV/	[b]	9	69%	73%	4
		[p]	0	0%	5%	5
		[f]	0	0%	1%	1
		[v]	4	31%	21%	10
7	/VgV/	[g]	11	59%	58%	1
		[k]	0	0%	3%	3
		[x]	0	0%	2%	2
		[ɣ]	8	41%	37%	4
8	/VkV/	[g]	8	16%	53%	37
		[k]	17	34%	9%	25
		[x]	4	8%	5%	3
		[ɣ]	21	42%	33%	9
9	/VfV/	[b]	11	30%	9%	21
		[p]	1	3%	1%	2
		[f]	0	0%	13%	13
		[v]	25	67%	77%	10
9	/VθV/	[d]	8	31%	10%	21
		[t]	0	0%	2%	2
		[θ]	1	4%	13%	9
		[ð]	17	65%	75%	10

This table shows that the removal of initial /#θ/ and /#h/, both after sonorants and non-sonorants, allows the grammar to better predict the frequencies of variants for the other initial fricatives, also both after sonorants and non-sonorants. This discrepancy is perhaps due to a decrease in overall noise by removing items with comparatively small amounts of tokens: /#θ/ has four tokens, /[+son]|θ/ has nine, /#h/ has nine, and /[+son]|θ/ has nine. The table above also shows that while the removal of intervocalic /VpV/ coincides with an improvement in the accuracy of the grammar's predicted variants for the other intervocalic stop, /VkV/, the removal of /VsV/ has not much helped the grammar predict the frequencies of variants of the other intervocalic voiceless fricatives. It is possible the decrease in accuracy for intervocalic fricatives is in part due to complicated interactions resulting from the removal of the word-initial fricatives /#θ/, /[+son]|θ/, /#h/, and /[+son]|θ/. Another possible explanation for these differences in the performance of the grammar is that the constraint inventory fails to account for some phonological process or processes. However, it is difficult to identify this as the decisive factor when small amounts of tokens could be impacting the output grammar as well. This conflict lends support to collapsing the data in Trial 3 to increase the amount of tokens for each input, which somewhat controls for token amounts so that I can more closely examine the efficacy of the constraint inventory. The following statistics show that this subset of the data in Trial 2 output an overall more accurate grammar than the Trial 2 grammar that uses the entire dataset:

Calculation	**Value**
Standard Deviation of Actual Frequencies	31.00
Standard Deviation of Praat's Frequencies	31.34
Pearson Coefficient	.94
Average Difference (in % Points)	7.07

These statistics, which have improved from the trial with the full dataset for Trial 2, suggest that eliminating inputs with comparatively few tokens increases the accuracy of the grammar's predictions. However, in order to more closely examine the effectiveness of the constraint inventory in accounting for the phonological patterns in the grammar, the token data from Trial 2 is collapsed into natural classes in Trial 3.

3.2.3 Trial 3: Natural Classes

In Trial 3, the algorithm generated a grammar in Stochastic OT based on a list of natural classes in word-initial and intervocalic positions totaling in 9 underlying representations and 444 tokens. I examined the frequencies for each variant that the output grammar predicted and compared those with the actual frequencies in the corpus data. The statistics for Trial 3 are presented in the following table:

Calculation	**Value**
Standard Deviation of Actual Frequencies	29.82
Standard Deviation of Praat's Frequencies	29.39
Pearson Coefficient	.93
Average Difference (in % Points)	8.03

The results show a further decrease from the full dataset of Trial 2 in the difference between the standard deviations of each set of frequencies, while the Pearson coefficient indicates that the individual pairs of frequencies were highly correlated. Additionally, the average difference in percentage points between the observed and predicted frequency of each variant decreased substantially from the values in the full dataset run of Trial 2. The following table displays the observed frequencies and the grammar's output frequencies for Trial 3:

Condition	Underlying Representation	Variant	Observed Number of Tokens	Observed Frequency	Praat Output Frequency	Difference
1	/#+voice, -cont/	[+voice, -cont]	83	69%	66%	3
		[-voice, -cont]	38	31%	34%	3
2	/#-voice, -cont/	[-voice, -cont]	70	100%	100%	0
		[+voice, -cont]	0	0%	0%	0
3	/#-voice, +cont/	[-voice, +cont]	25	81%	81%	0
		[-voice, -cont]	6	19%	19%	0
4	/[+son]|+voice, -cont/	[-voice, +cont]	33	72%	70%	2
		[+voice, +cont]	6	13%	6%	7
		[-voice, -cont]	7	15%	24%	9
		[+voice, -cont]	0	0%	0%	0
5	/#h/	[h]	3	33%	37%	4

		[ʔ]	0	0%	3%	3
		Ø	6	67%	60%	7
6	/[＋son]|h/	[h]	4	44%	34%	10
		[ʔ]	0	0%	2%	2
		Ø	5	56%	61%	5
		[ɦ]	0	0%	3%	3
7	/V [＋voice, -cont] V/	[＋voice, -cont]	21	64%	53%	11
		[-voice, -cont]	0	0%	3%	3
		[-voice, ＋cont]	0	0%	2%	2
		[＋voice, ＋cont]	12	12%	42%	6
8	/V [-voice, -cont] V/	[＋voice, -cont]	8	13%	46%	33
		[-voice, -cont]	21	36%	10%	26
		[-voice, ＋cont]	4	7%	8%	1
		[＋voice, ＋cont]	26	44%	36%	8
9	/V [-voice, ＋cont] V/	[＋voice, -cont]	21	32%	7%	25
		[-voice, -cont]	1	1%	2%	1
		[-voice, ＋cont]	2	3%	16%	13
		[＋voice, ＋cont]	42	64%	75%	11

This table indicates that the final grammar for Trial 3 was entirely accurate in predicting the variants of initial voiceless stops and initial voiceless fricatives after non-sonorants, and it was very accurate in predicting the variants of initial voiced stops, initial fricatives after sonorants, and initial /h/ after sonorants and non-sonorants. In the intervocalic position, the grammar was better at predicting the variants of underlyingly voiced stops than voiceless stops or voiceless fricatives. The largest difference in empirically observed percentages and the grammar's output percentages was for intervocalic voiceless stops, which the grammar predicted would surface as voiced stops more frequently than the observed data indicated; it is possible that there are phonological processes affecting these segments that the current constraints do not account for.[89] The following chart presents the final rankings for the constraints in Trial 3:

Constraint[90]	**Ranking Value**
AGREE(voice)	102.817
*GLOTTAL	100.238
IDENT(cont)	100.118
AGREE(cont)	99.979
MAX	99.762
IDENT(voice)	99.744
*VOICEOBS	98.560
*FRIC	97.604
*[+son]||[-voice]	96.856

This table shows that the final stochastic grammar for Trial 3 prefers forms that abide by intervocalic voicing and frication and that it often allows voiced obstruents and fricatives, even word-initial voiceless ones after sonorants. However, all of the constraints are relatively close together, which translates to most of them overlapping (with a standard deviation of 2.0) to result in complicated patterns of variation. I also ran a subset of the data in Trial 3 without /h/ to examine the patterns more closely. The following table presents the results:

---

[89] See below for trials with additional constraints.

[90] Because I eliminated word-initial /s/ from this dataset and /s/ does not vary with affricates intervocalically, the IDENT(strid) and *[+strid] were not included in this trial.

Condition	Underlying Representation	Variant	Observed Number of Tokens	Observed Frequency	Praat Output Frequency	Difference
1	/#＋voice, -cont/	[＋voice, -cont]	83	69%	77%	8
		[-voice, -cont]	38	31%	23%	8
2	/#-voice, -cont/	[-voice, -cont]	70	100%	100%	0
		[＋voice, -cont]	0	0%	0%	0
3	/#-voice, ＋cont/	[-voice, ＋cont]	25	81%	66%	15
		[-voice, -cont]	6	19%	34%	15
4	/[＋son]|＋voice, -cont/	[-voice, ＋cont]	33	72%	55%	17
		[＋voice, ＋cont]	6	13%	3%	10
		[-voice, -cont]	7	15%	42%	27
		[＋voice, -cont]	0	0%	0%	0
7	/V [＋voice, -cont] V/	[＋voice, -cont]	21	64%	43%	21
		[-voice, -cont]	0	0%	1%	1
		[-voice, ＋cont]	0	0%	2%	2
		[＋voice, ＋cont]	12	36%	54%	19
8	/V [-voice, -cont] V/	[＋voice, -cont]	8	13%	35%	22

		[-voice, -cont]	21	36%	9%	27
		[-voice, +cont]	4	7%	11%	4
		[+voice, +cont]	26	44%	45%	1
9	/V [-voice, +cont] V/	[+voice, -cont]	21	32%	11%	21
		[-voice, -cont]	1	1%	3%	2
		[-voice, +cont]	2	3%	18%	15
		[+voice, +cont]	42	64%	68%	4

This table shows that the exclusion of /h/ actually causes the grammar to predict the variants of word-initial fricatives less accurately than when /h/ is included. The grammar for this subset also makes less accurate predictions for the variants of word-initial voiced stops and intervocalic voiceless segments than the grammar from the initial run of Trial 3. The following statistics indicate that this subset of Trial 3 is less accurate than the trial with the complete dataset:

Calculation	**Value**
Standard Deviation of Actual Frequencies	30.80
Standard Deviation of Praat's Frequencies	29.05
Pearson Coefficient	.88
Average Difference (in % Points)	10.86

Perhaps the addition of constraints that deal more specifically with faithfulness and markedness of different natural classes in specific phonological environments—along with a reanalysis of some of the underlying forms—would improve the accuracy of the Trial 3's grammar. In order to examine this, I ran a number of trials with different combinations of the following three changes: I reanalyzed all instances of underlying intervocalic fricatives as underlying voiced stops,[91] I added a constraint that penalized underlyingly voiceless stops that were unfaithful in continuancy and voicing in their output correspondents, and I added a constraint that penalized voiceless fricatives in the intervocalic position. The formal definitions of these two constraints are as follows:

IDENT[voice, cont]$_{VS}$: Assign one violation mark for every feature of an input voiceless stop that changes in its output correspondent, where the set of features considered is [±voice, ±continuant].
***V[-voice, +cont]V:** Assign one violation mark for every voiceless fricative in the intervocalic position.

These constraints are an effort to improve the accuracy for predicted variants of intervocalic voiceless stops, one of the inputs for which the original output grammar of Trial 3 was the least successful at modeling. The first constraint is intended to increase the output grammar's frequency of voiceless stop variants from underlying intervocalic voiceless stops;[92] the original output grammar from Trial 3 predicted that underlying intervocalic voiceless stops would surface more frequently as their voiced stop variants and less frequently as their voiceless stop variants than the empirical data indicated. The second constraint is meant to decrease the output grammar's frequency of voiceless fricative variants from underlyingly voiceless stops, because when the underlying forms of intervocalic voiceless fricatives are changed to be underlying

---

[91] I ran a previous set of trials with only the underlying fricatives in /nɛfə/, /oʊfə/, and /ɹɪfə/ changed to stops (see footnote 80). In that set of trials, I only changed those three forms because I determined the other forms with intervocalic fricatives to be less ambiguously fricatives in their underlying forms, and I also determined that intervocalic underlyingly voiceless fricatives had some variants that were unattested for intervocalic underlyingly voiced stops. In this trial, however, I changed all of the underlying forms of intervocalic voiceless fricatives with the understanding that it somewhat changes my analysis of the allowable variants of underlyingly voiced intervocalic stops (i.e. my previous analysis determined that they only surface as voiced stops and voiced fricatives, while changing the underlying forms of intervocalic fricatives to intervocalic voiced stops necessitates that they have both voiceless fricative and voiceless stop variants as well). In changing all of the forms with underlyingly voiceless intervocalic fricatives to voiced stops, I also had to remove one item's tokens—/seɪbəm/—because its distributions of tokens differed significantly with some of the newly altered items.

[92] The constraint does not refer to the intervocalic position specifically but is instead stated as a general faithfulness constraint because in word-initial position, underlyingly voiceless stops are always undominated and thus always surface. In other words, formulating this constraint in the most general way allows it to capture more of the phonological trends in the language.

voiced stops, the original output grammar from Trial 3 ends up predicting more voiceless fricative variants from underlyingly voiceless intervocalic stops than the empirical data exhibits.

Out of the seven additional trials that I ran combining the three changes listed above, the following trials were more accurate at predicting variants than the original run of Trial 3: the trial with just the underlying forms changed, with the underlying forms changed and with the IDENT[voice, cont]$_{VS}$ constraint added, with the underlying forms changed and the *V[-voice, +cont]V constraint added, with the underlying forms changed and both constraints added, and with just the addition of the two constraints.[93] Out of these trials, the one that showed the most improvement from the original run of Trial 3 was the grammar that combined the changed underlying representations and the *V[-voice, +cont]V constraint. The statistics for this rerun are presented below:

Calculation	**Value**
Standard Deviation of Actual Frequencies	30.04
Standard Deviation of Praat's Frequencies	31.06
Pearson Coefficient	.98
Average Difference (in % Points)	4.98

Although the difference between the standard deviations of the observed and predicted frequencies of variants has increased slightly from the original run of Trial 3, the increase is negligible. The table above shows that the Pearson coefficient has increased and the average difference in percentage points between the empirically observed and output frequencies has decreased substantially, indicating that the grammar predicted the frequencies of variants much more accurately than the original Trial 3 grammar. These results suggest that perhaps some of the original underlying forms were transcribed incorrectly. More importantly, the results also show that the use of additional constraints, such as a markedness constraint against intervocalic voiceless fricatives, improves the accuracy of the grammar and thus is a better model of phonological variation in Kimberley Kriol. Below, I discuss the overall patterns from the three trials run in Praat.

---

[93] The other two trials—one with just the *V[-voice, +cont]V constraint added and one with just the IDENT[voice, cont]$_{VS}$ constraint added—both performed worse than the original run of Trial 3.

3.3.3 Overall Trends

Generally speaking, the further I abstracted away from the individual lexical items across the three trials, the more accurately the constraint rankings of the final grammar predicted the frequencies of variants in Kimberley Kriol. This is evident in the steady increase in the Pearson coefficient, the steady decrease in the difference between standard deviations, and the decrease in the average difference between the actual frequencies and the output grammar's predictions.[94] Similarly, the constraint ranking became more refined across the trials and ultimately resulted in a series of constraints that overlap substantially. Although the grammar generated from the subset of Trial 2 was the most accurate in predicting frequencies of variants out of all of subsets of the Trials, its performance was only negligibly better than the grammar learned from the full dataset in Trial 3. These results suggest that the most accurate model is Trial 3, which makes general claims about natural classes. The addition of new constraints and the changing of multiple underlying forms in the reruns of Trial 3 greatly improved the grammar's accuracy, suggesting that there were most likely some phonological processes that the original grammar for Trial 3 failed to take into account. The next section discusses possible accounts for the performance of the grammars and presents potential improvements for the model.

4 Discussion and Conclusion

4.1 Discussion

Overall, the three trials run through Praat's Stochastic OT learning program indicate that a Stochastic OT account of phonological variation among singleton stops and fricatives in Kimberley Kriol is possible. The tableaus presented in Section 3.1 show the necessary constraint rankings for all of the possible variants of word-initial and intervocalic stops and fricatives in the dataset. In Section 3.1, I also provide a comprehensive list of constraints that deal with the phenomena and identified constraints that should clump together—voicing constraints, stridency constraints, continuancy constraints, and constraints dealing with retaining or deleting glottal segments—such that they overlap in a Stochastic OT account. The resulting constraint rankings in the stochastic grammars accurately reflected the clumped constraints; additionally, the constraint rankings showed that the individual clumps overlap to a large extent. The constraint

---

[94] The improvement across the trials could in part be due to a simple decrease in statistical noise as the amounts of tokens increased for each input form. See Section 4.1 for a more in-depth discussion of the results.

rankings that the stochastic grammars output show that the constraints under discussion must *all* be ranked very close to one another in order to create the kind of overlap that will allow for all the attested variants in the correct proportions.

Generally speaking, the grammars generated by Praat's implementation of the GLA improved their predicted frequencies of specific variants as the input underlying forms abstracted further away from individual lexical items. In other words, the average difference between the observed frequency of variants and those predicted by the final stochastic grammars decreased substantially from Trial 1 to Trial 3. However, the individual trials revealed a number of intriguing nuances that a model of Kimberley Kriol phonology should ultimately take into account or at least acknowledge. The results of Trial 1 suggest that a small number of lexical items are exceptions from the general phonological patterns. Additionally, the success of the final reruns of Trial 3 suggests that additional constraints are needed to account for the behavior of specific natural classes in certain phonological environments. The fact that the full dataset in Trial 3 generated a more accurate grammar than the subset of Trial 3 suggests that while natural classes generally pattern together, some constraints addressing specific places of articulation—*GLOTTAL, for example—are necessary to fully account for the observed patterns of variation.[95] Overall, the results from the Praat trials indicate that the most accurate and comprehensive model of phonological variation in Kriol would describe phonological tendencies of overall natural classes while incorporating constraints that address the behavior of specific classes or places of articulation when necessary. Out of the three grammars tested, the closest model of the grammar of Kimberley Kriol is the rerun of Trial 3 with altered underlying forms for some intervocalic segments and an added markedness constraint against intervocalic voiceless fricatives.

The Praat models did not ultimately converge on a perfectly correct grammar, but the overall improvement across the trials suggests that such convergence is possible with further refinement of the input and constraints employed. Although there is no universal metric within the field by which to measure the success of a learned grammar, a good point of reference is the differences in the observed and predicted frequencies Boersma and Hayes' case studies (2001), which range from a .39% to 1.09% difference between the empirical frequencies and those

---

[95] Perhaps adding additional constraints dealing with specific places of articulation to Trial 2 would improve that trial's accuracy. I only implemented reruns with additional constraints for Trial 3.

output by the learned grammar.[96] Clearly, none of my trials predicted frequencies of variants as accurately as Boersma and Hayes' models. However, although my models are imperfect, they are a step in the right direction. The reruns of Trial 3 show that adding more specific constraints to the grammars in Trials 1-3—such as markedness constraints against specific places of articulation, individual markedness constraints for different natural classes in specific places of articulation, and individual voicing or continuancy faithfulness constraints for different natural classes—might increase the accuracy of the grammar's predicted frequencies for specific variants.

Apart from refining the constraint inventory for the analysis of Kimberley Kriol phonology, cleaner and more abundant data would likely improve the accuracy of the models. As I have mentioned throughout the analysis, there are numerous lexical items in my dataset that have significantly fewer tokens than others, as well as certain segments (such as word-initial /#θ/ and /[+son]||θ/, intervocalic /VpV/, and intervocalic /VsV/) or natural classes that had fewer tokens. Because of this, there are some variants that are not attested that should theoretically be possible given the posited constraints and their necessary amount of overlap. For example, it should theoretically be possible for the voiceless fricative /θ/ to surface as the voiced variant [ð] word-initially after a word ending in a sonorant, but the data has no tokens that represent this variation. These holes in the corpus data may be accidental gaps; it is possible that they exist in the speech patterns of the participants and I just do not have any recorded instances of them. Trial 2 and Trial 3 decreased statistical noise by collapsing the data into fewer underlying representations with more tokens. However, it would ideally be better to collect more tokens from more recorded data and stick to an analysis of the lexical items rather than an abstract analysis of phonemes and natural classes. The abstraction away from individual forms also potentially glosses over any anomalous lexical items that would defy general phonological patterns.

Additionally, the details of my data elicitation may have created subsequent challenges for my data analysis. Although I knew at the outset that I intended to do a phonological analysis of variation in Kimberley Kriol, I had no previous fieldwork and only a basic knowledge of creole studies. Because I was not sure exactly what kind of data I was looking to use, I used a wide variety of elicitation tactics with my subjects. As a result, I ended up pulling my tokens

---

[96] See section 2.2.2 for a more in-depth discussion of the procedure in Boersma and Hayes 2001.

from a number of very different elicitation tasks that may have affected the nature of the speech differently. For example, tokens from my wordlist elicitations were probably pronounced more carefully than tokens from the more natural speech during the Frog Story description. However, this variety of elicitation tactics is arguably a strength of my data in that it allows for more variation to occur in the subjects' speech. An added confounding factor was that two out of three of my recordings were of female research participants, so most of my data—and all of my data from the Frog Story—comes from their speech. There is no discernable effect of gender, but it is true that my data is dominated by the two female speakers. On one hand, this makes the data more cohesive and allows for stronger arguments of intra-speaker variation. On the other hand, more data from the male participants would strengthen claims about variation across the entire lect. In an ideal situation, I would have data from more than four speakers with the tokens evenly distributed across the subjects.

There are a few ways in which I could have conducted my data acquisition differently. I could have had the participants do the same elicitation tasks that I used, but I could have left the room and let them do the tasks on their own. This would have better facilitated natural speech and might have engendered a more basilectal lect of Kriol, seeing as my presence in the room most likely caused the speakers to tend towards the acrolectal end of the continuum.[97] My absence from the room would also have eliminated their use of Aboriginal English, which they used occasionally in order to speak to me.[98] I could also have recorded speakers in a more informal setting, such as in the community or on field trips. This would potentially have allowed me to capture more variation in a more natural setting, but it would have also been less controlled and more difficult to analyze.

4.2 Conclusion

Despite the challenges that my data presented for my analysis, I was able to represent phonological variation in Kimberley Kriol fairly effectively in a Stochastic Optimality Theory model. The improvement in accuracy of predicted frequencies of variants from the grammars across the trials shows promise for the learnability of the phonological grammar of Kimberley

---

[97] This is why I determined the lect of my speakers to be the upper mesolect.

[98] I made the decision to remain in the room during elicitation in order to better explain the elicitation tasks, ask for clarification, and elicit such things as sentences and wordlists translated from English into Kriol.

Kriol in the Gradual Learning Algorithm. The fact that the average difference in observed and predicted frequency decreased across the trials suggests that while none of the trials output a flawless grammar, it is possible to account for free variation among stops and fricatives in the upper mesolect of Kimberley Kriol within a single grammar in the Stochastic Optimality Theory framework. In this way, my hypothesis was borne out; my results indicate that Stochastic OT, which can be used to account for variation in other natural languages, is compatible with the amount and type of variation present in Kimberley Kriol.

My hypothesis predicted that the most abstract trial, Trial 3, would most closely model the phonological variation in Kimberley Kriol because it is the most general model that describes phonological tendencies of overall natural classes. The fact that the grammar generated from Trial 3—in particular the rerun that changed underlying forms for intervocalic fricatives and added a markedness constraint against intervocalic voiceless fricatives—predicted the most accurate frequencies for variants out of the three models supports my hypothesis. The results suggest that the most accurate model of the phonological variation in Kimberley Kriol accounts for patterns in general natural classes while allowing specific places of articulation, such as glottal segments, to exhibit slightly different patterns of variation. The resulting stochastic grammars from Praat's GLA are admittedly imperfect and stand to improve their accuracy. However, with more refined and comprehensive constraints relating to the behavior of specific natural classes in particular environments—as well as cleaner and more abundant data—Praat's OT Learning Program would converge on a grammar that would output frequencies for each form with differences in the range of Boersma and Hayes' case studies (2001).

These results have broader implications for creole studies in general, in which creole languages are often described as highly variable. If the phonological variation of the upper mesolect of a creole—which can either be thought of as a single 'lect' or as multiple lects that bleed into each other—can be encompassed within one stochastic grammar, perhaps the entire creole continuum can be described by a single grammar. The challenge of containing the entire creole continuum within a single overarching grammar has been acknowledged before: "…the 'creole continuum' poses a greater challenge to the systematic analysis of variation because of the distance between the end points, and the difficulty of writing one grammar that can span all the variants" (Labov 1971 in Sankoff 1980: 156). In discussing panlectal grids and implicational scaling[99], Rickford says,

> What we aim to be describing when writing a grammar of a continuum is a set of *varieties*…which would involve combinations and interrelations of elements from different [grammatical] subsystems and levels. But what we have been increasingly restricted to is the analysis of *variants* and sets of variants within individual subsystems…it remains true that people do not speak with pronoun systems along, nor with copulas, nor with tense markers, but with combinations of all these subsystems and more. The key problem, therefore, especially as we try to bring our analyses closer to the realities of language use in pidgin-creole continua, is to attempt descriptions of such combinations…However, attempts to apply this concept in actual studies of pidgin-creole continua have not met with much success. (Rickford 1980: 169).

As Rickford points out, a comprehensive description of the grammar of a creole or any other language cannot be restricted to one subset of the language. Stochastic Optimality Theory provides a promising method of describing the full amount of variation across a creole continuum within one grammar, although it is most frequently used for phonological analysis. In order to fully describe the grammar of a creole continuum, one would have to combine OT analyses of variation in multiple areas of the language, such as syntax and phonology. However, this is a daunting and perhaps impractical task that perhaps serves more as a theoretical possibility than something to be physically implemented.

Overall, there are undeniably ways in which to improve the data collection and analysis of the study presented here, but the increasing accuracy of the GLA trials in Praat suggests that it is entirely possible to account for the phonological variation present in Kimberley Kriol by using Stochastic Optimality Theory. Future extensions of this Stochastic OT approach would preferably include a more in-depth analysis of the constraints at play, more subjects, cleaner data, and more types of phonological variation—both if applied to Kimberley Kriol and other to creole languages. With further refinement of the methods, materials, and analysis, Stochastic Optimality Theory would be an effective way to accurately describe and account for the phonological variation in Kimberley Kriol and other creole languages.

---

[99] See Section 1.1.4

References

Alber, Birgit, and Ingo Plag. 2001. Epenthesis, deletion and the emergence of the optimal syllable in creole: The case of Sranan. *Lingua* 111:811–840.

Alleyne, Mervyn C. 1980. Introduction: Theoretical orientations in creole studies. In *Theoretical orientations in creole studies*, 1–18. New York: Academic Press, Inc.

Anttila, Arto, and Young-mee Yu Cho. 1998. Variation and change in Optimality Theory. *Lingua* 104:31–56.

Anttila, Arto. 1997. Deriving variation from grammar. In *Variation, change and phonological theory*, 35–68. Amsterdam: John Benjamins Publishing.

Artstein, Ron. 1998. The incompatibility of underspecification and markedness in Optimality Theory. *RuLing Papers* 1:7–13.

Bickerton, Derek. 1980. Decreolisation and the creole continuum. In *Theoretical orientations in creole studies*, 109–128. New York: Academic Press, Inc.

———. 1973. The nature of a creole continuum. *Language* 49(3):640–669.

Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm.

Boersma, Paul. 1997. How we learn variation, optionality, and probability.

Boersma, Paul, and David Weenink. 1992-2010. *Praat*.

Coetzee, Andries, and Joe Pater. 2009. The place of variation in phonological theory.

DeCamp, David. 1971. Toward a generative analysis of a post-creole speech continuum. In *Pidginization and creolization of languages*, 349–70. Cambridge: Cambridge University Press.

Hume, Elizabeth, and Georgios Tserdanelis. 2002. Labial unmarkedness in Sri Lankan Portuguese Creole. *Phonology* 19(3):441–458.

Inkelas, Sharon, Orhan Orgun, and Cheryl Zoll. 1997. The implications of lexical exceptions for the nature of grammar. In *Derivations and constraints in phonology*, 393–418. Oxford: Oxford University Press.

Inkelas, Sharon. 1994. The consequences of optimization for underspecification.

Ito, Junko, and Armin Mester. 2001. Covert generalizations in Optimality Theory: The role of stratal faithfulness constraints. *Studies in Phonetics, Phonology and Morphology* 7.

Ito, Junko, Armin Mester, and Jaye Padgett. 1995. Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26(4):571–613.

Jarosz, Gaja. 2011. Learning with violable constraints.

Kimberley Kriol - English Wordlist. 2006.

Klein, Thomas B. 2006. Creole phonology typology: Phoneme inventory size, vowel quality distinctions and stop consonant series.

———. n.d. Syllable structure and lexical markedness in creole morphophonology: Determiner allomorphy in Haitian and elsewhere.

Lipski, John M. 1999. Epenthesis vs. elision in Afro-Iberian language: A constraint-based approach to creole phonology. In *Current issues in Romance languages*, 173–188. Philadelphia: John Benjamins Publishing.

Migge, Bettina, and Norval Smith. 2007. Substrate influence in creole formation. *Journal of Pidgin and Creole Languages* 22(1):1–15.

Mühlhäusler, Peter. 1980. Structural expansion and the process of creolization. In *Theoretical orientations in creole studies*, 19–55. New York: Academic Press, Inc.

Pater, Joe. 2008. Gradual learning and convergence. *Linguistic Inquiry* 39(2):334–345.

Plag, Ingo. 2009. Creoles as interlanguages: Phonology. *Journal of Pidgin and Creole Languages* 24(1):119–138.

Prince, Alan, and Paul Smolensky. 2993. Optimality Theory: Constraint interaction in generative grammar.

Rickford, John R. 1980. Analyzing variation in creole languages. In *Theoretical orientations in creole studies*, 165–184. New York: Academic Press, Inc.

———. 1973. Phonological reduction of *doz* and some other tense-aspect markers.

Sandefur, John, R. 1982. *An introduction to conversational Kriol*. Vol. 5. B. Summer Institute of Linguistics: Australian Aborigines Branch.

Sandefur, John R. 1986. *Kriol of North Australia: A language coming of age*. Vol. 10. A. Summer Institute of Linguistics: Australian Aborigines Branch.

Sankoff, Gillian. 1980. Variation, pidgins and creoles. In *Theoretical orientations in creole studies*, 139–164. New York: Academic Press, Inc.

Schultze-Berndt, Eva, Felicity Meakins, and Denise Angelo. 2012. Kriol.

Siegel, Jeff. 2008. *The emergence of pidgin and creole languages*. Oxford; New York: Oxford University Press.

Webb, Eric Russell. 2010. Creole phonological restructuring: The role of perception in contact induced change. *Journal of Pidgin and Creole Languages* 25(2):263–288.

Winford, Donald. 2008. Processes of creole formation and related contact-induced language change. *Journal of Language Contact* 2(1):124–145.