

The Development of Phonological Stratification: Evidence from Stop Voicing Perception in Gurindji Kriol and Roper Kriol

Jesse Stewart

University of Saskatchewan
stewart.jesse@usask.ca

Felicity Meakins

University of Queensland
f.meakins@uq.edu.au

Cassandra Algy

Karungkarni Arts
karungkarniart@gmail.com

Angelina Joshua

Ngukurr Language Centre
contact@ngukurrlc.org.au

Abstract

This study tests the effect of multilingualism and language contact on consonant perception. Here, we explore the emergence of phonological stratification using two alternative forced-choice (2AFC) identification task experiments to test listener perception of stop voicing with contrasting minimal pairs modified along a 10-step continuum. We examine a unique language ecology consisting of three languages spoken in Northern Territory, Australia: Roper Kriol (an English-lexifier creole language), Gurindji (Pama-Nyungan), and Gurindji Kriol (a mixed language derived from Gurindji and Kriol). In addition, this study focuses on three distinct age groups: children (group I, 8>), preteens to middle-aged adults (group II, 10–58), and older adults (group III, 65+). Results reveal that both Kriol and Gurindji Kriol listeners in group II contrast the labial series [p] and [b]. Contrarily, while alveolar [t] and velar [k] were consistently identifiable by the majority of participants (74%), their voiced counterparts ([d] and [g]) showed random response patterns by 61% of the participants. Responses

to the voiced stimuli from the preteen-adult Kriol group were, however, significantly more consistent than in the Gurindji Kriol group, suggesting Kriol listeners may be further along in acquiring the voicing contrast. Significant results regarding listener exposure to Standard English in both language groups also suggests constant exposure to English maybe a catalyst for setting this change in motion. The more varied responses from the Gurindji, Kriol, and Gurindji Kriol listeners in groups II and III, who have little exposure to English, help support these findings.

Keywords

language contact – perception – stop voicing – Gurindji Kriol – Kriol – mixed language

1 Introduction

Contact languages provide a unique opportunity for analysing extensive language change in a considerably short period of time. While language change under normal circumstances can take generations before variation is quantifiable, the effects of language contact can often be seen in as little as a single generation. In the case of mixed languages, an extreme variety of contact language, entire lexical and grammatical elements can transfer from one language into another, often within a single generation. Mixed languages are often characterised as containing the lexical or grammatical patterns of linguistic elements from different languages, referred to as stratification. While there is a wealth of literature that explores lexical and grammatical stratification in mixed languages and languages under other intense contact scenarios (Hickey, 2010b; Matras and Bakker, 2003; S. Thomason, 1997), we are only just beginning to understand how the phonological systems from different languages interact at the phonetic level (both in production and perception) in situations of language contact. In the more classical sense, we are only just beginning to understand the effects of phonological interference in mixed languages. This paper adds to this literature with a perceptual study that explores a specific phonemic conflict site (a conflicting area of phonological convergence) involving stop voicing contrasts. Here, we provide a synchronic description of stop perception in three Australian languages that have either emerged or changed through considerable contact via English: Kriol, Gurindji Kriol, and Gurindji.

These languages provide an interesting test case for exploring the perceptual effects of language contact in that the clear majority of Australian languages

do not have contrastive stop voicing, while English, the primary lexifier of Kriol and Gurindji Kriol, clearly does. We show that, based on perceptual data (and with further evidence from production data from previous studies (Bundgaard-Nielsen and Baker, 2016; Jones and Meakins, 2013)), it is likely that during the development of Kriol and Gurindji Kriol, the stop contrast was not initially present. Through constant and increasing contact with English and the recognition of Kriol/English cognates, however, there is evidence that a voicing contrast is now developing in both languages—an example of language contact adding complexity to a linguistic system. This study also demonstrates that Kriol is more advanced in the development of a voicing contrast than Gurindji Kriol, which is likely the result of 50 years more exposure to English through the earlier presence of formal Western education.

Importantly, this study provides further evidence that linguistic systems and complexity develop incrementally and with variation. Previous studies of loan phonology typically characterise loan words as either conforming wholesale to the recipient language phonology or as categorically introducing new phonemes into the recipient language in restricted areas of the lexicon (and thereby creating stratification) (Bullock, 2009; Campbell, 1996; Hyman, 1970; Itô and Mester, 1995; Matras, 2009). This study provides a more nuanced picture of how stratification occurs. In this respect, this study joins the growing morphosyntactic literature on mixed languages that demonstrates the complex nature of language development under intensive contact with other languages. Earlier studies often characterised mixed languages as faithfully replicating the morphosyntactic patterns from both of their source languages, however subsequent studies have noted that transferred grammatical elements often undergo change when they are absorbed into the recipient language (often under the influence of patterns in the recipient language). Furthermore, absorption is not categorical but is an incremental process, resulting in variation among speakers. For example, the transfer of the Gurindji ergative suffix into Gurindji Kriol in the genesis of this mixed language saw its transformation into an optional nominative case suffix under the influence of Kriol argument structure (Meakins, 2015). Similarly, this study captures two languages at different stages of developing phonological stratification, demonstrating how the contrast has developed in individual bilingual speakers and is incrementally propagating through the speaker communities.

1.1 *Gurindji Kriol*

Gurindji Kriol is a mixed language spoken in the Victoria River District of northern Australia that is located 470 kilometres from the nearest town of Katherine. It emerged around 40 years ago and is now spoken by Gurindji

people in the Aboriginal communities¹ of Daguragu and Kalkaringi, and by Bilinarra and Ngarinyman people in two communities north of Kalkaringi—Pigeon Hole and Yarralin.

Gurindji Kriol originates in Gurindji (Ngumpin Yapa, Pama-Nyungan), the traditional Australian language of the region, and Kriol, the English-lexifier creole language spoken across much of northern Australia. It combines the lexicon and structure of these two languages. The structural mix of Gurindji Kriol is well documented, with Gurindji providing much of the noun phrase system and Kriol contributing the verb phrase system (e.g., Meakins, 2011). This type of mixed language is referred to as a V(erb)-N(oun) mixed language and includes Michif and Light Warlpiri (Matras and Bakker, 2003; Meakins, 2013b). The lexicon of Gurindji Kriol is also highly mixed. Based on a 200 word Swadesh list, 36.6% of vocabulary is derived from Kriol and 35% finds its origins in Gurindji. The remaining 28.4% contains synonymous forms from both languages (Meakins, 2011: 19). The extent of lexical mixing is shown in (1) below where Gurindji forms are given in italics and Kriol forms in plain font.

(1) (Meakins, 2011: 18)

(a) *Karu* *i=m* *top* *la=i*m *kankula* *diya-ngka*.
 child 3SG.S=PRS be OBL=3SG.O up deer-LOC
 ‘The child is perched on top of the deer.’

(b) *I=m* *teik-im* *rarra*j *dat* *karu-ma* *nyanuny*.
 3SG.S=PRS take-TR run the Child-TOP 3SG.DAT

ngarlaka-ngka *an* *warlaku* *kanyjurra-ngka*.
 head-LOC and dog down-LOC
 ‘The deer takes the child running on its head, with the dog below.’

(c) *Dat* *diya-ngku* *i* *bin jak* *im* *na* *karu*.
 the deer-NOM 3SG.S PST fall 3SG.O FOC child
 and *warlaku* *kanyjurra-k* *klif-nginyi-ma*.
 and dog down-ALL cliff-ABL-TOP
 ‘The deer threw the child and the dog downwards off the cliff.’

¹ Aboriginal communities in Australia are similar to many Indian reservations/reserves in the United States/ Canada, in that the majority of residents are Indigenous.

- (d) Tubala baldan *kujarrap-pa-rni karu an warlaku ngawa-ngkirri jirpu.*
 3DU fall pair-PA-ONLY child and dog water-ALL dive
 ‘The pair of them, the child and dog fell down, plummeting into the water.’

Gurindji Kriol now has around 700 speakers. It is the main language spoken and acquired at Kalkaringi. Gurindji is still spoken by people over the age of 40 years, albeit generally code-switched with Kriol. All Gurindji people speak Kriol to varying extents when they visit Kriol-speaking areas to the north, for example Katherine and Timber Creek, but do not speak it at home. Standard Australian English is the language of the school despite the fact that children enter school with no background in English. English is also the language of the media and government services but it plays little role in people’s home lives (Meakins, 2008: 287–295).

Gurindji Kriol originated from contact between non-Indigenous colonists and the Gurindji people. In the early 1900s, white pastoralists set up cattle stations in the Victoria River District area, including on the homelands of the Gurindji. Many Gurindji people were killed in skirmishes over land, and the remaining people were put to work on Wave Hill Station in the early 1900s as stockmen and kitchen hands in slave-like conditions together with other Aboriginal groups such as the Bilinarra and Ngarinyman. In 1966 the Gurindji initiated a workers’ strike to protest against the poor working and living conditions and to ultimately regain control of their traditional lands. Today the Gurindji continue to live on their traditional lands at Kalkaringi (Charola and Meakins 2016).

The linguistic practices of the Gurindji are closely tied to these social circumstances. The establishment of the cattle stations by colonisers saw the introduction of the cattle station pidgin (the basis of Kriol) into the linguistic repertoire of the Gurindji. Code-switching was a common practice and it is likely that it provided a fertile ground for the formation of the mixed language (McConvell and Meakins, 2005; Meakins, 2011, 2012). The shift to a mixed language rather than monolingual Kriol was probably the result of the fact that Kalkaringi had only one dominant language (with other languages present such as Bilinarra and Ngarinyman mostly mutually intelligible) rather than many disparate languages spoken in one community that is a characteristic of Kriol-speaking communities (see below). English has had little foothold in the community, perhaps due to its late introduction. It is not entirely clear when a school was established in Kalkaringi but probably not before the 1960s. Most access to English before then was in the limited communication Gurindji

people had with station people who, in any case, mostly addressed Gurindji people using the cattle pidgin.

1.2 *Kriol*

Kriol is an English-lexifier creole language and the first language of most Aboriginal people across the Top End of Australia with the exception of northern Arnhem Land and the Daly River region (Munro, 2000; Sandefur and Harris, 1986). Kriol-speaking communities include Ngukurr (where Roper River Kriol, the variety discussed in this paper, originated), Beswick, Barunga, Bulman, Katherine, Timber Creek, Bulla and Amanbidji (Fig. 1). Kriol is now the main language of these communities, with traditional Australian languages rarely used except by the oldest generations. English is the second most widely used language in most of these places, although is only learnt when children enter school. Like Kalkaringi, all education and government services are provided in English.

Structurally, Kriol is an isolating language with little bound morphology, for example core arguments are differentiated using word order or marked by prepositions. Similarly, tense, mood, and aspect (TAM) categories are expressed through auxiliary verbs rather than inflections (Sandefur, 1979). The lexicon of Kriol is almost entirely derived from English, with a small amount of vocabulary maintained from surrounding substrate languages, in particular Marra (Dickson, 2016). Some examples are given below.

- (2) Main mami bin oldei gemp langa gemp.
 1SG.POSS Mother PST CONT live PREP camp
 'My mum always stayed at home.' (Munro 2005: 87)
- (3) Det boi im-in gaj-im-ap-bat dat biginini.
 the boy 3SG-PST carry-TR-up-CONT the child.
 'The boy was carrying the child.'
- (4) Olabat wandi gaman burrum Darwin.
 3PL Want come ABL Darwin
 'They want to come from Darwin.' (Sandefur 1979: 148)
- (5) Langa gud-wan kantri blanga im dadi.
 LOC good-NMLZ country POSS 3SG father
 'In the good country which belongs to his father.' (Sandefur, 1979: 143)

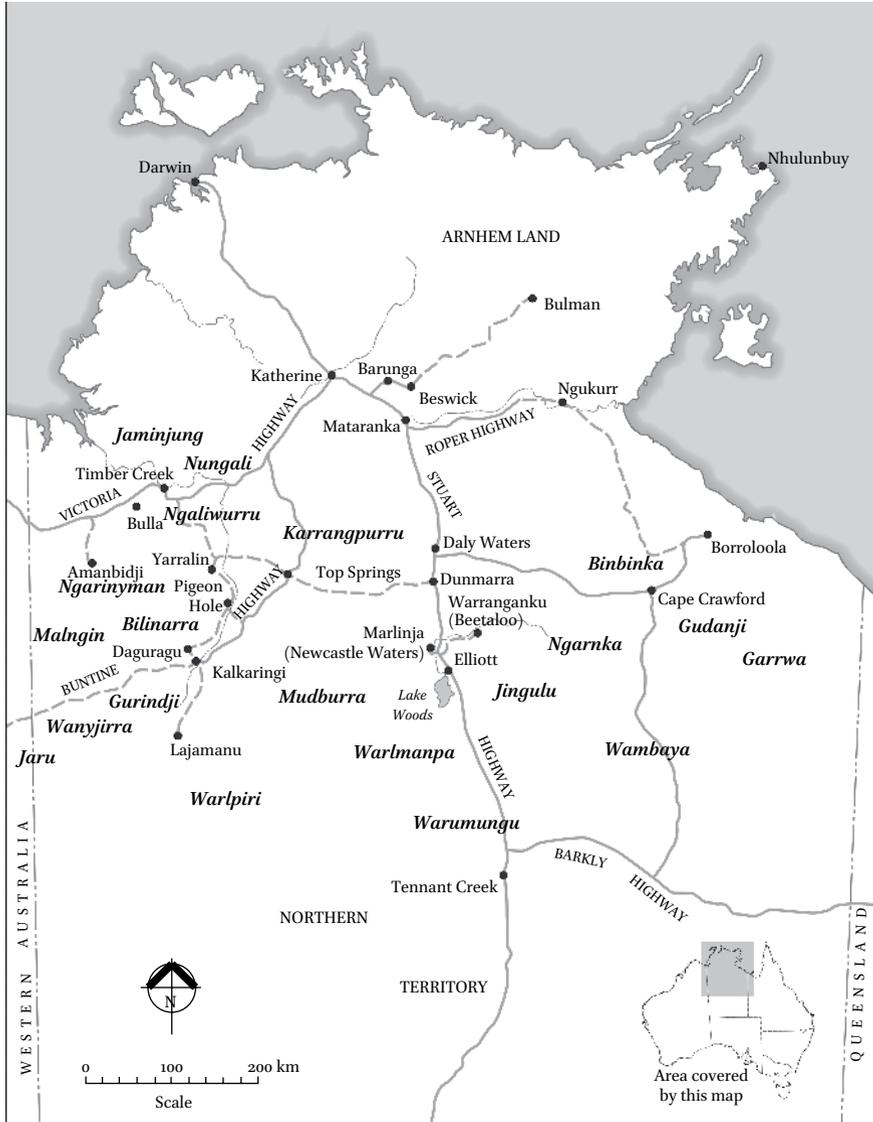


FIGURE 1 *Towns and Aboriginal communities in northern Australia*

Kriol originated in NSW Pidgin and spread north to Queensland and the Northern Territory in the early 1900s through the pastoral industry (via Aboriginal labour imported from Queensland) and nativised in different places (Meakins, 2014; Sandefur and Harris, 1986; Simpson, 2000 for an overview). One of the

earliest varieties of this cattle station pidgin to nativise was Roper River Kriol at Roper River Mission (now Ngukurr) in the early 1900s. Roper River Mission was established as a refuge for Aboriginal people from nine different language groups including Alawa, Marra, Warndarrang, Ngalakgan, and Ngandi who were escaping massacres. Most Aboriginal people were fluent in two or more of these languages. In addition, they would have spoken the pidgin English that arose from interaction with the colonists at least 30 years prior to the establishment of the mission. For many Aboriginal people at the mission, the cattle station pidgin became their *lingua franca*, with traditional languages reserved for in-group communication. The mission also separated children from their parents so a combination of community-level multilingualism and lack of access to traditional languages most likely contributed to the formation of Kriol (rather than a mixed language, as was the case for Gurindji Kriol). The presence of English was also strongly felt in the mission with children taught in English right from its establishment in the early 1900s (Harris, 1986). In this respect, Ngukurr is a community that has around 50 years and two generations more contact with English than Kalkaringi where Gurindji Kriol developed.

1.3 *Stop Consonants*

It has been shown that listeners weight relevant cues encoded in the speech stream to identify contrasts (Lisker, 1986; Scobbie, 1988). Some cues are given priority over others and experiments involving the removal of specific cues (e.g., vowel duration vs. spectral cues in English /i/ vs. /ɪ/ (Escudero, 2000)) can reveal the importance or weight of such cues. Escudero (2000) reveals that spectral cues in the tense/lax high front vowel pair in English take priority over duration.

When languages have a distinction between stop consonants in the same place of articulation, one of the primary cues used to distinguish such categories involves voice onset time (VOT). This cue refers to the temporal duration from the moment of release of the closure to the onset of voicing in the following vowel (Lisker and Abramson, 1964). When a stop series is contrastive, it often conforms to one of three patterns: voiced, voiceless unaspirated, and aspirated (Keating, 1984). While the differences in duration are language specific, voiceless aspirated stops ([p^h, t^h, k^h]), like those found in word-initial position in English, Australian Kriol, and Gurindji Kriol, are shown to have overall longer durations compared to voiceless unaspirated stops ([p, t, k]). The VOT of stop consonants, like those found in French dialects (Caramazza and Yeni-Komshian, 1974; Hoonhorst et al., 2009), can also be negative, meaning vocal fold vibration begins before release. English contrasts between aspirated and unaspirated stops and speakers interpret the latter, both phonemically and orthographically, as <b, d, g> although they are not true voiced stops in the since voicing begins post-release.

Other secondary cues thought to be involved in stop production and perception include pitch (F_0) depression after voiced stops (Abramson and Lisker, 1973). This can be observed as a decrease in the fundamental frequency right after release. Another secondary cue involves the loss of the initial transition of the first formant (F_1) in vowels following a voiceless stop (known as F_1 cut-back) (Liberman, Delattre, and Cooper, 1958; Lisker and Abramson, 1964). The duration of the post-stop vowel has also been shown to correlate with stop voicing contrasts (Miller and Dexter, 1988; Summerfield, 1981).

1.4 *Stop Production under Contact*

While a substantial number of studies investigate the effects of bilingualism on VOT values compared to those of monolinguals (MacLeod and Stoel-Gammon, 2005 for French-English; Delano, 2012 for Spanish-Creole English; Flege, 1991 for Spanish-English; Kehoe, Lleó, and Rakow, 2004 for Spanish-German; *inter alia*), studies that examine sound production in lexical borrowings in monolingual speech are only now emerging. Those described here all come from the mixed language or Kriol literature and all suggest that phonology, like the lexicon and grammar of a language, also does not conform to any clear systematic paradigmatic patterns in situations of borrowing but rather variation is commonplace, perhaps as an intermediate step in the development of a system.

Specifically related to this study, Jones and Meakins (2013) look at VOT production in Gurindji Kriol and Northern Australia English. Unlike English, traditional Gurindji does not have a voicing contrast in the stop series, which consists of [p, t, c, k]. One particularly relevant finding to this study describes VOT variation in Kriol-derived and Gurindji-derived words produced by adult speakers of Gurindji Kriol. Here, they tested whether the values systematically relate to those in English cognates. Based on data gathered from a picture naming task and natural speech, their results show that there is little effect of English voicing in Gurindji Kriol among words of Kriol or Gurindji origin in word-initial position, although there is some degree of variability (Jones and Meakins, 2013: 216).

These findings raise the questions: How are stops categorically perceived in Gurindji Kriol and is there any variation based on age or exposure to Australian English? And how do their results compare with those of Kriol? Based on impressionistic data, Kriol has been described as not having a stop voicing contrast, at least not in basilectal varieties, in existing published literature (Hudson, 1985; Munro, 2004; Sandefur, 1979) as well as in recent surveys (Butcher, 2008; Schultze-Berndt, Meakins, and Angelo, 2013). However, Bundgaard-Nielsen and Baker (2016) and Baker, Bundgaard-Nielsen, and Graetzer (2014), show that second and third generations of monolingual Roper

Kriol speakers both produce and perceive stop-voicing contrasts ([p-b, t-d, k-g]) while first generation speakers show variability. For Gurindji Kriol, Jones and Meakins (2013) show that Gurindji Kriol speakers tend to assimilate any form of stop voicing perceptually to that of Gurindji's phonological system though there is some degree of variation. What makes this situation worthy of further investigation, however, is the fact that variation between the voiced and voiceless series shows that speakers are at least able to make the correct articulatory gestures needed to produce such sounds. This means speakers might be able to take advantage of such variability perceptually when needed (e.g., under ambiguous conditions such as contrasting minimal pairs out of context e.g., *boring/poring* in the phrase *Nyantuma i bin tok im rili poring/boring* 'She said it's really pouring/boring').

Because Kriol, as it is spoken at Ngukurr, developed earlier than Gurindji Kriol and has been in contact with English (which has a clear stop voicing contrast) for longer and more intensively through an extended period of schooling, we might expect Kriol listener perception to be more contrastive than their Gurindji Kriol counterparts. Through constant modern day contact with English, however, both languages may be adopting the stop voicing contrast—Kriol in all parts of speech and the Kriol origin lexicon in Gurindji-Kriol (e.g., *pak* and *bak* from English 'park' and 'bark' may be perceived as distinct instead of both defaulting to homonym *pak*). If the adoption process was merely for sociolinguistic reasons, we would expect a quicker diffusion of the contrast as speakers would be made consciously aware of the difference. However, an incremental and variable change may signify the structure of the language is benefiting from adopting the contrast (e.g., reducing functional load of the voiceless series that might level out phoneme frequency and distribution allowing for a greater number of contrasts leading to greater phonological optimization (Surendran and Niyogi, 2006; Wedel, Kaplan, and Jackson, 2013)). Regarding perception, there are four primary outcomes that will reveal how stop consonants are categorized in the phonology of these languages: (1) the voiced series assimilates to the voiceless series, (2) both series are perceptually contrastive, (3) both series exist in free variation, and (4) the voiceless series is established while the voiced series is in flux.

Bundgaard-Nielsen and Baker (2016) show that for elicited stops from three Roper Kriol speakers, there is a clear contrast between voiced and voiceless stop production in the English origin lexicon.² For spontaneous speech data from a single speaker, there also appears to be a contrast, though their results

2 It should be noted that due to the limited number of participants tested in these studies, results may be more variable.

are non-significant; a result they claim is due to the small number of tokens. Moreover, they also show variability in stop voicing production in a Kriol dominant Wubuy L1 speaker that suggests Wubuy speakers make use of a single stop category regarding voicing. With respect to perception, Bundgaard-Nielsen and Baker (2015) showed that Wubuy listeners had a difficult time discriminating between both English and Kriol labial stops that differed in VOT duration; a result they attribute to the lack of native experience in dealing with the voicing contrast. The adoption of the stop voicing contrast by Kriol speakers might be expected before that in Gurindji Kriol since the functional load of the contrast would affect the entire Kriol lexicon rather than just the Gurindji Kriol verb phrase elements.

In *Media Lengua*, a lexicon-grammar (LG) mixed language (Matras and Bakker, 2003; Meakins, 2013b) spoken in Ecuador, with Imbabura Quichua systemic elements and an Ecuadorian Rural Spanish-derived lexicon, Stewart (2015) showed the Spanish voiced stop series has been adopted, both productively and perceptually, by Quichua³ and *Media Lengua* speakers with varying ages and levels of Spanish proficiency. The VOT values of these adopted stops, however, are longer in duration than their original Spanish counterparts suggesting some degree of overshoot during acquisition. For the Quichua speakers, a significant number of stops also undergo variable weakening to [β, ð, γ]. Stewart (forthcoming, 2014) also claims a similar tendency for Spanish-derived vowels in both Quichua and *Media Lengua*.

Based on the differences in formation between these two mixed languages (code-switching in Gurindji Kriol (McConvell and Meakins, 2005)) versus relexification in *Media Lengua* (Muysken, 1981) and the type of splits (50/50 Gurindji and Kriol lexicon in Gurindji Kriol (Meakins, 2011: 11)) versus 10/90 Quichua and Spanish lexicon respectively in *Media Lengua* (Muysken, 1997), the amount of 'weight' placed on the phonological system in *Media Lengua* by Spanish may have been large enough to warrant adopting the series; while this might not have been the case in Gurindji Kriol. To illustrate this point, in Michif, which, like Gurindji Kriol, is a (V)erb-(N)oun mixed language (Bakker, 2003: 122; Meakins, 2013a: 179), with Cree-derived verb phrases and French-derived noun phrases, Rosen, Stewart, and Cox (2016) show that speakers have actually only adopted a small number of French vowels while the rest assimilate to their Cree counterparts.

It should be noted that there have been attempts to systematically categorize these phonological processes. Van Gijn (2009) provided an in-depth analysis suggesting that mixed languages borrow phonological material based on

3 With the exception of older Quichua speakers.

type of lexical and grammatical material they adopt. Here, a language with a lexical-grammar split, where the lexicon of one language and the grammar from another combine to make a new language (e.g., *Media Lengua*, *Ma'a*), should share lower level material such as individual segments since phrases are more likely to be made of individual linguistic parts from each language. On the other hand, noun-verb mixed languages, which borrow lexical items categorically (e.g., *Gurindji-Kriol*, *Michif*) should maintain language-specific phonological material at levels higher than the segment since entire phrases may be of a single source language. Recent studies referenced above that explore the phonetic properties of these sound systems, however, paint a more complex picture involving mergers, near-mergers, segments with substantial overlap in acoustic space, and category maintenance. While some of these patterns align with Van Gijn's (2009) analysis, the degree of alignment can seem peculiar (e.g., vowel spaces with such a high degree of overlap that they would seem to have little perceptual benefits to listeners). Other patterns (e.g., the number of actual French vowels borrowed in *Michif*), do not align with Van Gijn's hypotheses.

Turning briefly to the bilingual literature, Pasquale (2005) revealed that when speaking Quechua, Quechua-Spanish bilinguals dominant in Quechua produced overall shorter *vOT* values than Quechua monolinguals; values that trended towards Spanish-like production. Spanish-dominant bilinguals, on the other hand, showed no noticeable shift toward Quechua-like *vOT* production when speaking Quechua. MacLeod and Stoel-Gammon (2005) suggest simultaneous French-English bilinguals produce *vOT* with French monolingual-like values, which also carried over into their English *vOT* production. Flege (1991) shows that Spanish-English late bilinguals produced the *vOT* values of /t/ in between those of standard monolingual Spanish unaspirated values and monolingual English aspirated values. On the other hand, early bilinguals (Spanish L1, English L2) produced *vOT* values that matched those of English monolinguals. These findings suggest that, for the most part, simultaneous and early bilinguals typically maintain separate L1 and L2 *vOT* values while late bilinguals usually do not reach native-like *vOT* production in their L2. Similarly, Chang, Yao, Haynes, and Rhodes (2011) show that the younger a heritage speaker is when exposed to both languages, the more successful they will be at maintaining distinctions within and across their languages.

Beyond this clear effect of age of acquisition, studies also show that language exposure (use, length of residence, practise etc.) is also a relevant factor in sound production and perception. Flege, Takagi, and Mann (1996) show Japanese speakers living in the US for 21+ years were able to identify liquids with higher consistency compared to Japanese speakers who only lived there for

2 years. About 10% of the improvement in the production of the English [eɪ] diphthong by Italian speakers (with native [e]) could be attributed to the frequency of a speaker's L2 usage; suggesting that practise can improve production in adult speakers (Flege, Schirru, and MacKay, 2003). At the same time, Flege and Liu (2001) conclude that for adults, length of residency is not enough to improve L2 performance. Instead, improvement is only measurable if a speaker receives constant input from L1 speakers. Finally, Klein (2013) shows that French and Mandarin L1 speakers with a substantial length of residency in an English-speaking area tend to produce more native-like English voiced stops. These findings might be applied to the Gurindji Kriol and Kriol context as a way to understand the roles of age and length of exposure to English in the formation of these languages.

1.5 *Categorical Perception*

Since Liberman (1957) researchers have been aware that humans (and later other animals (Kuhl and Iverson, 1995; Kuhl and Miller, 1979)) perceive individual speech sounds as homophonous-like categories, meaning distinct sounds within a single category are perceived as similar while neighbouring sounds in a separate category are perceived as distinct—even if cross-category sounds are closer in acoustic space. For bilingual listeners, however, the categorization of phonemes is more complex and varies based on age of acquisition of the L2. It is often thought that listeners establish phonemic categories within the first year of life (Kuhl, 2004; Werker and Tees, 1984), yet the organization of such categories for bilinguals has been shown to be distinct from their monolingual counterparts. Caramazza, Yeni-Komshian, Zurif, and Carbone (1973) show that for simultaneous and early bilinguals, a single intermediate boundary in VOT perception was established for both a listener's L1 and L2. Bosch, Costa, and Nuria (1997) however, show that L1 phonemic categories of early bilinguals remain essentially unchanged even when exposed to similar categories in the L2. On the other hand, Hazan and Barrett (2000) suggest the refinement of phonemic categories can take place until adolescence. Furthermore, Guion (2003) established that simultaneous bilinguals maintain separate categories even when faced by sounds that have the same phonemic function and articulatory shape across both languages (e.g., Spanish /i/ and Quichua /i/). For early bilinguals, however, these sounds merged while late bilinguals, who typically acquired Spanish under 'unguided' conditions, also merged the Spanish mid-vowels with Quichua high vowels (Quichua being a three vowel system consisting of /i, u, a/, Spanish consisting of /i, u, e, o, a/). Although these studies may differ as to when categories become solidified, it is clear that simultaneous and early bilinguals have distinct categorical

arrangements compared to late bilinguals who rely on their L1 for perceptual cues in both their languages.

When investigating categorical perception of speech sounds two common task-based experiments are often implemented—identification-based and discrimination-based. The first involves identifying sounds as belonging to a given category—often presented in a forced choice format. In such an identification task, modified audio tokens along a continuum between two canonical phonemes might be presented at random and participants would be asked to label the audio stimuli by selecting a corresponding image/text or with a gestural/oral response. Two-alternative forced choice experiments, similar to the one presented in this paper, are considered advantageous for identifying categories for several reasons: (1) they are considered simple tasks for participants to complete, (2) they minimize bias as participants are only given two options for identification where one is known to be correct, (3) the need for distributional assumptions is typically not necessary (McGuire, 2010), and (4) according to Borden, Harris, and Raphael (1994) categorical boundaries can be estimated if the stimuli are contrastive. Pitfalls to this experimental method involve asking participants to sit through a lengthy experiment with a large number of trials that may become monotonous. The stimuli also need to be explicitly defined for the participants that may require a brief training session.

While not used in this study, it is worth briefly mentioning discrimination task-based experiments that are also often implemented for labelling categories perceptually (see e.g., Bundgaard-Nielsen and Baker, 2015). One common construction involves an AX design where participants are asked to label two audio samples as *same* or *different*. The benefits of such a discrimination task experiment of this nature involve (1) a smaller number of trials compared to identification tasks and (2) the ability to pin point categorical boundaries. Two main disadvantages of this experiment type involve a bias towards the *same* response when pairs are more difficult to contrast. The second involves a substantial amount of trial rejections with the *same* response due to their uninterpretability (McGuire, 2010).

2 Methodology

This section details our two-alternative forced-choice (2AFC) identification task experiments (Section 2.1) including the stimuli used, how the continua were designed (Section 2.1.1), and presented (Section 2.1.2). Moreover, this section provides demographic information provided by the participants (Section 2.2) and the procedures used to implement the experiments (Section 2.3).

2.1 2AFC Identification Tasks

The primary 2AFC identification task used in this experiment was designed to look for stop voicing contrasts based on the intuitions of native speakers of both Gurindji Kriol and the Roper dialect of Kriol. This task-based experiment made use of Kriol lexical borrowings that also make up part of the Gurindji Kriol lexicon. In addition, a simplified version of the experiment was designed to test the intuitions (devised from a listener's native experience with their L1 phonology) of Gurindji Kriol and Kriol speaking children typically younger than 8 years of age and Gurindji and Kriol older adults typically over the age of 65 (see Section 2.2 for further details). The goal of this experiment was not to seek out categorical boundaries, but rather to simply learn whether listeners in these age ranges could identify voicing contrasts in word-initial stops. Going forward, we refer to these 2AFC identification task experiments as the 'standard experiment' (for preteens to middle-aged adult participants) and the 'simplified experiment' for participants with limited exposure to Standard English (young children under the age of approximately 8 and older adults over the age of approximately 65).

2.1.1 Stimuli

To gather stop perception data for our standard experiment, we used seven word-initial minimal pairs that contrasted in stop voicing quality in word initial position in Kriol (e.g., *pak* 'park' and *bak* 'bark'). Each minimal pair has its origins in Kriol and has made its way to the Gurindji Kriol lexicon where it maintains a nearly identical phonological shape. Table 1 presents the stimuli used in our word identification task. Both the [t-d] and [k-g] series contain two minimal pairs each, while the [p-b] series contains three. For the simplified experiment, the minimal pairs: *traiyimat-draiymat* and *katim-gatim*, were removed.

TABLE 1 *Minimal pair words used in the standard experiment.*

[p-b]	[t-d]	[k-g]
<i>pak</i> [pɔ:] 'park'	<i>tai</i> [tɔ:] 'tie'	<i>kol</i> [kol] 'coal'
<i>bak</i> [bɔ:] 'bark'	<i>dai</i> [dɔ:] 'dye'	<i>gol</i> [gol] 'gold'
<i>pai</i> [pɔ:] 'pie'	<i>traiyimat</i> ['tɔɪ.mat] 'try'	<i>katim</i> ['ka.rəm] 'cut'
<i>bai</i> [bɔ:] 'baby sleep'	<i>draiymat</i> ['dɔɪ.mat] 'dry'	<i>garram</i> ['ga.rəm] 'have'
<i>poring</i> [pɔɪɪŋ] 'pouring'		
<i>boring</i> [bɔɪɪŋ] 'boring'		

Instead of using synthetic audio tokens for the stimuli, we chose to modify natural speech tokens to minimize issues with quality that have been attributed to synthetic speech (Vainio, Järvikivi, Werner, Volk, and Välikangas, 2002). For both experiments, one of the authors, a female speaker of Roper River Kriol from Ngukurr (a Kriol-English bilingual) with a clear stop voicing contrast produced the minimal pairs in Table 1. An Editor Rog portable digital recorder with a Sony lapel mic (40–20,000 Hz response) was used to record the stimuli with a sample rate of 44.1 kHz. After rendering the recordings to 16-bit stereo WAV format, we manually modified several primary and secondary acoustic cues known to carry weight in stop perception. These included, the voice onset time of each word-initial stop, formant transitions at the onset of voicing, and the overall pitch and duration of the following vowel. The removal of aspiration during the modification of the VOT took place immediately following the release burst to insure consistency in the resulting token stimuli. We then combined any remaining portion of the original voiced minimal pair token to create a more naturalistic sound sample. Each continuum was tested for naturalness before it was integrated into the experiment.

For the standard experiment, we chose to modify the sound tokens along a 10-step continuum that transitioned gradually from the word-initial voiceless stop minimal pair to its word-initial voiced stop counterpart to cover a reasonably large range of token samples. As the modified values become more distant from their prototypical form along the continuum, we hypothesized that if a participant does indeed contrast the minimal pairs, identification consistency would be reduced. On the other hand, if a participant perceived the minimal pairs as the same, we would expect random responses throughout the continua. If, however, only one token appears phonemically in a participant's inventory, we hypothesize she/he will accurately identify tokens at one end of the continuum while at the other responses would be more randomized. If there is indeed a voicing contrast found, using the 10-step continua will allow us to hone in on the categorical boundaries of each minimal pair.

For the simplified experiment, we opted to do away with the continua and only use the canonical tokens of each minimal pair (i.e., those used at step 1 and step 10 in the standard experiment). For the standard experiment, all the aforementioned modified values were evenly spaced along the continuum as per the original values from the voiceless and voiced stop tokens. All modifications took place using the open source programs Praat version 6.0.8 (Boersma and Weenink, 1996) and Audacity 2.1.0 (Ash, Chinen, Dannenberg, Johnson, and Martyn, 2012). Praat scripts to help automate portions of the token modification process were written by the authors. A sample of the resulting values for the *poring-boring* 'pouring-boring' continuum are provided in Table 2. The

TABLE 2 *vOT duration and mean post-vowel formant, pitch, and duration values of pak-bak continuum*

Acoustic Cues	Poring		Continuum steps						Boring	
	1	2	3	4	5	6	7	8	9	10
VOT (ms)	65	59	53	48	42	36	30	25	19	11
1 st F1 point (Hz)	924	900	876	851	826	802	777	752	728	703
1 st F2 point (Hz)	1779	1695	1612	1528	1445	1361	1278	1194	1111	1027
1 st F3 point (Hz)	2790	2790	2790	2791	2791	2791	2791	2792	2792	2792
Pitch (Hz)	151	149	147	145	143	141	139	137	135	133
Vowel Duration (ms)	42.7	42.8	42.9	43.1	43.2	43.3	43.5	43.6	43.7	43.9

poring-boring values of the simplified experiment are the same as those in step 1 and step 10 in Table 2.

2.1.2 Presentation

In this section we describe the user interface of the experiment. For the standard experiment, the 10 tokens of each minimal pair, described in Section 2.1.1, were placed in a Microsoft PowerPoint presentation along with images corresponding to each minimal pair (Fig. 2).

To attain more accurate results, we designed the presentation for the standard experiment to have more repeats of more distant stimuli from the canonical forms. The resulting *vOT* production values from Jones and Meakins (2013) provided an additional basis for determining that tokens should be repeated (Table 3). Therefore, the participants listened to the same minimal pair series along the continuum 16 to 17 times for a total of 115 token samples when considering all seven minimal pairs. Table 3 provides an example of the repeated tokens along the continuum.

For both experiments, we configured the PowerPoint presentation to play each token 50 ms after each new slide appeared on the screen. Participants were given the option to repeat the audio sample if they so desired. The participants could also take as much time as they saw fit to respond to the stimuli as there was no form of time pressure. The presentation was configured to use 'Kiosk' mode, which restricted where the participant could click on the screen to move to the following slide. By doing so, the participant had to click one of the two images that avoided accidental clicks on the surrounding areas that would otherwise not record their response. Each image was scripted

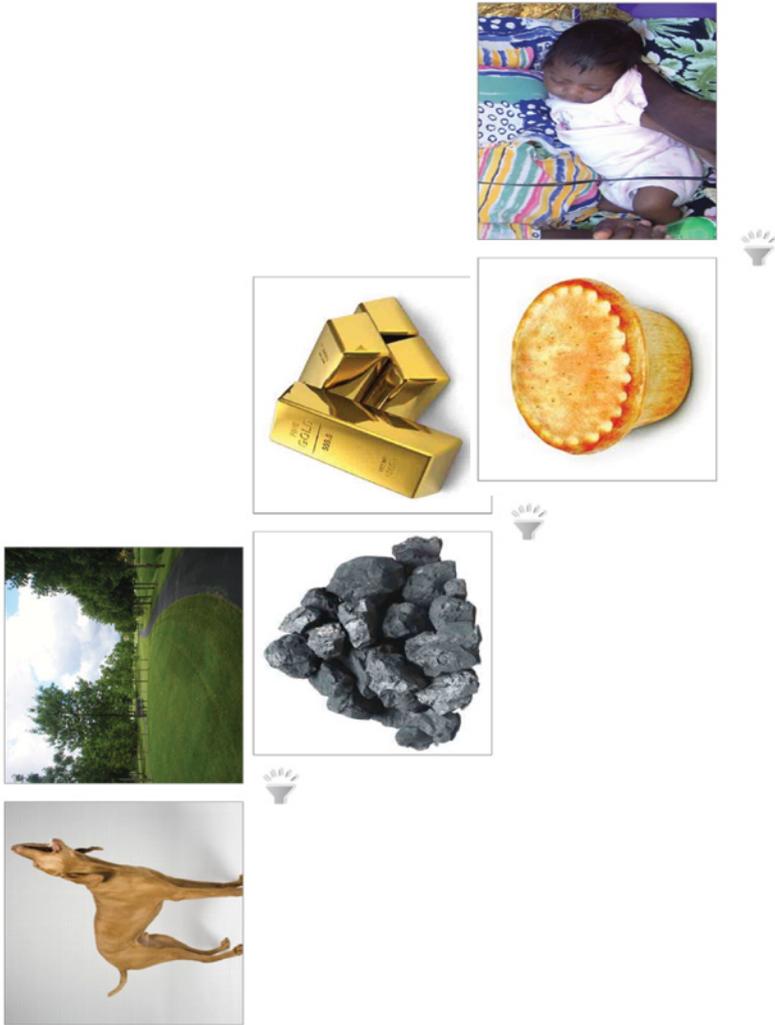


FIGURE 2 Three slides from the identification task. The top left shows the minimal pairs bak 'bark' and pak 'park'. The middle slide shows the minimal pairs kol 'coal' and gol 'gold'. The bottom right slide shows the minimal pairs pai 'pie' and bai 'baby sleep'.

TABLE 3 *Example of the repeated steps along the continua for each minimal pair.*

Minimal pairs	Number of repeats per continuum step										Total # of tokens
	1	2	3	4	5	6	7	8	9	10	
bak-pak	1	1	2	3	3	2	1	1	1	1	16
boring-poring	2	3	3	2	2	1	1	1	1	1	17
bai-pai	1	1	2	3	3	2	1	1	1	1	16
gol-kol	1	1	1	1	1	2	3	3	2	2	17
garram-katim	1	1	1	1	1	2	2	3	3	2	17
dai-tai	1	1	1	2	3	3	2	1	1	1	16
draiyimat-traiyimat	1	1	2	3	3	2	1	1	1	1	16

using the Visual Basic for Applications (VBA) add-on in PowerPoint to record the participant's individual response for each slide. To avoid any type of pattern recognition in the data the slides were reordered using a randomization macro. The slides were then further adjusted to make sure no two contained the same images one after the other. At the beginning of the experiment two trail tokens were presented to introduce the participants to the experiment. Finally, for the standard experiment, one slide containing an audio sample from step 10 was placed at the beginning of the presentation. This provided the participants with a canonical form to get their bearings before being presented with non-canonical forms at random and all the participants heard the stimuli in the same randomised order. For both experiments, distractor tokens involving stop-fricative minimal pairs, produce by the same speaker, were added to the experiment to reduce the constant repetition of the same seven minimal pairs. All instructions were given in the participants' L₁ (Gurindji Kriol and Kriol respectively). At the end of the experiment a text file was created containing all the participant's responses and demographic information collected on the first slide.

2.2 *Participants*

Table 4 provides demographic information provided by all 103 participants who took part in the experiments. These data include numbers and percentages for: our groupings based on language and age, number of participants per group, each group's mean age and one standard deviation of the group's age, the group's gender distribution, and their level of exposure to Standard English and a traditional Australian language (Gurindji in the case of the Gurindji Kriol

TABLE 4 *This table provides demographic information provided by the participants who took part in both experiments. Data here includes: our groupings based on language and age, the number of participants per group, the group's mean age and one standard deviation of the group's age, gender distribution in each group, and the participant's exposure to both Standard English and a traditional language (inc. Gurindji): H(igh) and L(ow).*

Language and Age Groups	Participants (Count)	Age		Gender		Exposure to English		Exposure to Traditional	
		Mean	1 SD	F	M	H	L	H	L
Gurindji Kriol (G I)	11	5	0.75	3	8	0	11	0	11
Kriol (G I)	13	6	3	12	1	4	9	0	13
Gurindji Kriol (G II)	48	32	18	41	7	42	6	30	18
Kriol (G II)	18	36	11	14	4	17	1	3	15
Gurindji (G III)	5	70	18	5	0	0	5	5	0
Kriol (G III)	8	62	11	6	2	2	6	8	0
Totals	103			81	22	65	38	46	57

participants). Henceforth the participant's age will be referred to by groups, with the children forming group I (ages 8+), the preteens to middle-aged adults forming group II (ages 10–58), and the older adults forming group III (65+).

It should be mentioned that the simple experiment was required because the length and complexity of the standard experiment was not ideal for participants in group I and group III. Moreover, these participants differed from those in group II as they had substantially less exposure to Standard English through formal education, which may make a difference in the acquisition of the stop voicing contrast as it is present in English. By including the simplified experiment, we could analyse stop voicing perception in six distinct groups of participants differing in their use and exposure to a traditional Australian language and English, but who all currently live and interact in the same language ecology.

2.3 Procedure

For both experiments, the participants were told that they would hear several words and their task was to choose the image that corresponded to the word they heard. The participants were also told that if they would like to hear the

audio sample again, they could click on the speaker icon at the bottom of the screen. We urged them, however, to go with their first instinct. There was no time restriction or time pressure placed on the participants. They were also told that the words would be repeated many times and that some of them might be harder than others to understand but to try their best. Before beginning the experiments, we reviewed the minimal pairs with each participant with a printout of the picture pairs. This was to help avoid any confusion matching the images with the token samples during the task. For the standard experiment, the participants were told the entire task would last about 15–20 minutes and there were no right or wrong answers. For the simplified experiment, participants were told the entire task would last about 5 minutes. Participants who took part in the experiments were monetarily compensated for their time.

The participants were provided with a PC laptop and noise cancelling headphones for the experiments. For the experiment, we asked them to point at the images they heard in the audio sample and we or our assistant would click the picture for them.

3 Results

Our results section is broken down into two subgroups. The first (Section 3.1) details the standard experiment, while the second (Section 3.2) looks at the simplified version. Each section provides the results with line plots detailing the mean averages of the responses, followed by a statistical analysis of the results.

Based on production studies of stop voicing in both Gurindji Kriol (Jones and Meakins, 2013) and Kriol (Baker et al., 2014), we would expect the Gurindji Kriol participants will not have a strong perceptual contrast between voiced and voiceless stops, while the Kriol participants should, in turn, maintain two separate categories. In addition, if results from the standard experiment differ substantially from the simplified version, it may be possible that exposure to Standard Australian English through formal Western education is influencing group 11's perception. On the other hand, if English exposure is not a defining factor in stop voicing perception, we would expect similar results in the simplified experiment and standard experiment. Similarly, continuing exposure to a traditional Australian language may be predicted to influence results, i.e. older adults at Ngukurr, and older adults and the preteens to middle-aged adults at Kalkaringi have higher levels of exposure to traditional languages that may be expected to affect their ability to make a strong perceptual contrast between voiced and voiceless stops.

To test these hypotheses, we built two generalized linear mixed effects models fit by the Laplace approximation, one for each experiment, to analyse the results from the perceptual experiments described in Section 2. Generalised linear (logistic) regressions allow for the analysis of a discrete dependent variable (e.g., the binary voiced/ voiceless response to the stimuli in these experiments) along with independent variables that include an entire population (e.g., age, gender, exposure to English etc.). The mixed effects version of a generalised linear regression incorporates an additional layer of analysis by including variables whose populations cannot easily be exhausted (e.g., listener or word (where it is impossible to test every word in a language and its variation each time its uttered)). These models help answer two basic questions: (1) is there a difference between Kriol and Gurindji Kriol at the intercept⁴ (i.e., do listener responses to the stimuli differ significantly at the first step of a continuum)? And (2) do the slopes of the curves differ across the continuum by language (i.e., do listener responses to the stimuli deviate significantly across the continuum)? To answer the latter question, the models contain interactions between *continuum* and *language*. These models also look for differences across the age range of the participants and differences based on place of articulation (bilabial, alveolar, and velar).

The mixed effects models were created in R 3.2.1 with the *lmer* function of the *lme4* package (Bates, Maechler, Bolker, and Walker, 2015). Ninety-five percent confidence intervals (CI₉₅) were computed using *confint* function from the *lmerTest* package (Kuznetsova, Brockhoff, and Bojesen, 2014). Each model included *participant* and *word* as random effects. We considered the following predictors (fixed effects) for each model: *continuum* (steps 0–9), *gender* (female, male), *language group* (Gurindji Kriol, Kriol), *place of articulation* (labial, alveolar, velar), *age, exposure to English* (low, medium, high),⁵ *exposure to a traditional language* (low, medium, high) and *word frequency of each word in the minimal pairs* (low, medium, high).⁶ The model was fit using a backward step-wise procedure where non-significant predictors were removed from the

4 In this case by subtracting 1 from each step of the continuum so that the model treats step 1 as the intercept. All line plots are also presented in the same format (0–9).

5 Exposure to English was based on the number of the number of years of schooling where people have the greatest exposure to English (lower primary school = low; mid-high primary school = medium; high school/tertiary = high). This functions as a reliable correlate to judge overall English proficiency. Standard English is often taught under unguided conditions (i.e., explicit instruction in ESL is not part of the curriculum).

6 Frequency thresholds were based on an 80:20hr (59,933 clause) morphologically-tagged corpus of speech from 73 Gurindji Kriol speakers.

model one-by-one based on the closest z-value to zero, until only significant predictors remained. At the same time, the Bayesian Information Criteria (BIC) score was also used to identify the best model to avoid overfit.

This section includes three line plots containing the perceptual trajectories of each language group along the continuum. For additional analysis, line plots broken down by word are also included. This section also includes the results from the model summary of the generalized linear mixed effects model. When a result is significant, we are most interested in the coefficient estimate (β), which is a conservative estimate of the average difference in log-odds (a measurement of probability) response between the predictors in question. For example, a negative log-odd result for *continuum* means the likelihood of a participant choosing a voiceless token decreases x amount per step, while a positive log-result for *language* simply means a given variable (e.g., *alveolar*) was chosen significantly more than another by a specific language group. Because the continuum has voiceless stops on the left and voiced stops on the right, the continuum effect should not be positive if there is indeed any degree of contrast. Fig. 3 with the [p-b] contrast, Fig. 4, with the [t-d] contrast, and Fig. 5, with the [k-g] contrast, all contain a line plots that illustrate the mean trajectories of the responses from each language along the continuum—Gurindji-Kriol (solid line) and Kriol (dashed line).

3.1 *Standard Experiment*

The results in this section detail the participant responses to the standard experiment. This section contains line plots for all three minimal pair responses ([p-b], [t-d], and [k-g]) in addition to the results of the linear mixed effects model.

Fig. 3 suggests that, overall, both the Gurindji Kriol and Kriol participants appear to perceptually contrast [p] from [b]. It is worth mentioning again, however, that while the results are not near-ceiling/near-floor, the consistent negative trend line suggests a contrast is present. Based on the mean average, it appears the Kriol participants had more consistent responses than the Gurindji Kriol participants regarding the voiced-like stop, while the responses to the voiceless-like stop tokens, towards the beginning of the continuum, are nearly identical. Regarding the individual words, all three appear to be contrastive to the participants of both languages even though the Kriol participants showed fewer random responses to the voiced-like series. The least contrastive pair appears to be *poring-boring* ‘pouring-boring’, but for the Kriol participants, the results suggest the optimal point of contrast for this pairing was at step 6. This might suggest that allowing slightly more aspiration is beneficial for identifying [b] (30 ms).

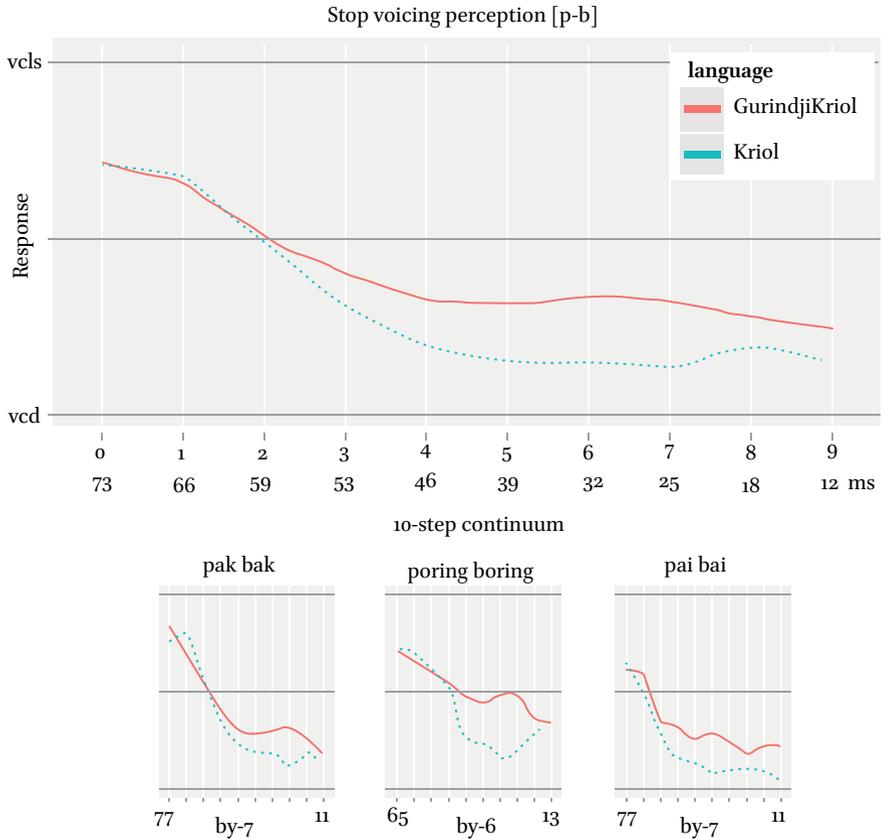


FIGURE 3 The top image represents the mean average of the responses to the [p-b] stimuli across the continuum. The canonical voiceless token is on the left (step 0) while the canonical voiced token is on the right (step 9). The lower images represent the mean average of the responses to each one of the individual words across the continuum (from left to right, pak-bak, boring-poring, pai-bai).

Fig. 4 suggests that, overall, both the Gurindji Kriol and Kriol participants appear to perceptually identify [t], though as the continuum wears on responses to the voiced-like stop stimuli become more and more random. With the *traiyimat-draiymat* pair, responses from the Gurindji Kriol group suggest they overwhelmingly preferred [t] over [d] all the way through the continuum while responses from the Kriol group became more random in the negative VOT range. It should also be noted that this was the only token pair that was produced with a negative VOT by the native Kriol speaker who provided the minimal pairs. Interestingly, pre-voicing had virtually no effect when compared to the *tai-dai* pair.

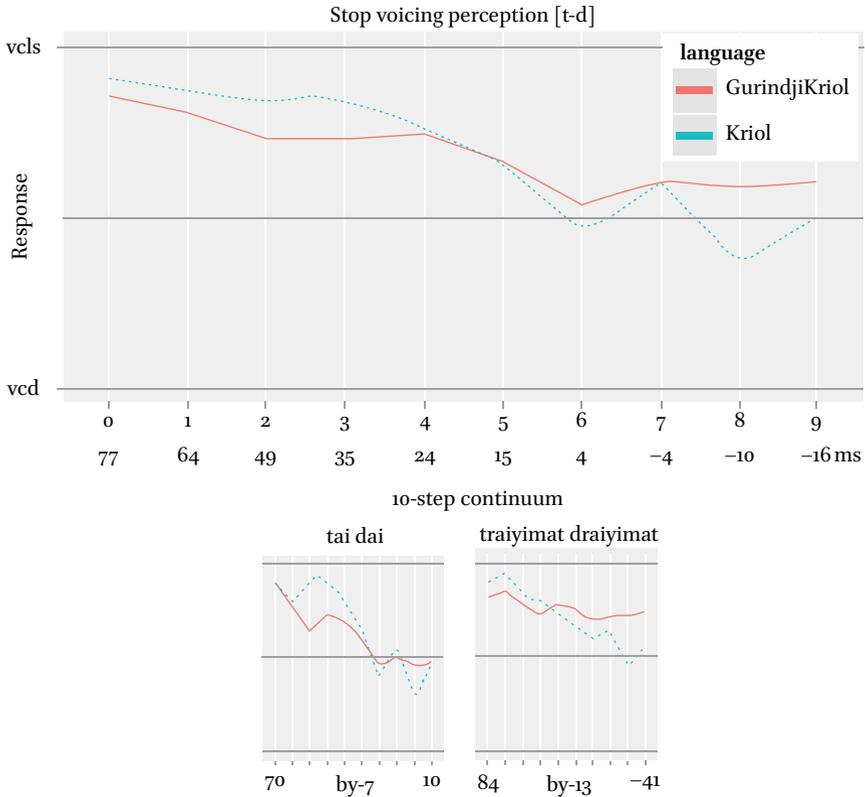


FIGURE 4 *The top image represents the mean average of the responses to the [t-d] stimuli across the continuum. The canonical voiceless token is on the left (step 0) while the canonical voiced token is on the right (step 9). The lower images represent the mean average of the responses to each one of the individual words across the continuum (from left to right, tai-dai, traiyimat-draiyimat).*

Similarly to Fig. 4, Fig. 5 suggests that, overall, both the Gurindji Kriol and Kriol participants appear to perceptually identify the voiceless-like token ([k]) with a high level of consistency. At approximately step 6, consistency begins to steadily decrease until it reaches the 50% mark at approximately step 8. There is little variation in this observation with the individual words.

Table 5 contains the results from the generalized linear mixed effects model using *response* as the dependent variable. Based on the model output, the intercept contains the following five baseline categories: (1) the first step of the continuum, (2) the Kriol participant responses, (3) the alveolar and velar series, (4) stimuli with high word frequency, and the participants who were considered to have a high level of exposure to English.

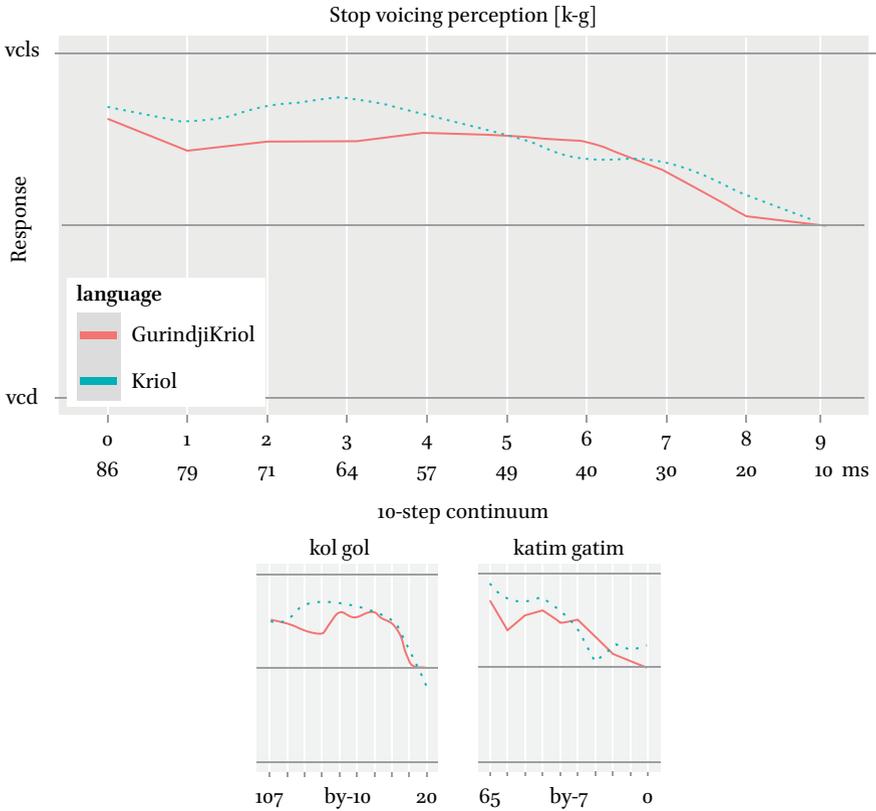


FIGURE 5 The top image represents the mean average of the responses to the [k-g] stimuli across the continuum. The canonical voiceless token is on the left (step 0) while the canonical voiced token is on the right (step 9). The lower images represent the mean average of the responses to each one of the individual words across the continuum (from left to right, tai-dai, traiyimat-draiymat).

TABLE 5 Generalized linear mixed effects model results for stop voicing perception.

	Estimate	Std. Error	2.5%	97.5%	z-value	Pr(> z)	
(Intercept)	2.96	0.24	2.49	3.43	12.34	<2e-16	***
Continuum	-0.36	0.02	-0.41	-0.31	-14.79	<2e-16	***
Gurindji Kriol	-1.31	0.20	-1.69	-0.92	-6.66	2.78E-11	***
Labial	-2.48	0.28	-3.03	-1.93	-8.89	<2e-16	***
Low Exposure to English	0.40	0.15	0.10	0.69	2.61	0.00896	**
Continuum*Lang: GK	0.17	0.027	0.12	0.22	6.34	2.30E-10	***
Lang: GK*Labial	1.10	0.14	0.82	1.37	7.84	4.50E-15	***

The intercept, with a ‘base’ value⁷ of 2.96 log-odds, suggests that Kriol participants, with *high* exposure to English, selected [t] or [k], on average, 95% (19:1 odds) of the time when presented with a canonical voiceless stop. As per the *continuum* predictor result, the probability of selecting a voiceless-like stop ([t, k]) decreased by, on average, -0.36 log-odds per step along the continuum. This suggests that by the final step, Kriol participants only chose the voiceless tokens, on average, 43% (-0.3 log-odds) of the time—indicating a steady decrease in the overall slope across the [t-d] and [k-g] continua and only slight preference towards the canonical voiced tokens ([d] and [g]) at the opposing end of the continua.

As indicated by the significant *labial* predictor, the chances of a Kriol participant selecting [p] over [b] at the first step of the continuum decreased by -2.48 log-odds; reducing the probability to 62% (0.48 log-odds). This decrease, however, suggests that by the end of the [p-b] continua, Kriol speakers were only selecting [p] 6% (-2.8 log-odds) of the time when presented with the canonical [b] tokens. This result provides evidence of a perceptual contrast between labial stops [p] and [b] in Kriol.

For the Gurindji Kriol group, there was a significant interaction with *continuum* (0.17 log-odds). This result suggests that Gurindji Kriol participants have a more moderate decline in their slope compared to Kriol participants, which based on the model output, correlates to a greater preference towards the voiceless series based on response patterns. This result, however, is offset by the significant *Gurindji Kriol* predictor (-1.31 log-odds), making the differences between the languages less extreme along the continua: Here, the model output suggests Gurindji Kriol participants, with a level of *high* exposure to English, selected [t] and [k], on average, 84% (1.66 log-odds) of the time at the first step of the continua, and 49% (-0.04 log-odds) of the time at the final step. Adding in the significant *labial* predictor, the Gurindji Kriol participants selected canonical [p] only 57% (0.27 log-odds) of the time that decreased to 19% (-1.42 log-odds) when presented with canonical [b] tokens. For both language groups, the model results to the [t-d] and [k-g] continua suggest a high level of perceptual consistency for the voiceless-like stops and more randomized responses to the voiced-like stops. These results reflect similar trends found in the mean average line plots shown in Figs. 4 and 5.

Turning to the *English exposure* predictor, with a log-odds value of 0.40, participants who have had less exposure to Standard English in their adult lives had a slight preference towards the voiceless series in all three places of articulation. Those in the Kriol group with *low* exposure to English had a probability

7 All results are displayed to two decimal places, though calculated to the fifth decimal place.

of selecting [t] and [k] 97% (3.4 log-odds) of the time when presented with the voiceless series, and 53% (0.12 log-odds) of the time when presented with the canonical voiced tokens at the last step of the continua. For [p-b], the same group selected [p] 71% (0.87 log-odds) of the time when presented with the canonical voiceless stop that decreased to 9% (-2.4 log-odds) when presented with the canonical voiced stop ([b]). For the Gurindji Kriol participants, those with *low* exposure to English had a probability of selecting [t] and [k] 89% (2.05 log-odds) of the time when presented with the voiceless series, and 59% (0.36 log-odds) of the time when presented with the canonical voiced tokens at the last step of the continua. For [p-b], the Gurindji Kriol participants selected [p] 66% (0.67 log-odds) of the time when presented with the canonical voiceless stop that decreased to 26% (-1.03 log-odds) when presented with the canonical voiced stop ([b]).

Another point of interest is that of the categorical boundary, which is considered the most distant point between both prototypical forms. Honing in on the categorical boundary is typically achieved by identifying the point at which perceptual accuracy is at its lowest (i.e., the data point closest to the 50% mark where responses are the most random) (Abramson and Lisker, 1973; Borden et al., 1994). Based on our results, it is only possible to identify the categorical boundary of the bilabial [p-b] series as it was the only minimal pair series to meaningfully cross the 50% boundary point; according to both the statistical results and Fig. 3. The results from the statistical analysis, shown in Table 6 and Table 7, suggest the categorical boundary for the bilabial series in Kriol (with both *high* and *low* exposure to English) falls in between step 1 and 2 (with a VOT between 66 and 59 ms) while for Gurindji Kriol (with both *high* and *low* exposure to English), it is found between steps 2–3 (with a VOT between 59 and 53 ms). These results are very similar to the mean averages estimated in Fig. 3 that suggest the [p-b] categorical boundaries for both languages fall at step 2 with an average VOT duration of 59 ms.

In summary, while Gurindji Kriol participants had overall more random responses to the alveolar and velar stimuli than the Kriol participants, responses from both language groups revealed greater consistency to the voiceless-like stimuli compared to the more random responses shown in the voiced-like stimuli from the alveolar and velar continua. Contrarily, responses to the labial stimuli revealed most participants from both language groups consistently contrasted voiceless [p] from [b]. Exposure to English also played a slight role in participant response. Here, those with lower exposure had slightly more random responses than those with medium or high exposure, suggesting their L2 plays some role in improving the identification of voiceless vs. voiced stops. Details of individual speaker responses are detailed in the following section.

TABLE 6 *Log-odds, odds, and probability results for the Kriol group (with high exposure to English) along the 10-step continuum according to the intercept (2.96 log-odds), continuum (-0.36 log-odds per step), and labial (-2.48 log-odds) data.*

Continuum step	0	1	2	3	4	5	6	7	8	9
Intercept	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48
Log-odds	0.48	0.12	-0.24	-0.60	-0.96	-1.32	-1.68	-2.03	-2.39	-2.75
Odds	1.61	1.13	0.79	0.55	0.38	0.27	0.19	0.13	0.09	0.06
Probability	62%	53%	44%	35%	28%	21%	16%	12%	8%	6%
p/b (ms. avg.)	73	66	59	53	46	39	32	25	18	12

TABLE 7 *Log-odds, odds, and probability results for the Gurindji-Kriol group (with high exposure to English) along the 10-step continuum according to the following predictors: intercept (2.96 log-odds), continuum (-0.19 log-odds per step (including the interaction)), Gurindji-Kriol (-1.31 log-odds), and labial (-1.38 log-odds (including the interaction)).*

Continuum step	0	1	2	3	4	5	6	7	8	9
Intercept	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87
Log-odds	0.87	0.52	0.16	-0.20	-0.56	-0.92	-1.28	-1.64	-2.00	-2.36
Odds	2.40	1.67	1.17	0.82	0.57	0.40	0.28	0.19	0.14	0.09
Probability	71%	63%	54%	45%	36%	28%	22%	16%	12%	9%
p/b (ms. avg.)	73	66	59	53	46	39	32	25	18	12

3.1.1 Word and Individual Participant Analysis

The *word frequency* predictor overfit the model and was therefore removed in the final version. It is worth noting, however, that the random effects revealed some degree of variation in their intercepts. The results in Table 8 provide the degree of deviance of each minimal pair from the baseline intercept presented in the model (Table 5). A positive log-odds result indicates an increase in the model's overall intercept (baseline value = 2.96 log-odds). Based on our experiment this correlates to an increased number of responses to the voiceless tokens at the first step of the continuum. When calculated into the model, these

TABLE 8 *Random effects from the word variable in the generalized mixed effects model in Table 5.*

Minimal pairs	Intercept
<i>kol-gol</i>	0.19
<i>katim-gatim</i>	0.036
<i>pak-bak</i>	0.008
<i>pai-bai</i>	-0.5
<i>poring-boring</i>	0.51
<i>tai-dai</i>	-0.36
<i>traiyimat-draiyimat</i>	0.128

results translate to a decrease in the response preference at the last step of the continua. This suggests a potential decrease in contrastability in the given minimal pair set. Based on this logic, a negative log-odds response in Table 8 suggests a decrease in response preference to the voiceless series at the first step of the continua, while the response preferences to the voiced series at the end of the continua increase. For example, participants from both language groups had more responses to *poring* ‘pouring’ at the first step of the *poring-boring* continuum compared to *pai* ‘pie’ in the *pai-bai* continuum.

The only minimal pairs with opposing levels of frequency were *pak-bak* ‘park-bark’ (high-low respectively) and *tai-dai* (low-high respectively). The intercepts of both these pairs trend in a way that might suggest frequency might play a role. There are, however, counter examples e.g., high frequency word pairs such as *pai-bai* and *poring-boring* that show opposing trends—a preference towards *poring* (voiceless) but also for *bai* (voiced).

It should be mentioned that when we combine all the minimal pairs into a single average, there was a relatively high degree of variation seen in some of the participant results for both groups. Some are able to contrast minimal pairs in every place of articulation ([p-b], [t-d], and [k-g]) with a high degree of consistency, as illustrated in the first graph in Fig. 7. Others showed a high degree of contrast for the [p-b], but not for [t-d] and [k-g] that appears as random responses in the voiced-like tokens, as illustrated in the second graph in Fig. 7. Several others preferred the voiceless series across the continuum (the third graph in Fig. 7) while a limited few actually showed a reverse trend (fourth graph in Fig. 7), choosing the voiceless token during the more voiced-like tokens and vice versa. These results will be further discussed in Section 4.

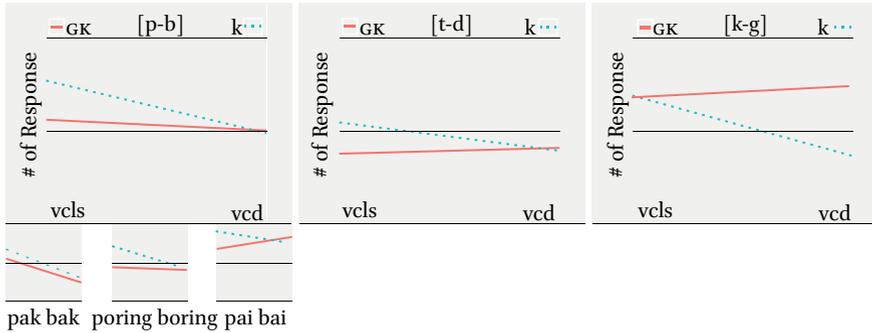


FIGURE 6 The image on the left represents the mean average of the responses to the canonical [p-b] stimuli. The centre image represents the mean average of the responses to the canonical [t-d] stimuli, while the image on the right represents the mean average of the responses to the canonical [k-g] stimuli. For all three images, the responses to the voiceless token is on the left while the response to the voiced token is on the right.

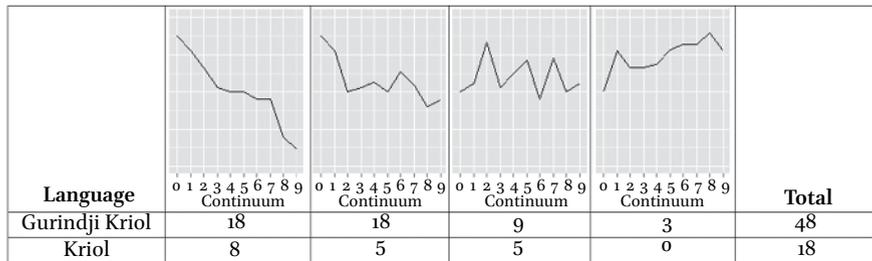


FIGURE 7 Individual participant perceptual trends across all three places of articulation combined.

3.2 Simplified Experiment

The results found in this section detail the participant responses to the simplified experiment. This section contains line plots for all three minimal pair responses ([p-b], [t-d], and [k-g]) in addition to the results of the linear mixed effects model.

Figure 6 suggests that, overall, the Kriol participants typically selected the voiceless stop token when heard in its canonical form, apart from [t-d] that had more random responses. When presented with the canonical voiced stop stimuli, there appears to be a slight preference for the voiced token apart from [p-b] that were responded to at random—with the clear exception of *pak-bak*. For the Gurindji Kriol group responses to the stimuli were even more random; responses to [p-b] were nearly completely in free variation, while there was a slight impressionistic preference for [d] over [t]. On the other hand, the Gurindji Kriol group appear to overwhelmingly prefer [k] over [g].

TABLE 9 *Generalized linear mixed effects model results for voiced-voiceless perception in group I and III.*

	Estimate	Std. Error	2.5%	97.5%	z-value	Pr(> z)	
(Intercept)	0.85	0.41	0.05	1.6	2.1	0.036	*
Continuum	-1.23	0.26	-1.7	-0.7	-4.7	2.10E-06	***
Gurindji Kriol	-1.11	0.35	-1.8	-0.4	-3.1	0.002	**
Velar	-0.16	0.87	-1.9	1.5	-0.2	0.85	
Continuum : GK	1.20	0.38	0.5	1.9	3.2	0.001	**
GK : Velar	1.49	0.47	0.6	2.4	3.2	0.002	**

Table 9 contains the results from the generalized linear mixed effects model using *response* as the dependent variable. Based on the model output, the intercept contains the responses to the canonical voiced stimuli.

The intercept, with a 'base' value⁸ of 0.85 log-odds, suggests that participants from the Kriol group selected the image of the voiceless stop token, when they heard its canonical form, 70% of the time. The probability of selecting [p] and [t] group decreased, on average, to 41% (-0.4 log-odds) when presented with a canonical token containing the voiced stop. These results suggest a slight preference towards the 'correct' forms when heard.

The Gurindji Kriol participants showed more random response preferences to the [p-b] and [k-g] stimuli. Here, when [p] and [t] were presented in their canonical forms, participants selected them, on average, 44% (-0.26 log-odds) of the time vs. 43% (-0.29 log-odds) when presented with the voiced counterpart. This suggests there was virtually no difference in how the participants categorized these categories perceptually.

Responses to the velar pair *kol-gol*, however, differed significantly from the other stimuli. For the Kriol participants, they chose *kol* 67% (0.69 log-odds) of the time when presented with its canonical form, while they only chose it 37% (-0.54 log-odds) of the time when presented canonical *gol*. For the Gurindji Kriol speakers on the other hand, there was an overwhelming preference in favour of the voiceless token no matter the token presented. (*K*)*ol* was selected 75% (1.07 log-odds) of the time in its canonical form while (*k*)*ol* was selected 74% (1.04 log-odds) of the time when canonical (*g*)*ol* was presented. No other predictors or interactions of predictors were shown to be significant and were thus removed from the final version of the model.

8 All results are displayed to two decimal places, though calculated to the fifth decimal place.

4 Discussion

This study was designed to investigate how listeners from six distinct groups perceive stop voicing in Kriol words that have English cognates. The six groups of participants included the Kriol and Gurindji Kriol listeners in group I (children), the Kriol and Gurindji Kriol listeners in group II (the preteens to middle-aged adults), and Kriol and Gurindji Kriol listeners in group III (older adults). Groups I and III were tested using the simplified version of our experiment because we identified these participants as having little exposure to mainstream English. It was thought that this might provide us with a comparative basis for testing if the increased exposure to English in the group II participants may be a factor in stop voicing perception.

Based on the statistical results from Section 3.2, Kriol listeners in groups I and III show a very limited, yet significant, degree of contrastability suggesting, tentatively, that Kriol listeners may be in the midst of acquiring the stop voicing contrast, although it has not fully come to fruition. Unlike the [p-b] minimal pairs in the standard experiment, only *pak-bak* 'park-bark' showed any degree of contrast while [k-g] unexpectedly showed a stronger degree of contrast than that of the standard experiment. The results do, however, suggest those in groups I and III appear to identify the voiceless series with a relatively high degree of consistency. This suggests the voiceless series is interpreted as distinct from the voiced, even though the voiced series is not reliably identifiable.

Contrarily, for the Gurindji Kriol participants in groups I and III there is virtually no difference in the responses to the [p-b] and [t-d] stimuli, suggesting free variation. In contrast, for [k-g] there was an overwhelming preference for [k]. It is also of interest that non-significant differences were revealed for *age* suggesting little difference in response patterning between those in group I and III. A word of caution however: the simplified version of the experiment only contained canonical forms of voiceless and voiced minimal pairs that does not permit us to identify any sort of categorical boundary and only provides us with general trends. It should also be noted that while this was a simplified version of the standard experiment, it still may have been difficult for the participants in groups I and III to understand the task at hand, which may be responsible for some of the varied responses in the data.

Overall the results from the older adults (group III) from Ngukurr and Kalkaringi who participated in the simplified version of the experiment suggest that the cattle station pidgin (from where Kriol and Gurindji Kriol later emerged) had no stop voicing contrast. Older adults at Ngukurr do show some contrast, though, in comparison with older adults at Kalkaringi, which is likely the result of higher levels of exposure to English (recall that Ngukurr has had

50 more years of Western formal schooling). The results from the Kriol children (group I), which show some degree of contrast, suggest that the children are exposed to a voicing contrast when they are acquiring Kriol. This fits with Bundgaard-Nielsen and Baker's (2016) observation that Kriol adults show a voicing contrast in production. Conversely, the random response results from the Gurindji Kriol children suggest that they are not exposed to a voicing contrast when they are acquiring Gurindji Kriol, which fits with Jones and Meakins' (2013) study that shows no VOT contrast in the production of stops.

Turning to group II, a clear result pertaining to *exposure to English* had a significant effect on the degree of contrast where those with more exposure showed an increased probability of differentiating between the minimal pair stimuli, while those with low exposure showed the reverse trend; a result indicative of those found by Klein (2013) and Flege et al (2003; 1996) suggesting that substantial exposure to a language may improve perception and production. This suggests that through constant and increasing contact with English, in addition to the recognition of Kriol/English cognates, a voicing contrast is now developing. While the individual participant results suggest 39% of listeners, from both language groups combined, appear to have adopted the stop voicing contrast perceptually, a nearly equal number of listeners recognize only the voiceless series with any degree of certainty. Moreover, an additional 25% had no particular preference. This finding might indicate that exposure to English may not account for all the variation, since approximately 75% of the participants (across both groups) were reported to have *high* exposure to English, which does not fully account for the outcomes presented in Fig. 7. Having a significant number of *high* exposure participants with higher degrees of contrastability, and a smaller group of mixed exposure backgrounds with varying response patterns supports the notion that linguistic complexities develop incrementally and with variation (Harrington, Kleber, Reubold, and Stevens, 2016). There may be a number of factors involved that could explain why this shift towards English is taking place. For example, it is common knowledge that the 'prestigious' language, in this case Standard English, often has a unidirectional influence on the 'non-prestigious' language, in this case, Kriol and Gurindji Kriol, under contact (Fought, 2010; Hickey, 2010a). Therefore, it might be that Kriol and Gurindji Kriol speakers are subconsciously assimilating their stop series to that of English, analogous to how Pasquale's (2005) Quechua participants were shifting to more Spanish-like stop production. Another possibility that might partially explain this subtle shift includes benefits in reducing the cognitive load by optimising the contrasts in the phonology. When a large proportion of the lexicon is assimilated to the phonology of another language (e.g., Kriol → Gurindji in Gurindji Kriol; or French → Cree in Michif), certain

segmental contrasts with a high functional load (i.e., phonemes that played a substantial role in distinguishing large portions of a language's lexicon, like stop voicing in English), may be lost. Since both Kriol and Gurindji Kriol have been under constant and increasing contact with English, allowing for the recognition of Kriol/English cognates, speakers/listeners of these languages now have at their disposal an additional and potentially beneficial contrast that was initially lost during assimilation in the pidgin phase of Kriol's development. It should be noted, however, that these explanations are speculative, a corpus based analysis of the phonological distributions of these languages may prove insightful.

The results from this study also show that Kriol, as a language, appears to be further along in developing the stop voicing contrast than Gurindji Kriol. This can be seen in the data from groups I and III where Gurindji Kriol responses are more varied than those of Kriol. Moreover, this is apparent in the significant degree of contrast in the adult Kriol and Gurindji Kriol data, with the former showing more consistent response patterns than the latter. These results fit with Jones and Meakins' (2013) study that shows no VOT contrast in the production of stops in Gurindji Kriol speakers and Bundgaard-Nielsen and Baker's (2016) findings that demonstrate that Kriol speakers are able to contrast voiceless from voiced stops both perceptually and productively, although their low sample size (3 speakers) may have meant they inadvertently tested Kriol speakers who are indeed able to fully contrast the stops (Fig. 7). Additional studies examining the production of stop consonants in Kriol across a larger sample of speakers will be a welcome addition to the literature. It is also worth noting that while response patterns in both language groups are in greater flux in the post-labial voiced series, the [p-b] minimal pairs show a great deal of contrast. This suggests that the stop voicing contrast may already be entering both languages perceptually. While it is not impossible to come across languages with voicing gaps in their inventory (e.g. Chickasaw), the following stop pattern does not appear to be documented [p-b, t, k] (Maddieson, 2013). Therefore, it appears both languages are in the midst of adopting the voiced series perceptually, which just so happens to be beginning at the labial place of articulation.

One of the reasons for this incremental and variable change might be the difficulties in acquiring perceptual contrasts after early adolescence (Bosch et al., 1997; Caramazza et al., 1973; Guion, 2003; Hazan and Barrett, 2000; Kuhl, 2004; Werker and Tees, 1984), which means not every adult, even those with a high degree of exposure to English, will pick up on the acoustic cues required to differentiate between voicing categories. Being that the production of word-initial stops in Gurindji Kriol shows little evidence of a voicing

contrast (manifesting primarily with short-lag VOT) (Jones and Meakins, 2013), it should come as little surprise that the acoustic cues needed by a listener to signal a voicing distinction are lacking. The fact that listeners from groups I and III, with low exposure to English, show little indication that they have acquired such cues supports our results that listeners, with the contrast, from group II are primarily relying on non-native perceptual cues from their L2—a language acquired during school, not taught explicitly using an ESL framework, and rarely used outside the classroom. Being that a number of Gurindji Kriol listeners from this study could assess such contrastive cues at all, highlights our linguistic resilience for acquiring aspects of language even in less than optimal learning environments. However, as the acoustic cues pertaining to stop voicing from the English become more robust and speakers begin to make more associations between English cognates, we expect the stop voicing contrast will continue to slowly disseminate through the communities until it eventually becomes nativised.

Returning to the central question posed by this paper—that of whether the stratification of the grammar of contact language extends to the phonology and how it develops—if we define stratification as the phonological differences between the lexicon from language X and the lexicon from language Y, Gurindji Kriol may fit this definition. This can be seen in the fact that Gurindji Kriol maintains its phonological system while the lexicon from English (via the cattle pidgin and Kriol) appears to be conforming to a separate set of rules with the ongoing adoption of the stop voicing contrast. Unlike other mixed languages, which deal with quite different phonological systems (e.g., Michif (Cree and French) and Media Lengua (Quichua and Spanish)), Kriol overwhelmingly conforms to the systemic phonology of the traditional languages from which it is derived—Gurindji being one of these. This means when Gurindji Kriol was formed, there was little competition for phonological material from each source language, which made assimilation a straight forward task. Only now with the constant exposure to standard English, are we beginning to see the stop system enter these languages post-formation.

Acknowledgements

This study was funded by the ARC Centre of Excellence for the Dynamics of Language (CE140100041). Thanks to participants at Kalkaringi, Daguragu and Ngukurr. Particular thanks to Salome Harris, the then co-ordinator of Ngukurr Language Centre, who facilitated the organisation of the Kriol participants.

References

- Abramson, Arthur S. and Leigh Lisker. 1973. Voice-timing perception in Spanish word-initial stops. *Journal of Phonetics* 1: 1–8.
- Andrea, MacLeod and Carol Stoel-Gammon. 2005. Are bilinguals different? What VOT tells us about simultaneous bilinguals. *Journal of Multilingual Communication Disorders* 3: 118–127.
- Ash, Richard, Michael Chinen, Roger Dannenberg, Vaughan Johnson, and Shaw Martyn. 2012. Audacity (Version 2.0.0). SourceForge.net.
- Baker, Brett, Rikke Bundgaard-Nielsen, and Simone Graetzer. 2014. The obstruent inventory of Roper Kriol. *Australian Journal of Linguistics* 34(3): 307–344. <http://doi.org/http://dx.doi.org/10.1080/07268602.2014.898222>.
- Bakker, Peter. 2003. Mixed languages as autonomous systems. In Yaron Matras and Peter Bakker (eds.), *The Mixed Language Debate: Theoretical and Empirical Advances*, 107–149. Berlin: Walter de Gruyter.
- Bates, Douglas M., Martin Maechler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1): 1–48. <https://doi.org/doi:10.18637/jss.v067.i01>.
- Boersma, Paul and David Weenink. 1996. Praat, a system for doing phonetics by computer, version 3.4. *Institute of Phonetic Sciences of the University of Amsterdam, Report*, 132.
- Borden, Gloria, Katherine Harris, and Lawrence Raphael. 1994. *Speech Science Primer: Physiology, Acoustics, and Perception of Speech* (3rd ed.). Baltimore: Williams and Wilkins.
- Bosch, Laura, Albert Costa, and Sebastian Nuria. 1997. Vowel discrimination in early bilinguals and the perceptual magnet effect. In *Paper presented at the the 5th International Congress of the ISAPL*. Porto, Portugal.
- Bullock, Barbara. 2009. Phonetic reflexes of code-switching. In Barbara Bullock and Almeida J. Toribio (eds.), *The Cambridge Handbook of Linguistic Code-switching*, 163–181. Cambridge: Cambridge University Press.
- Bundgaard-Nielsen, Rikke, and Brett Baker. 2015. Perception of voicing in the absence of native voicing experience. Presented at the Interspeech 2015, Dresden.
- Bundgaard-Nielsen, Rikke and Brett Baker. 2016. Fact or furphy? The continuum in Kriol. In Felicity Meakins and Carmel O'Shannessy (eds.), *Loss and Renewal: Australian Languages Since Colonisation*, 175–216. Berlin: De Gruyter Mouton.
- Butcher, Andrew. 2008. Linguistic aspects of Australian Aboriginal English. *Clinical Linguistics and Phonetics* 22: 625–642.
- Campbell, Lyle. 1996. Phonetics and phonology. In Hans Goebel, Peter H. Nelde, Zdeněk Starý, and Wolfgang Wölck (eds.), *Contact Linguistics: An International Handbook of Contemporary Research*. Berlin: Mouton de Gruyter.

- Caramazza, Alfonso and Grace Yeni-Komshian. 1974. Voice onset time in two French dialects. *Journal of Phonetics* 2: 239–245.
- Caramazza, Alfonso, Grace Yeni-Komshian, Edgar B. Zurif, and Ettore Carbone. 1973. The acquisition of a new phonological contrast: the case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America* 54(2): 421–428.
- Chang, Charles B., Yao Yao, Erin F. Haynes, and Russell Rhodes. 2011. Production of phonetic and phonological contrast by heritage speakers of Mandarin. *Journal of the Acoustical Society of America* 129(6): 3964–3980. <https://doi.org/10.1121/1.3569736>.
- Charola, Erika and Felicity Meakins (eds.). 2016. *Yijarni: True Stories from Gurindji Country*. Canberra: Aboriginal Studies Press.
- Delano, Lamy. 2012. *A Variationist Account of Voice Onset Time Among Bilingual West Indians in Panama*. University of Florida.
- Dickson, Greg. 2016. Rethinking the substrates of Roper River Kriol: The case of Marra. In Felicity Meakins and Carmel O'Shannessy (eds.), *Loss and Renewal: Australian Languages Since Colonisation*, 145–173. Berlin: Mouton de Gruyter.
- Escudero, Paola. 2000. The perception of English vowel contrasts: Acoustic cue reliance in the development of L2 perception. In *Fourth International Symposium on the Acquisition of Second Language Speech*. University of Amsterdam, The Netherlands.
- Flege, James. 1991. Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America* 89: 395–411.
- Flege, James and Serena Liu. 2001. The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition* 23: 527–552.
- Flege, James, Carlo B. Schirru, and Ian R.A. MacKay. 2003. Interaction between the native and second language phonetic subsystems. *Speech Communication* 40: 467–491.
- Flege, James, Naoyuki Takagi, and Virginia Mann. 1996. Lexical familiarity and English-language experience affect Japanese adults' perception of /r/ and /l/. *Journal of the Acoustical Society of America* 99: 1161–1173.
- Fought, Carmen. 2010. Linguistic contact. In Raymond Hickey (ed.), *The Handbook of Language Contact*, 229–319. West Sussex: Wiley-Blackwell.
- Guion, Susan. 2003. The vowel systems of Quichua-Spanish bilinguals: Age of acquisition effects on the mutual influence of the first and second languages. *Phonetica* 60: 98–128.
- Harrington, Jonathan, Felicitas Kleber, Ulrich Reubold, and Mary Stevens. 2016. The relevance of context and experience for the operation of historical sound change. In Anna Esposito and Lakhmi Jain (eds.), *Toward Robotic Socially Believable Behaving Systems: Modeling Social Signs*, Vol. 11: 61–85. Springer.
- Harris, John. (1986). *Northern Territory Pidgins and the Origin of Kriol*. Canberra: Pacific Linguistics.

- Hazan, Valerie and Sarah Barrett. 2000. The development of phonemic categorisation in children aged 6 to 12. *Journal of Phonetics* 28: 377–396. <http://doi.org/doi:10.1006/jpho.2000.0121>.
- Hickey, Raymond. 2010a. *Handbook of Language Contact*, 1–29. West Sussex: Wiley-Blackwell.
- Hickey, Raymond. 2010b. *The Handbook of Language Contact*. West Sussex: Wiley-Blackwell.
- Hoonhorst, Ingrid, Cécile Colin, Emily Markessis, Monique Radeau, Paul Deltenre, and Willy Serniclaes. 2009. French native speakers in the making: From language-general to language-specific voicing boundaries. *Journal of Experimental Child Psychology* 104: 353–366.
- Hudson, Joyce. 1985. *Grammatical and semantic aspects of Fitzroy Valley Kriol*. Darwin: SIL.
- Hyman, Larry. 1970. *The role of borrowing in the justification of phonological grammar*. Unpublished Manuscript.
- Itô, Junko and Armin R. Mester. 1996. Japanese Phonology. In John A. Goldsmith (ed.), *The Handbook of Phonological Theory*, 817–838. Cambridge (Mass): Blackwell Publishing Ltd.
- Jones, Caroline and Felicity Meakins. 2013. Variation in voice onset time in stops in Gurindji Kriol: Picture naming and conversational speech. *Australian Journal of Linguistics* 33: 196–220. <http://doi.org/10.1080/07268602.2013.814525>.
- Keating, Patricia. 1984. Phonetic and phonological representation of stop consonant voicing. *Language* 60: 286–319.
- Kehoe, Margaret M., Conxita Lleó, and Martin Rakow. 2004. Voice onset time in bilingual German-Spanish children. *Bilingualism: Language and Cognition* 7: 71–88. <http://doi.org/doi:10.1017/S1366728904001282>.
- Klein, Michael. 2013. *Aspir(at)ing to speak like a native: Tracking voice onset time in the acquisition of English stops*. Working paper, Fairfax. Retrieved from http://www.gmu.edu/org/lingclub/WP/texts/9_MikeKlein.pdf.
- Kuhl, Patricia K. 2004. Early language acquisition: cracking the speech code. *National Review of Neuroscience* 11: 831–841. <http://doi.org/doi:10.1038/nrn1533>.
- Kuhl, Patricia K. and Paul Iverson. 1995. Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *Journal of the Acoustical Society of America* 97: 553–562.
- Kuhl, Patricia K. and James D. Miller. 1979. Speech perception by the chinchilla: Voiced voiceless distinction in alveolar plosive consonants. *Science* 190: 69–72.
- Kuznetsova, Alexandra, Per B. Brockhoff, and Rune Bojesen. 2014. ImerTest (Version 2.0-11). Cran.r Project.
- Lieberman, Alvin. 1957. Some results of research on speech perception. *Journal of the Acoustical Society of America* 29(1): 117–123.

- Liberman, Alvin, Pierre C. Delattre, and Franklin S. Cooper. 1958. Some cues for the distinction between voiced and voiceless stops in initial position. *Language and Speech* 1: 153–167.
- Lisker, Leigh. 1986. “Voicing” in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech* 29: 3–11.
- Lisker, Leigh and Arthur S. Abramson. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20: 384–422.
- Maddieson, Ian. 2013. Voicing and Gaps in Plosive Systems. In Dryer Matthew and Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Matras, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press.
- Matras, Yaron and Bakker, Peter. 2003. *The Mixed Language Debate: Theoretical and Empirical Advances*. Berlin: Walter de Gruyter.
- McConvell, Patrick and Felicity Meakins. 2005. Gurindji Kriol: A mixed language emerges from code-switching. *Australian Journal of Linguistics* 25: 9–30.
- McGuire, Grant. 2010. *A brief primer on experimental designs for speech perception research*. Unpublished Manuscript. Retrieved from http://people.ucsc.edu/~gmcguir1/experiment_designs.pdf.
- Meakins, Felicity. 2008. Unravelling languages: Multilingualism and language contact in Kalkaringi. In Jane Simpson and Gillian Wigglesworth (eds.), *Children's Language and Multilingualism: Indigenous Language Use at Home and School*, 247–264. New York: Continuum.
- Meakins, Felicity. 2011. *Case-Marking in Contact: The development and function of case morphology in Gurindji Kriol*. Amsterdam: John Benjamins.
- Meakins, Felicity. 2012. Which Mix? – Code-switching or a mixed language – Gurindji Kriol. *Journal of Pidgin and Creole Language* 27(1): 105–140.
- Meakins, Felicity. 2013a. Gurindji Kriol. In Susanne Michaelis, Philippe Maurer, Martin Haspelmath, and Magnus Huber (eds.), *The Survey of Pidgin and Creole Languages*, Vol. 111: 131–139. Oxford: Oxford University Press.
- Meakins, Felicity. 2013b. Mixed languages. In Yaron Matras and Peter Bakker (eds.), *Contact Languages: A Comprehensive Guide*, 159–228. Berlin: Mouton de Gruyter.
- Meakins, Felicity. 2014. Language contact varieties. In Harold Koch and Rachel Nordlinger (eds.), *The Languages and Linguistics of Australia: A Comprehensive Guide*, 365–416. Berlin: Mouton de Gruyter.
- Meakins, Felicity. 2015. From absolutely optional to only nominally ergative: the life cycle of the Gurindji ergative suffix. In Francesco Gardani, Peter Arkadiev, and Nino Amiridze (eds.), *Borrowed Morphology*, 189–218. Berlin: Mouton De Gruyter.
- Miller, Joanne and Emily Dexter. 1988. Effects of speaking rate and lexical status on phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance* 14: 369–378.

- Munro, Jennifer. 2000. Kriol on the move: A case of language spread and shift in Northern Australia. In Jeff Siegel (ed.), *Processes of Language Contact: Studies from Australia and the South Pacific*, 245–270. Saint-Laurent (Quebec): Fides.
- Munro, Jennifer. (2004). *Substrate Language Influence in Kriol: The Application of Transfer Constraints to Language Contact in Northern Australia*. University of New England, Armidale.
- Muysken, Pieter. 1981. Halfway between Quechua and Spanish: The case for relexification. In Arnold R. Highfield (ed.), *Historicity and Variation in Creole Studies*, 57–78. Ann Arbor: Karoma Publishers.
- Muysken, Pieter. 1997. Media Lengua. In Sarah G. Thomason (ed.), *Contact languages: A Wider Perspective*, 365–426. Amsterdam: John Benjamins.
- Pasquale, Michael. 2005. Variation of voice onset time in Quechua-Spanish bilinguals. In Luis A. Ortiz López and Manel Lacorte (eds.), *Contactos y contextos lingüísticos. El español en los Estados Unidos y en contacto con otras lenguas. Segunda edición*, Vol. 27. Madrid: Lingüística Iberoamericana.
- Rosen, Nicole, Jesse Stewart, and Olivia Cox. 2016. Presented at the NWAV 45, Vancouver. Retrieved from <http://linguistics.arts.sfu.ca/nwav45/> (accessed March 30, 2017).
- Sandefur, John R. 1979. An Australian Creole in Northern Territory: A description of Ngukurr-Bamyili dialects (Part 1). In *Work Papers of SIL-AAB*. Darwin: SIL, Australian Aborigines Branch.
- Sandefur, John, and Harris, J. 1986. Variation in Australian Kriol. In Joshua Fishman (ed.), *The Fergusonian Impact*, 179–190. Berlin: Mouton de Gruyter.
- Schultze-Berndt, Eva, Felicity Meakins, and Angelo, Denise. 2013. Kriol. In Susanne Michaelis, Philippe Maurer, Martin Haspelmath and Magnus Huber (eds.), *The Survey of Pidgin and Creole Languages*, Vol. 1: 241–251. Oxford: Oxford University Press.
- Scobbie, James. 1988. Interactions between the acquisition of phonetics and phonology. In M. Catherine Gruber, Derek Higgins, Kenneth Olsen, and Tamra Wysochi (eds.), *Papers from the 34th Annual Meeting of the Chicago Linguistics Society, 11*. Chicago: Chicago Linguistics Society.
- Simpson, Jane. 2000. Camels as pidgin-carriers: Afghan cameleers as a vector for the spread of features of Australian Aboriginal pidgins and creoles. In Jeff Siegel (ed.), *Processes of language contact: Studies from Australia and the South Pacific*, 195–244. Saint-Laurent (Quebec): Fides.
- Stewart, Jesse. forthcoming. Vowel perception in Spanish, Media Lengua, and Quichua.
- Stewart, Jesse. 2014. A comparative analysis of Media Lengua and Quichua vowel production. *Phonetica* 7: 159–182. <http://doi.org/10.1159/000369629>.
- Stewart, Jesse. 2015. *Production and perception of stop consonants in Spanish, Quichua, and Media Lengua*. University of Manitoba, Dspace.

- Summerfield, Quentin. 1981. Articulatory rate and perceptual constancy in phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance* 7: 1074–1095.
- Surendran, Dinoj and Partha Niyogi. 2006. Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals. In Nedergaard Thomssen (ed.), *Current trends in the theory of linguistic change. In commemoration of Eugenio Coseriu (1921–2002)*. Amsterdam / Philadelphia: John Benjamins.
- Thomason, Sarah. 1997. *Contact Languages: A Wider Perspective*. Amsterdam: John Benjamins.
- Vainio, Martti, Juhani Järvikivi, Stefan Werner, Nicholas Volk, and Jarmo Välikangas. 2002. Effect of prosodic naturalness on segmental acceptability in synthetic speech, 143–146. Presented at the Speech Synthesis, Proceedings of 2002 IEEE Workshop. <http://doi.org/10.1109/WSS.2002.1224394>.
- van Gijn, Rik. 2009. The phonology of mixed languages. *Journal of Pidgin and Creole Languages* 24: 91–117. <http://doi.org/10.1075/jpcl.24.1.04gij>.
- Wedel, Andrew, Abby Kaplan, and Scott Jackson. 2013. High functional load inhibits phonological contrast loss: a corpus study. *Cognition* 128(2): 179–86. <http://doi.org/10.1016/j.cognition.2013.03.002>.
- Werker, Janet F. and Richard C. Tees. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7: 49–63.