Cross-language Influence in the Stop Voicing Contrast in Heritage Tagalog

Yoonjung Kang, Sneha George, and Rachel Soo
University of Toronto Scarborough

ABSTRACT

In heritage bilinguals' sound structure, some aspects of the sound system are more prone to cross-language influence than others. In this study, we compare two different models of cross-language influence, a phonological markedness based model, which proposes that influence selectively affects a phonologically marked structure, and a phonetic category based model, where influence is mediated through cross-language equivalence classification of similar phones. The empirical data for the study comes from the production of the voicing contrast in English and Tagalog stops by heritage Tagalog speakers in Toronto. We compare the heritage speakers' production with native control productions and also probe the effect of lexical stress in voicing realization as evidence for the underlying target structure of stop categories. The key empirical findings are that the heritage speakers produce their voiceless stops in both languages nearly native-like, including a native-like stress effect, but voiced stops exhibit considerable cross-language influence and assimilatory stress effects. We propose that the heritage speakers successfully establish separate phonetic categories for English and Tagalog voiceless stops, but form a partially merged category for English and Tagalog voiced stops. The findings provide partial support for the phonetic category based model of influence over the phonological markedness based model.

Keywords: *Tagalog, English, voice onset time, stress, stop consonants*

INTRODUCTION

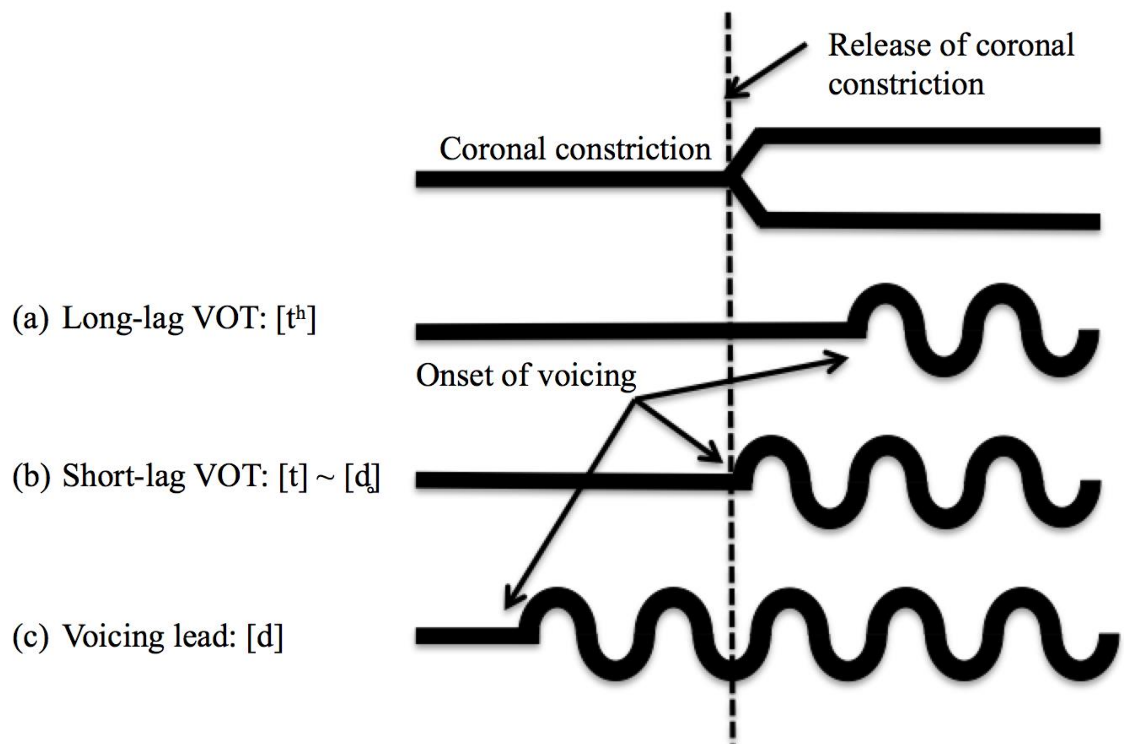
A heritage speaker is commonly defined as a speaker who learns a minority language as his/her first language at home and then learns the language spoken by the majority in the wider community, which often ends up becoming his/her dominant language (Chang, Yao, Haynes, & Rhodes, 2011; Polinsky & Kagan, 2007).[1] Studies on heritage language sound patterns often focus on understanding the extent to which the speech of this special population of bilingual speakers is native-like and developing a theory of how and why their speech patterns differ from those of monolinguals. The current paper contributes to this body of literature by studying the speech of Tagalog heritage speakers in Toronto, Canada. In particular, we examine the acoustic realization of the plosive voicing contrast in the two languages of this heritage population, Tagalog and English. Our study examines the timing and duration of phonetic voicing in stop consonant production and probes how stress modulates the realization of voicing contrasts. We compare the speech patterns of a heritage population with those of native speaker comparison groups and examine where differences occur and what those differences tell us about the sound structure of heritage speakers.

PREVIOUS RESEARCH

Typology of Voicing Contrast

The primary acoustic dimension that differentiates voiced and voiceless stop consonants is Voice Onset Time (VOT), or "the duration of the time interval by which the onset of periodic pulsing either precedes or follows release" of a stop closure (Lisker & Abramson, 1964). Languages differ in how they divide the VOT continuum to distinguish voiced and voiceless stops. For example, English and German are *aspirating* languages where phonemically voiceless stops are produced with long-lag VOT, while phonemically voiced stops are produced with short-lag VOT and limited voicing lead in word-initial position. Other languages, such as Russian, Spanish, Dutch, Tamil, Hungarian, and Tagalog (as we will see below), are *true voicing* languages where phonemically voiced stops are realized with voicing lead (i.e., negative VOT), while phonemically voiceless stops are produced with short-lag VOT. Figure 1 illustrates the relative timing of oral and laryngeal articulations that underlie the three types of Voice Onset Time configurations.

Figure 1. A Schematic Illustration of Three Voice Onset Time Configurations:
(a) Long-lag; (b) Short-lag; (c) Voicing-lead



Keating (1984) proposes that phonetic categories of stop consonants of the world's languages are limited to three types—{voiced}, {vl.unasp.} (voiceless unaspirated), and {vl.asp.} (voiceless aspirated), corresponding to lead voicing, short-lag, and long-lag stops, respectively. The curly brackets are used to represent phonetic targets and to distinguish them from phonological features. Under Keating's model, true voicing languages and aspirating languages share the same phonological contrast of [±voice], but they differ in how the contrast is mapped to phonetic categories. In true voicing languages, voicing contrast is realized as {voiced} versus {vl.unasp.} and in aspirating languages the same phonological contrast is typically realized as {vl.unasp.} versus {vl.asp.}.[2]

Others argue that the phonetic difference between these two language types stems from different phonological representations (Beckman, Jessen & Ringen, 2013; Honeybone, 2005; Iverson & Salmons, 1995). According to this view, in true voicing languages, the contrast is defined by an active voicing gesture on voiced stops, which are marked by a monovalent feature [voice], while voiceless stops are not specified for any feature. In aspirating languages, on the other hand, the contrast is defined by an active aspiration gesture on voiceless stops ([spread glottis]), while voiced stops are unmarked. Table 1 summarizes how the voiced and voiceless stops of aspirating languages and true voicing languages are distributed along the VOT continuum and how they are represented in distinctive features under these two models of voicing contrast.

Table 1.

Stops of Aspirating Versus True Voicing Languages

(a) A phonological feature model (cf. Iverson & Salmons, 1995)

	Voicing lead	Short lag	Long lag
True voicing languages	/b/ ([voice])	/p/ (∅)	
Aspirating languages		/b/ (∅)	/p/ ([sp. glottis])

(b) A phonetic category model (cf. Keating, 1984)

	Voicing lead **{voiced}**	Short lag **{vl.unasp.}**	Long lag **{vl.asp.}**
True voicing languages	/b/ ([+voice])	/p/ ([-voice])	
Aspirating languages		/b/ ([+voice])	/p/ ([-voice])

Asymmetrical Influence in Heritage Languages

Interaction between a true voicing language and an aspirating language is one of the most widely studied topics in cross-language sound patterns and many studies have examined how bilingual speakers, including heritage speakers and second language (L2) learners, cope with the challenges of accommodating different laryngeal systems in their cross-language sound structure (Broselow & Kang, 2013; Fowler, Sramko, Ostry, Rowland, & Hallé, 2008; Kang & Guion, 2006; Kang & Nagy, 2016; MacLeod & Stoel-Gammon, 2009; Nagy, 2015; Simon, 2009; Sundara, Polka, & Baum, 2006). Studies on heritage speakers generally find that early exposure to the heritage language endows heritage speakers with an advantage over late L2 learners in approximating the native speaker norms of the heritage language (Au, Oh, Knightly, Jun, & Romo, 2008; Knightly, Jun, Oh, & Au, 2003; Oh, Jun, Knightly, & Au, 2003). Heritage speakers are also more likely to attain native-like production patterns in the majority language of the community compared to corresponding late L2 learners (Kang & Guion, 2006; MacKay, Flege, Piske, & Schirru, 2001; McCarthy, Evans, & Mahon, 2013; Newlin-Łukowicz, 2014).[3] Heritage speakers presumably have an advantage because they are more likely than late L2 learners to establish independence between similar sound categories of their two languages and retain cross-language contrast (Chang, Yao, Haynes, & Rhodes, 2011; Kang & Guion, 2006).

Most studies also find, however, that while the speech of heritage speakers closely approximates native speaker norms, it often does not exactly match them. The differences in heritage speakers' speech are usually attributable to assimilation between similar phones of the two languages, indicating that the two languages of bilingual speakers are not completely independent, even for highly proficient bilinguals (Antoniou, Best, Tyler, & Kroos, 2010; MacKay, Flege, Piske, & Schirru, 2001). In the domain of voicing contrast, influence is attested for both voiceless and voiced stops, even for early bilinguals, but some studies find that such assimilatory influences tend to occur more with voiced stops than with voiceless stops.[4] For example, Newlin-Łukowicz (2014) reports that Polish-English heritage speakers in New York City produce English voiced stops with prevoicing more frequently (around 50%) than non-Polish English speakers (17%). Polish is a true voicing language and the high rate of prevoicing of English voiced stops is attributed to the influence of Polish voiced stops. Interestingly, however, these heritage speakers produce English voiceless stops with VOT values comparable to non-Polish English speakers and do not show influence from short-lag Polish voiceless stops.

Newlin-Łukowicz (2014) explains that cross-language transfer is filtered through phonological markedness and that Polish voiced stops, being specified as [voice], are more likely to influence English voiced stops than Polish voiceless stops, which are unmarked for laryngeal features. Newlin-Łukowicz's study did not examine heritage speakers' Polish stop production, but extending the logic of her markedness based account would predict that the influence of English on a prevoicing language like Polish should affect voiceless stops more than voiced stops, as voiceless stops are featurally marked structures in English, while voiced stops are not. We can think of this asymmetry as a hierarchy of influence rather than categorical restrictions; that is, other things being equal, a true voicing language (with its phonologically marked voiced stops) would be more likely to interfere with English voiced rather than voiceless stops. On the other hand, English (with its phonologically marked voiceless stops) would be more likely to interfere with the voiceless rather than voiced stops of a true voicing language. This prediction is summarized in Table 2(a).

A different account of cross-language influence on voiced stop production is proposed by MacKay, Flege, Piske and Schirru (2001), who examined word-initial voiced stops of English and Italian (the latter a true voicing language) produced by Italian-English bilinguals in Canada. They found that these bilinguals, both early as well as late arrivals, produced more prevoiced tokens of English voiced stops than monolingual English speakers. At the same time, they produced their Italian voiced stops with full prevoicing less often than Italian monolinguals. According to Flege's (1995) *Speech Learning Model,* production of L2 speech sounds is affected by similar L1 sounds when the sounds
are perceived as equivalent and form a single merged category. Building on Flege (1995), MacKay, Flege, Piske and Schirru (2001) argue that the bidirectional influence indicates that these bilinguals did not form separate categories for the voiced stops of English and Italian, and they propose two reasons why. Firstly, while relatively infrequent, an Italian-like realization of voiced stops (i.e., prevoiced stops) is an acceptable variant for English in word-initial position and there is no communicative pressure to form separate categories for L2 English voiced stops, independent of L1 Italian voiced stops. Another reason is based on the observation that possible

stop categories of the world's languages are limited to three types (i.e., lead voicing, short-lag VOT, and long-lag VOT, based on Keating, 1984) and the same upper limit of three categories applies to individual speakers' inventories (Flege & Eefting, 1988). Given this limit, the presence of short-lag voiceless stops in Italian (and similarly, in other true voicing languages) preempts the option of creating a separate short-lag VOT category for English voiced stops.[5]

MacKay, Flege, Piske and Schirru's (2001) study does not examine voiceless stops and we do not have any information about the relative malleability of voiced versus voiceless stops in their bilingual population. However, if we extend MacKay, Flege, Piske and Schirru's logic, we would predict less influence in the voiceless stops of the two languages compared to voiced stops. The long-lag voiceless stops of English and short-lag voiceless stops of Italian (or a similar true voicing language) are well separated in their distribution in word-initial position and there is no L1 stop category that preempts the formation of an independent phonetic category for English long-lag voiceless stops (Flege & Eefting, 1988). As a result, equivalence classification is less likely for voiceless stops than voiced stops, which, in turn, predicts more cross-language influence for voiced stops rather than voiceless stops. Table 2(b) summarizes the prediction of this phonetic category based account. Again, we can think of this asymmetry as a hierarchy of malleability; that is, other things being equal, voiced stops (in both types of languages) are more likely to show influence than voiceless stops because the distributional asymmetry of voicing categories in the two languages makes equivalence classification more likely for voiced stops than for voiceless stops.

Table 2.

Predicted cross-language influence on the voicing contrast in bilinguals' two languages where the first language (L1) is a true voicing language and the L2 is an aspirating language (inequalities indicate that the category is more susceptible to influence).

(a) Phonological markedness based model

	L1 influence on L2	***L2 influence on L1***
voiceless	NA	long VOT
voiced	more prevoicing	NA
	voiced>voiceless	voiceless>voiced

(b) Phonetic category based model

	L1 influence on L2	***L2 influence on L1***
voiceless	NA	NA
voiced	more prevoicing	less prevoicing
	voiced>voiceless	voiced>voiceless

To test the full predictions of these two accounts, we need to examine the voiced and voiceless stops of both languages of a heritage speaker population. Antoniou, Best, Tyler and Kroos (2010) examine Australian-born Greek heritage speakers' production of English and Greek voiced and voiceless stops in word-initial and word-medial positions and found that these early bilinguals produced the stops of their two languages remarkably monolingual-like. The only exception was their production of English /b/ in word-medial position, which was significantly more prevoiced (i.e., Greek-like) than English control productions. These results exhibit the asymmetry between English voiced and voiceless stops predicted by both accounts discussed above. However, these speakers have essentially reached their ceiling in their heritage language, producing both the voiced and voiceless stops of Greek in monolingual-like fashion. As a result, we do not observe any asymmetry between voiced and voiceless stops and the two competing accounts in Table 2 are under-differentiated.

To summarize, previous studies provide an incomplete picture of the pattern of asymmetry in cross-language influence because they examined transfer effects in only one language of a heritage population and/or in only one type of stops. Our study fills this gap by examining the voiced and voiceless stops of both Tagalog and English produced by Tagalog-English heritage speakers. This allows us to investigate if and how cross-language influence affects voiced and voiceless stops differently across the two languages of heritage speakers and determine which model of asymmetrical transfer provides a better account.

Lexical Stress and Voicing Contrast

The voiced and voiceless stops of aspirating languages and true voicing languages also differ in how their production is modulated by lexical stress. Studies on the effect of lexical stress on VOT in English consistently find that stress enhances aspiration of voiceless stops; VOT is lengthened when voiceless stops occur in stressed syllables compared to unstressed syllables (Antoniou, Best, Tyler, & Kroos, 2010; Lisker & Abramson, 1967; Simonet, Casillas, & Díaz, 2014). Stress has a similar effect on the long-lag voiceless stops of German, an aspirating language (Keating, 1984). The effect of lexical stress on voiced stops in English, on the other hand, is inconsistent. Lisker and Abramson (1967) report more occurrences of prevoicing of word-initial voiced stops in stressed conditions than unstressed conditions and Keating (1984) similarly reports longer prevoicing for English voiced stops before a stressed vowel than an unstressed vowel. Simonet, Casillas and Díaz (2014) report a lower average VOT for English word-initial voiced stops in stressed conditions, but note that the stress effect on voiced stops is much smaller than what is found for voiceless stops. Antoniou, Best, Tyler and Kroos (2010), on the other hand, find shorter prevoicing (or an increased average VOT) for word-medial voiced stops in stressed conditions than unstressed conditions and Davidson (2016) reports that phrase-medially, fully voiced stops occur less frequently and devoiced stops occur more frequently in stressed than unstressed syllables. Keating (1984) finds a similar pattern for German word-medial voiced stops; that is, less prevoicing before a stressed than an unstressed vowel.

The effect of stress on the stops of true voicing languages also shows an asymmetry between voiced and voiceless stops. The small number of available studies shows that the effect is consistent for voiced stops, where voicing lead or closure voicing duration is lengthened in stressed compared to unstressed conditions; this is found for word-medial voiced stops in Dutch (Cho & McQueen, 2005) and Greek (Antoniou, Best, Tyler, & Kroos, 2010) and for word-initial voiced stops in Spanish (Simonet, Casillas, & Díaz, 2014). The effect of lexical stress on short-lag voiceless stops, however, is inconsistent. Cho and McQueen (2005) found that stress shortens the VOT of word-initial voiceless stops in stressed over unstressed syllables in Dutch, which they interpret as stress enhancing the {vl.unasp.} (=[-spread glottis]) feature of short-lag stops. Antoniou, Best, Tyler and Kroos (2010) and Simonet, Casillas and Díaz (2014), on the other hand, found no effect of lexical stress on the VOT of the voiceless stops of Greek and Spanish, respectively.

The consistent trend found in these previous studies is that stress has the effect of enhancing the voicing contrast by augmenting the VOT of the long lag stops in aspirating languages and by augmenting the voicing lead of the prevoiced stops in true voicing languages. The short-lag stops, on the other hand, show a minimal or inconsistent effect. A similar pattern of contrast augmentation is found in the effects of slow speech rate (Allen & Miller, 1999; Kessinger & Blumstein, 1997) and clear speech (Schertz, 2012; Smiljanic & Bradlow, 2008). The phonological model of cross-language voicing contrasts (Table 1(a)) accounts for this asymmetrical pattern of stress-conditioned enhancement as a function of asymmetrical feature specification. Under this view, stress enhances the marked phonological features of languages; that is, [spread glottis] for aspirating languages and [voice] for true voicing languages (Beckman, Helgason, McMurray, & Ringen, 2011; Simonet, Casillas, & Díaz, 2014), while short-lag stops, which are underspecified, do not have the same featural target for enhancement. These predictions are summarized in the first two columns of Table 3(a).

Table 3.

Predicted lexical stress effects on voicing contrasts and cross-language influence in monolingual and bilingual speech, where L1 is a true voicing language and L2 is an aspirating language.

(a) A phonological markedness based model

	monolingual		***bilingual***	
	Aspirating language (L2)	***True voicing language (L1)***	***L1 influenced L2***	***L2 influenced L1***
voiceless	longer VOT	NA	NA (=longer VOT)	long VOT
voiced	NA	more prevoicing	more prevoicing	NA (=more prevoicing)

(b) A phonetic category based model

	monolingual		***bilingual***	
	Aspirating language (L2)	***True voicing language (L1)***	***L1 influenced L2***	***L2 influenced L1***
voiceless	longer VOT	inconsistent	NA (=longer VOT)	NA (=inconsistent)
voiced	Inconsistent	more prevoicing	**more prevoicing**	NA (=more prevoicing)

Under the phonetic model of cross-language voicing contrasts (Table 1(b)), on the other hand, asymmetrical stress effects are explained as a convergence of two types of enhancement: of phonological contrast (i.e., *polarization* in Keating, 1984), which maximizes the distinction between voiced and voiceless stops, and of phonetic features, which promotes the maximally accurate realization of phonetic targets (Cho & McQueen, 2005). For the long-lag voiceless stops of an aspirating language and the prevoiced stops of a true voicing language, lengthening of aspiration and prevoicing enhances the contrast between voiced and voiceless stops and also enhances the phonetic targets {vl.asp.} (or {+spread glottis}) and {voiced} (or {+slack vocal folds}), respectively.[6] For short-lag stops, on the other hand, the two types of enhancements pull the stops in opposite directions. For the short-lag voiced stops of an aspirating language, enhancement of voicing contrast requires lengthening prevoicing away from voiceless stops, while enhancement of the phonetic target {vl.unasp.} ({-slack vocal folds}) requires shortening prevoicing in stressed conditions.[7] Similarly, for the short-lag voiceless stops of a true voicing language, enhancement of voicing contrast requires lengthening VOT away from prevoiced stops, while enhancement of the phonetic target {vl.unasp.}({-spread glottis}) requires shortening VOT in stressed conditions. In the case of Dutch voiceless stops, as examined by Cho and McQueen (2005), the phonetic feature enhancement (i.e., enhancement of {-spread glottis}) prevails and stress shortens the VOT of short-lag voiceless stops. It is conceivable, however, that the conflict between the two types of enhancements may be resolved differently in different conditions, leading to inconsistent lexical stress effects on short-lag stops. These predictions are summarized in the first two columns of Table 3(b).

Now the question arises as to how lexical stress affects voicing contrasts in heritage speakers' speech. Given the proposed link between the phonetic and phonological status of stops and the effect of lexical stress, stress effects can provide indirect evidence for how stop sounds are categorized in bilinguals' speech. Under the phonological model of cross-language influence, where the marked members of a contrast exert cross-language influence, we predict an asymmetrical pattern, which is schematically summarized in the last two columns of Table 3(a). In other words, we expect a true voicing language to affect the voiced stops of an aspirating language (stress induces more prevoicing) and an aspirating language to have an effect on the voiceless stops of a true voicing language (stress induces longer VOT). The unmarked members of contrasts (the voiceless stops of a true voicing language and the voiced stops of an aspirating language) are not expected to show an influence of the other language. The phonetic category based model, on the other hand, predicts a different asymmetry, which is summarized in the last two columns of Table 3(b). Under the assumption that the voiceless stops of the two languages in question are more likely to form separate phonetic categories than voiced stops, we predict less cross-language influence for voiceless stops than for voiced stops. Recall that the voiced stops of the two languages are expected to form a single category of {voiced} stops. This phonetic target specification predicts that stress will induce more prevoicing for the voiced stops of both languages involved, which is a divergence from the expected monolingual pattern for an aspirating language.

Simonet, Casillas and Díaz (2014) investigated this issue by examining stress-conditioned variation of word-initial voiced and voiceless stops in Spanish and English produced by Mexican-American bilinguals who were born and raised in Arizona. They found that the stress effect in heritage speakers' speech mirrored the corresponding monolingual pattern with the exception of English voiced stops; stress shortened the VOT of voiced stops in monolingual English, but raised the VOT of voiced stops in the bilinguals' English. The authors interpret this result in terms of the phonological model of voicing contrast (see Table 1(a)) and propose that the stress effect in the bilinguals' English voiced stops is due to the Spanish-influenced [voice] feature specification. However, note that the direction of the stress effect in the bilinguals' English voiced stops is, in fact, the opposite of what the [voice] feature specification would predict; that is, the [voice] feature predicts longer prevoicing for pretonic stops, as shown in Table 3(a), but instead, stress shortened prevoicing in the Spanish-influenced voiced stops of English.

The emergent stress effects in bilinguals' English voiced stops in Simonet, Casillas and Díaz's (2014) study is explained more naturally if we assume that English voiced stops have a phonetic target of {vl.unasp.}. Under this view, the seemingly conflicting stress effects in monolinguals' versus bilingual's English voiced stops is an enhancement of the same phonetic target {vl.unasp.}, or {-slack vocal folds, -spread glottis}. In monolinguals' English voiced stops, where stops are generally short-lag stops, this enhancement is achieved by shortening the VOT of short-lag stops, while in bilinguals' English voiced stops, where more prevoiced stops are found, the same phonetic feature enhancement is achieved by reducing prevoicing. In other words, contrary to the predictions of both models in Table 3, the emergent stress effect suggests that while bilinguals' English voiced stops are influenced by Spanish voiced stops and are pulled toward more prevoicing, this shift does not necessarily mean that bilinguals' stress effect differs from that of English monolinguals.

Antoniou, Best, Tyler and Kroos (2010) is another study on stop consonant production by heritage speakers that examined the effect of stress. They found that Australian-born Greek heritage speakers produced different stress effects on intervocalic stops of English and Greek and that those effects mirrored the patterns found in corresponding monolingual control groups: for English, voiceless stops in stressed syllables had longer VOTs and voiced stops in stressed syllables had shorter voicing leads, while for Greek, voiced stops in stressed syllables had longer voicing leads and voiceless stops were not affected by stress. It is notable that while the bilingual speakers produced English /b/ with more prevoicing (i.e., more Greek-like) than the monolingual norm, they still exhibited the monolingual-like stress effect on English /b/ (i.e., stress reduces prevoicing). This finding is also naturally explained as a consequence of the phonetic enhancement of {vl.unasp.}. Word-medial voiced stops are generally prevoiced in both monolingual and bilingual English and the stress-induced phonetic target enhancement of {vl.unasp.} reduced voicing in both speaker groups' data.[8] Note that while these bilinguals produced their English voiced stops with more prevoicing due to the influence of Greek voiced stops, the stress effects show that this drift toward the Greek norm does not necessarily imply a shift of stress effect toward the phonetic category {voiced}.

To summarize, the effect of stress on the realization of voicing contrast in heritage speakers' speech is vastly understudied. The results from two such previous studies are somewhat unexpected; Greek and Spanish heritage speakers in an English-speaking environment produce their English voiced stops with more prevoicing than English monolingual controls, suggesting a shift in their phonetic target to {voiced} but in their speech, stress reduces prevoicing, approximating a target of {vl.unasp}, rather than lengthens prevoicing which would have indicated a phonetic target of {voiced}. In this study, we examine the effect of stress on stop realization in a Tagalog heritage speaker population's speech, and if and how the stress effects can further illuminate the structural organization of these bilingual speakers' stop categories.

Tagalog
Tagalog is an Austronesian language that forms the structural basis of Filipino, the national language of the Philippines, spoken by at least 84% of the population. Along with Filipino, English is also an official language of the country and is spoken by 56% of the population (Gonzalez, 1998).[9] According to the 2011 census, there are about 491,075 Tagalog speakers in Canada (Statistics Canada, 2012). In Toronto, Tagalog is the fourth largest heritage language, with 179,975 speakers as of 2011 (Statistics Canada, 2012). Tagalog has three voiced stops, /b d g/, and three voiceless stops, /p t k/ (Ramos, 1971; Rubino & Llenado, 2002; Schachter & Otanes, 1972).[10] Tagalog /d/ and /t/ are dental, thus contrasting with English alveolar /d/ and /t/, and Tagalog /k/ is slightly more retracted than English /k/. Tagalog /k/ is also often realized as a velar fricative, [x], especially between low and back vowels (Schachter & Otanes, 1972). Voiceless stops are described as unaspirated, suggesting that Tagalog employs a short-lag versus voicing lead contrast, unlike English. Stops are also usually unreleased in word-final position. Most descriptions of Tagalog consider stress (or accent) to be contrastive in the language, where primary stress falls on either of the final two syllables of a word (French, 1991; Gonzalez, 1970; Ramos, 1971; Rubino & Llenado, 2002). Stress is signaled by a combination of length, amplitude and a rise in pitch (Gonzalez, 1970). With this background, we will now turn to our methods, results and the implication of our results.

METHODOLOGY
Participants
Nine heritage Tagalog speakers (four females and five males, ages 19-26) were recruited from the University of Toronto Scarborough campus.[11] Twelve native English speakers (six females and six males, ages 18-32) and ten native Tagalog speakers (five females and five males, ages 19-24) also participated in the study as comparison groups. All speakers reported normal speech, hearing, and vision and were paid for their participation. The heritage speakers are self-identified speakers of Tagalog whose parents are native speakers of Tagalog.[12] Seven were born in Canada and two came to Canada before school age.[13] Additionally, each heritage speaker stated that he/she had not lived in the Philippines for longer than a combined period of six months and had no formal schooling in Tagalog. Most speakers gave a higher communicative confidence rating for English (mean = 4.9, SD = 0.3) than for Tagalog (mean = 2.8, SD = 1.1). The native Tagalog speakers all lived in the Philippines,[14] specifically in the Metro Manila area, at least until the age of 15, and they grew up in households where both parents were native speakers of Tagalog.[15] They all had come to Canada within the previous five years of the time of recording and rated their confidence in Tagalog as five out of five. All native Tagalog speakers were also proficient in English (mean = 4.3, SD = 0.7). Participants for the native English control group were all born and raised in Canada and reported English as their home language. None of the English control speakers reported knowledge of Tagalog.

Materials and Procedure
The stimuli consist of 36 bisyllabic Tagalog words and 36 bisyllabic English words. The vowels that flanked the target stop were a low back vowel or a schwa where possible. The words were chosen to represent six stop consonants (/b d g p t k/) in three positions (i.e., word-initial, word-medial intervocalic and word-final) with two different stress placements (i.e., initial or final). For

the current paper, we report the analysis of the word-initial and word-medial stops. The full list of words with initial and medial stops is provided in Appendix A. The experimenter for the heritage group and the Tagalog group was a heritage Tagalog speaker and the experimenter for the English group was an English speaker without any Tagalog background. The heritage group and the Tagalog control group produced the words in the two lists in separate blocks (Tagalog followed by English), with a break between the blocks, while the English control group produced English words only.[16] Prior to the production of each word list, speakers were given time to familiarize themselves with the word list. If participants were unsure of the pronunciations of certain words, the experimenter would clarify and provide the intended pronunciation. Participants were also given a short passage in Tagalog to read to ensure that they were comfortable with the reading task and were in the relevant language mode when they began the word reading task.

Each word list was presented to speakers in a randomized order and each word was produced in isolation with three repetitions. Participants were encouraged not to use a list intonation and to produce each repetition as a separate utterance. After discarding tokens with excessive background noise, signal errors, incorrect pronunciations and/or disfluencies, a total of 2,769 tokens of word-initial and word-medial stops from the two languages, produced by the three speaker groups, were analyzed for this study. Recordings were made using an AT831B microphone and a Zoom H4N digital voice recorder at 44.1 kHz and 16-bit. After the main production task, participants completed a language background questionnaire (See APPENDIX B), which gathered information about each participant's linguistic background and proficiency in their relevant language(s).

Data Analysis

Sound recordings were segmented into separate files that contained individual word tokens and were analyzed using Praat (Boersma & Weenink, 2015). Files were manually annotated for the acoustic events listed in (1) based on a visual inspection of waveforms and spectrograms.

(1) Annotation labels of acoustic analyses

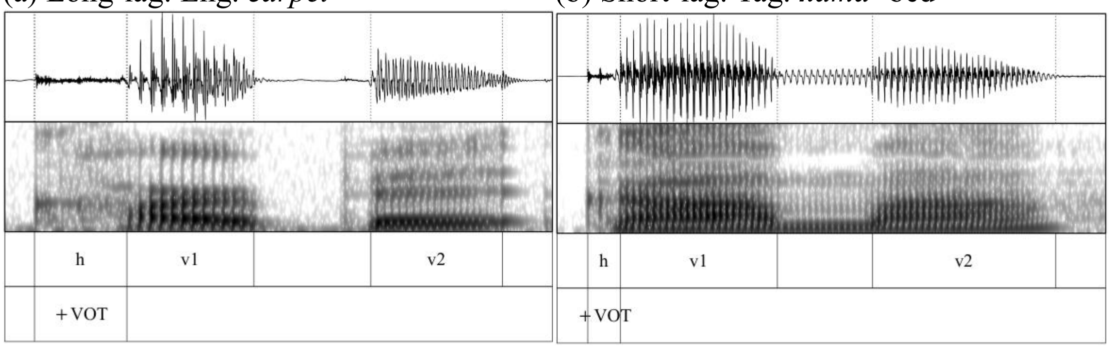
- Voiced closure (cv)
- Voiceless closure (cl)
- Voicing lag (h)
- First vowel (v1)
- Second vowel (v2)

Voiced closure marked the visible periodicity in the waveform during a stop closure. The portion of a stop closure without voicing was labeled as *voiceless closure*. As words were produced in isolation, without a carrier phrase, it was not always possible to identify the onset of a voiceless closure for word-initial stops. Therefore, voiceless closure was labeled for initial stops only when it was preceded by a voiced closure. *Voicing lag* was identified as the time from the release of a stop to the onset of voicing in the following vowel. In each word, the *first vowel* and the *second vowel* were labeled after identifying their respective onset and offset, based on the presence of higher formants (F2 and up) and voicing.

From these annotations, VOT measurements were extracted. However, many mixed voicing configurations were attested in our data, where lead voicing (i.e., voicing during a stop closure) and voicing lag (i.e., devoicing of a vowel at the release) coexisted in a single stop token or where there was residual voicing from the preceding vowel into the stop closure (Davidson, 2016; Flege & Brown, 1982; Keating, 1984). In such cases of mixed voicing, VOT, defined as a time lag between the stop release and the onset of voicing, did not always provide an accurate measure of voicing patterns. The general principle we follow is that for phonemically voiceless stops, which generally have a positive VOT value (with English post-tonic /t/ being an exception), the traditional VOT is used as a measure of voicelessness/aspiration. Examples of word-initial long-lag and short-lag stops are provided in Figure 2.

Figure 2. Examples of Word-initial Voiceless Stops: (a) Long-lag; (b) Short-lag

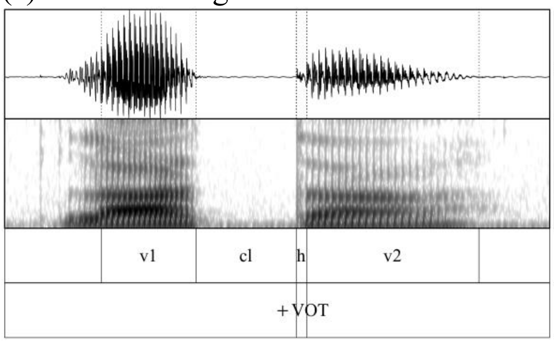
(a) Long-lag: Eng. ***cárpet*** (b) Short-lag: Tag. ***káma*** 'bed'



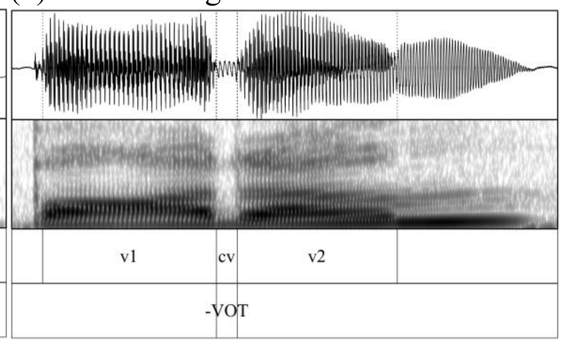
Medial voiceless stops can be categorized into three types with respect to closure voicing. The first type is a case of no closure voicing, illustrated by the example in Figure 3(a), where VOT can be straightforwardly measured as an interval between the stop release and the voicing onset in the following vowel. The second type is a case of complete voicing found in English /t/ lenition, where /t/ is reduced to a voiced tap [ɾ]. Also, one Tagalog native male speaker produced /k/ in *taká* ('to wonder') as a voiced approximant. For these stops, we treat the duration of the consonantal constriction, which is fully voiced, as voicing lead, as illustrated in (Figure 3(b)). The final type is a case of voicing leak, where there is residual voicing from the preceding vowel into the initial portion of stop closure, with the voicing tapering off before the stop release in most cases (Flege & Brown, 1982; Keating, 1984), as shown in Figure 3(c). This type of voicing leak is found in 49.5% of the medial voiceless stops in our data, distributed across all four data sets,[17] and the majority of these tokens are also accompanied by voicing lag at the stop release. For our VOT measurements, we did not take this type of voicing into account, but rather treated it as a property of the preceding vowel. Finally, a few Tagalog speakers produced some tokens of intervocalic /k/ as [x] and these tokens were excluded from our VOT calculations.[18]

Figure 3. Examples of Word-medial Voiceless Stops: (a) Voiceless; (b) Lenited; (c) Voicing leak

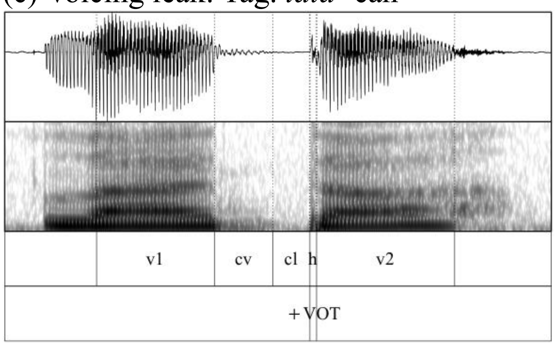
(a) Voiceless: Tag. *láta* 'can'



(b) Lenited: Eng. *bóttom*

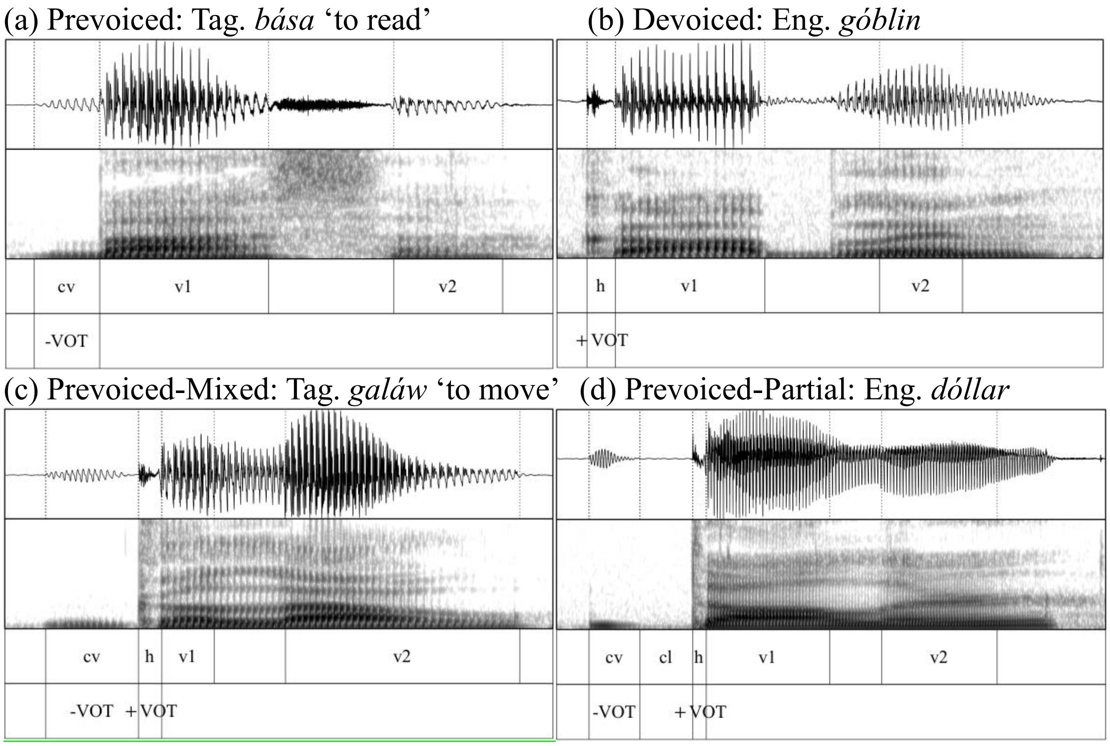


(c) Voicing leak: Tag. *láta* 'can'



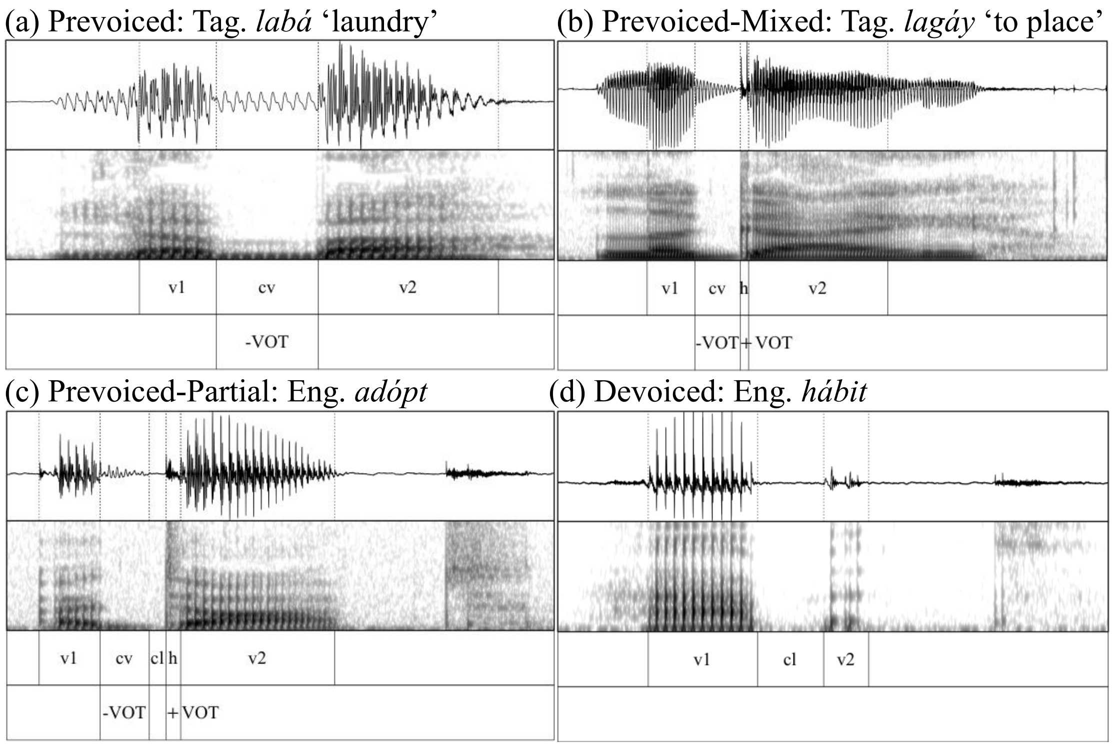
Turning to phonemically voiced stops, whose VOT values may be negative or positive, we quantified the overall degree of voicing by taking into account both voicing lead and voicing lag. For voiced stops in word-initial position, a number of different voicing configurations are attested. The first of the four voicing types is a straightforward case of prevoiced stops, where voicing starts sometime before the release of the stop constriction and is sustained through the release and into the following vowel. This pattern is produced mostly by the Tagalog control group. An example of a *prevoiced* stop is shown in Figure 4(a). The second pattern is a fully *devoiced* stop, where there is no prevoicing during the closure and voicing starts at or after the stop release, as shown in Figure 4(b). The majority of initial voiced stops produced by the English control group fall into this category. For these first two types of voiced stops, VOT measurements are straightforward. In addition, there are voicing configurations where both prevoicing and short-lag VOT are found in the same stop. In these tokens, prevoicing may be sustained throughout the stop closure, as shown in Figure 4(c) (i.e., *Prevoiced-Mixed*) or voicing tapers off during the closure, as shown in Figure 4(d) (i.e., *Prevoiced-Partial*). This last type of prevoicing is almost exclusively found in heritage speakers' speech.[19] For these cases, we calculated the VOT by adding up the negative of the duration of closure voicing and the positive duration of voicing lag. For example, we calculated the VOT of the initial stop in Figure 4(d) (-51.2 ms) by adding the negative of the duration of closure voicing (-70.1 ms) and the positive duration of voicing lag (18.9 ms). We will refer to this combined measurement of VOT as *adjusted* VOT in order to differentiate it from traditional VOT.[20]

Figure 4. Examples of Word-initial Voiced Stops: (a) Prevoiced; (b) Devoiced; (c) Prevoiced-Mixed; (d) Prevoiced-Partial



Word-medial voiced stops are categorized into the same four voicing types and the degree of voicing was quantified the same way as word-initial voiced stops. Figure 5 provides examples of each of the four voicing configurations attested for word-medial voiced stops. We did not find any clear case of prevoicing where voicing begins in the middle of a closure and is sustained into the stop release, which replicates the findings for English voiced stops by Davidson (2016). The fourth type, *devoiced*, was rare and produced by a single male English speaker.[21]

Figure 5. Examples of Word-medial Voiced Stops: (a) Prevoiced; (b) Prevoiced-Mixed; (c) Prevoiced-Partial; (d) Devoiced



Statistical Analysis

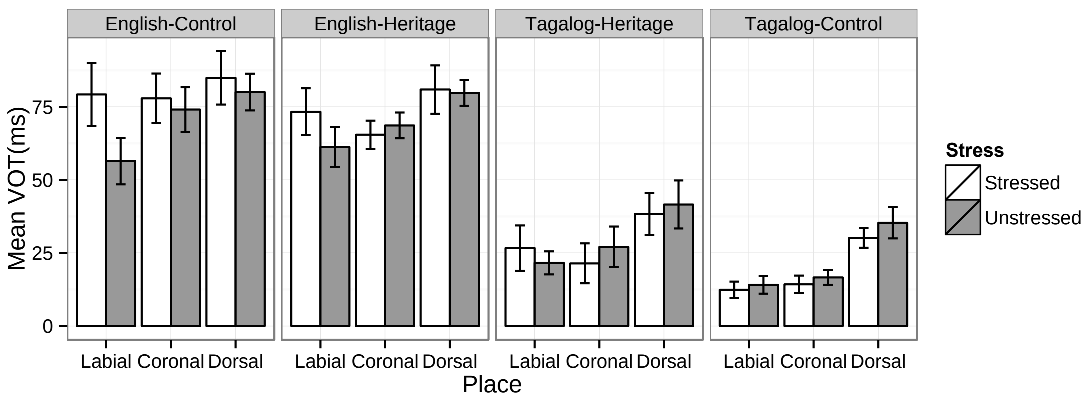
For statistical analyses, we examine voiced and voiceless stops in word-initial and word-final positions separately. For each condition, we test how the degree of voicing and the effect of stress differ by the speaker group using linear mixed effects models (Baayen, Davidson, & Bates, 2008). The statistical analyses are conducted in R (R Development Core Team, 2015) and the *lmer* function of the *lme4* package (Bates et al., 2015) is used. The fixed effects factors include GROUP, STRESS and PLACE as well as their interactions. PLACE of articulation is included because VOT is known to vary systematically as a function of place of articulation (Cho & Ladefoged, 1999) and to interact with the stress-conditioned lenition of word-medial stops in English (i.e., /t/ tapping). Simple coding was used for GROUP and STRESS, and PLACE was backward difference coded to compare labial versus coronal and then coronal versus dorsal. The random effects included a random intercept for *speaker* and a by-speaker random slope for STRESS, PLACE and their interaction. P-values are determined by the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2015). Follow up post-hoc comparisons are conducted using the *testInteractions* function of the *phia* package (De Rosario-Martinez, Fox, & R Development Core Team, 2015). The alpha level for significant tests was set at 0.05 and the threshold for marginal significance was set at 0.1.

RESULTS

Voiceless Stops: Word-initial Position

Voiceless stops in word-initial position, which was also utterance-initial position in our data, were consistently produced without prevoicing. Figure 6 summarizes the mean VOT values of word-initial voiceless stops for the four language-speaker groups based on stress and place of articulation. Table 4 presents the output (the β coefficients and p values only) of three linear mixed effects models.

Figure 6. Mean VOTs (in milliseconds, ms) of Word-initial Voiceless Stops (Error bars indicate 95% confidence intervals)



The first model in Table 4 compares the data from the two control groups: native speakers of English and Tagalog. The second and the third models compare the two speaker groups (i.e., heritage versus non-heritage) within each language. As our main concern is the comparisons across speaker groups, we will focus on the GROUP effect and its interaction with other factors in our discussion. These crucial factors are highlighted in the model output tables.

In the first model, which compares the two control groups, the main effect of GROUP is significant; English voiceless stops have significantly longer VOT values than Tagalog voiceless stops ($p < 0.001$). This is expected due to the fact that English is an aspirating language with long-lag voiceless stops, while Tagalog is a true voicing language with short-lag voiceless stops. The interaction of GROUP * STRESS is significant and so is the three-way interaction of GROUP*STRESS*PLACE. Post-hoc comparisons of the stress effect by each place of articulation of each language shows that the stress effect is significant only for English /p/ ($p < 0.001$); VOT is longer in stressed than unstressed syllables, in line with previous observations that stress has an additive effect on the VOT of voiceless stops in English. As can be seen in Figure 6, we find a trend in the same direction, but no significant effect, for /t/ and /k/ in English. We find no effect of stress for Tagalog voiceless stops. This is also in line with the previous observation of a lack of a consistent stress effect on short-lag stops.

Now we turn to the other two models, which test how closely heritage speakers approximate the native norms of their two languages. For the English model, there is no significant effect of speaker group ($p = 0.627$), indicating that heritage speakers produce English voiceless stops with native-like VOT values. There is no significant interaction of STRESS*GROUP or

STRESS*PLACE*GROUP, indicating that the two English speaker groups do not differ in their stress effects. When we compare the Tagalog productions of the two relevant speaker groups, we find that there is a marginally significant effect of GROUP (p = 0.069); heritage speakers produce Tagalog voiceless stops with longer VOT. There is no significant main effect of STRESS and there is no significant higher-level interaction, which means that neither speaker group demonstrates any stress effect.

Table 4.
Linear mixed effects models of VOT for word-initial voiceless stops. The reference level of each comparison is underlined.

	Control: Tagalog vs. <u>English</u>		English: Heritage vs. <u>Control</u>		Tagalog: Control vs. <u>Heritage</u>	
	β	*p*	*β*	*p*	*β*	*P*
Intercept	48.0	<0.001 ***	73.5	<0.001 ***	24.6	<0.001 ***
GROUP	-55.0	<0.001 ***	-3.8	0.627	-8.3	0.069 .
STRESS (unstressed vs. <u>stressed</u>)	-3.7	0.093 .	-7.3	0.011 *	2.3	0.109
PLACE (coronal vs. <u>labial</u>)	5.2	0.042 *	3.6	0.265	1.1	0.399
PLACE (dorsal vs. <u>coronal</u>)	11.9	<0.001 ***	10.2	0.001 **	17.0	<0.001 ***
GROUP*STRESS	13.6	0.004 **	6.4	0.229	1.6	0.573
GROUP*PLACE (cor vs. <u>lab</u>)	-5.9	0.230	-9.2	0.159	2.2	0.403
GROUP*PLACE (dor vs. <u>cor</u>)	10.8	0.022 *	7.4	0.146	0.6	0.899
STRESS*PLACE (cor vs. <u>lab</u>)	9.8	0.008 **	17.5	0.001 **	5.1	0.088 .
STRESS*PLACE (dor vs. <u>cor</u>)	0.9	0.730	-2.1	0.552	0.4	0.917
GROUP*STRESS*PLACE (cor vs. <u>lab</u>)	-18.3	0.012 *	-2.8	0.752	-9.0	0.129
GROUP*STRESS*PLACE (dor vs. <u>cor</u>)	3.8	0.470	-2.1	0.759	4.9	0.493

Significance codes: <0.001 '***'; <0.01 '**'; <0.05 '*'; <0.1 '.'

Voiceless Stops: Word-medial Position

The results for voiceless stops in word-medial position are summarized in Figure 7 and Table 5. The control group comparison shows a significant main effect of GROUP (p < 0.001); as in word-initial position, English voiceless stops have longer VOT values than Tagalog voiceless stops. The interactions of GROUP * STRESS and GROUP*STRESS*PLACE are also significant. Post-hoc comparisons confirm the patterns in Figure 7; the stress effect is significant for all English stops, with unstressed stops showing substantial VOT reduction (p < 0.001). On the other hand, there is no effect of stress on Tagalog stops. The significant three-way interaction is due to the further lenition of coronal stops (i.e., tapping) in English.

The second model compares the two groups of English speakers. In this comparison, we find a marginally significant effect of GROUP (β = -7.5, p = 0.095), where heritage speakers produce a slightly shorter VOT in medial stops. There is no significant interaction of GROUP*STRESS, indicating that the strong stress effect is found for heritage speakers' English as well. The third model compares the two groups of Tagalog speakers and finds a marginally significant effect of GROUP (p = 0.082), with Tagalog control speakers producing slightly shorter VOT values than heritage speakers.[22] There is also a significant effect of STRESS (p = 0.021), indicating that overall, stress reduces VOT, which is the opposite of the stress effect in English voiceless stops. The lack of an interaction between GROUP*STRESS indicates that the two Tagalog groups produce their medial voiceless stops comparably.

Figure 7. Mean VOT (ms) of Word-medial Voiceless Stops by Stress Condition and Language Group (Error bars indicate 95% confidence intervals)

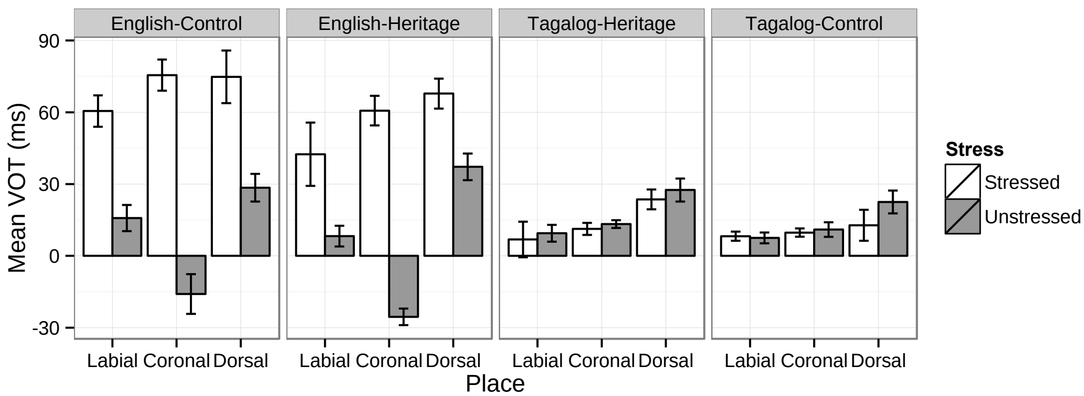


Table 5.

Linear regression models of VOT in word-medial voiceless stops. The reference level of each comparison is underlined.

	Control: Tagalog vs. <u>English</u>		English: Heritage vs. <u>Control</u>		Tagalog: Control vs. <u>Heritage</u>	
	β	*p*	*β*	*p*	*β*	*p*
Intercept	25.9	<0.001 ***	36.1	<0.001 ***	13.4	<0.001 ***
GROUP	-28.0	<0.001 ***	-7.5	0.095 .	-4.3	0.082 .
STRESS (unstressed vs. <u>stressed</u>)	-28.8	<0.001 ***	-55.8	<0.001 ***	2.2	0.021 *
PLACE (coronal vs. <u>labial</u>)	-2.9	0.358	-9.0	0.028 *	3.3	0.015 *
PLACE (dorsal vs. <u>coronal</u>)	14.3	0.001 **	28.2	<0.001 ***	9.1	0.011 *
GROUP*STRESS	64.1	<0.001 ***	10.0	0.282	-0.9	0.624
GROUP*PLACE (cor vs. <u>lab</u>)	10.9	0.096 .	-1.3	0.864	-1.6	0.523
GROUP*PLACE (dor vs. <u>cor</u>)	-15.2	0.046 *	12.8	0.089 .	-8.8	0.171
STRESS*PLACE (cor vs. <u>lab</u>)	-22.4	<0.001 ***	-48.3	<0.001 ***	0.7	0.764
STRESS*PLACE (dor vs. <u>cor</u>)	26.5	<0.001 ***	50.6	<0.001 ***	2.5	0.367
GROUP*STRESS*PLACE (cor vs. <u>lab</u>)	48.7	<0.001 ***	-3.2	0.771	2.6	0.581
GROUP*STRESS*PLACE (dor vs. <u>cor</u>)	-37.2	0.005 **	11.0	0.357	2.0	0.707

Significance codes: <0.001 '***'; <0.01 '**' ; <0.05 '*' ; <0.1 '.'

Voiceless Stops: Summary

To summarize, heritage speakers' voiceless stops closely approximated the native speaker norms of their two languages. Stress effects (or a lack thereof) in their voiceless stop production also mirrored their native comparison groups. While we found evidence of minor influence between the voiceless stops of the two languages in question in terms of absolute VOT values, the two languages of heritage speakers faithfully mirrored the respective native comparisons, including stress effects.

Voiced Stops: Word-initial Position

Turning to voiced stops in word-initial position, the distribution of voicing configurations across the four data sets is summarized in Figure 8. The English control group produced only a handful of prevoiced stops (10/212, 4.7%), while the Tagalog control group produced the majority of their word-initial voiced stops with some type of prevoicing (154/180, 85.6%). Heritage speakers' English and Tagalog mirror the cross-language difference; they produced more prevoiced stops in Tagalog (70/147, 47.6%) than in English (31/160, 19.4%). However, these prevoicing rates differ substantially from the respective native comparisons, indicating cross-language influence in both directions. The effect of stress on voicing type is marginal for both groups of English speakers and a comparison of the Tagalog groups shows trends in opposite directions; for Tagalog control speakers, stress increases prevoicing, while for heritage speakers, stress reduces prevoicing. This overall pattern holds true when the adjusted VOT values are compared through statistical tests. Figure 9 and Table 6 present a summary graph and the statistical models' outputs, respectively.

The model for the control group comparison shows a significant main effect of GROUP ($p < 0.001$). As expected, Tagalog voiced stops are produced with substantial prevoicing, while the voiced stops of English are produced with short-lag VOT. There is no effect of STRESS or an interaction between STRESS*GROUP, indicating that there is no effect of stress on the word-initial voiced stops of either language. A comparison of the two groups of English speakers shows a marginal effect of GROUP ($p = 0.074$), with shorter prevoicing for the heritage group than for the control group. Again, no effect of STRESS or a STRESS*GROUP interaction are found. A comparison of the two Tagalog speaker groups shows a significant main effect of GROUP ($p = 0.007$); heritage speakers produce stops with shorter prevoicing than the control group. We also find a marginally significant interaction of STRESS *GROUP ($p = 0.095$). A post-hoc comparison indicates a marginally significant effect of STRESS for the heritage group ($p = 0.067$); voiced stops have an increased VOT (i.e., less prevoicing) in stressed syllables. However, there is no effect of stress in the Tagalog control group. Note that the stress effect in heritage Tagalog is the opposite of what we expect for the voiced stops of a true voicing language.

Figure 8. Proportion of Voicing Types for Word-initial Voiced Stops by Stress Condition and Language Group

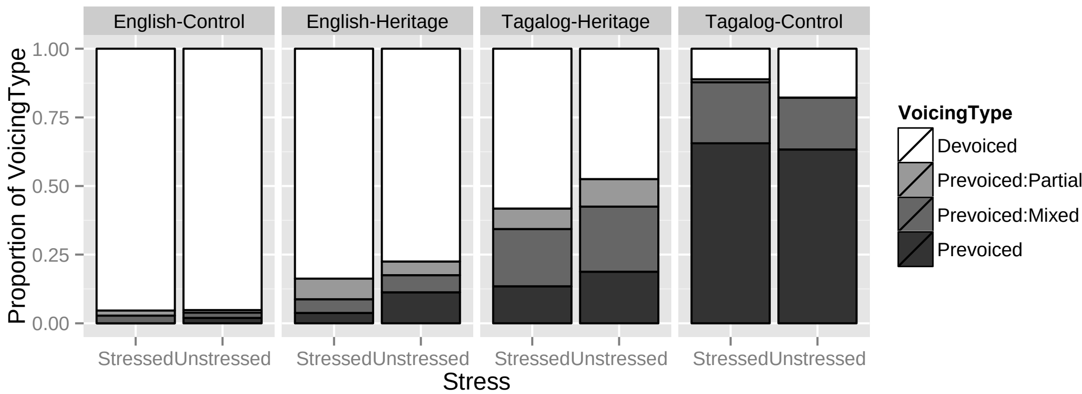


Figure 9. Mean Adjusted VOT (ms) of Word-initial Voiced Stops (Error bars indicate 95% confidence intervals)

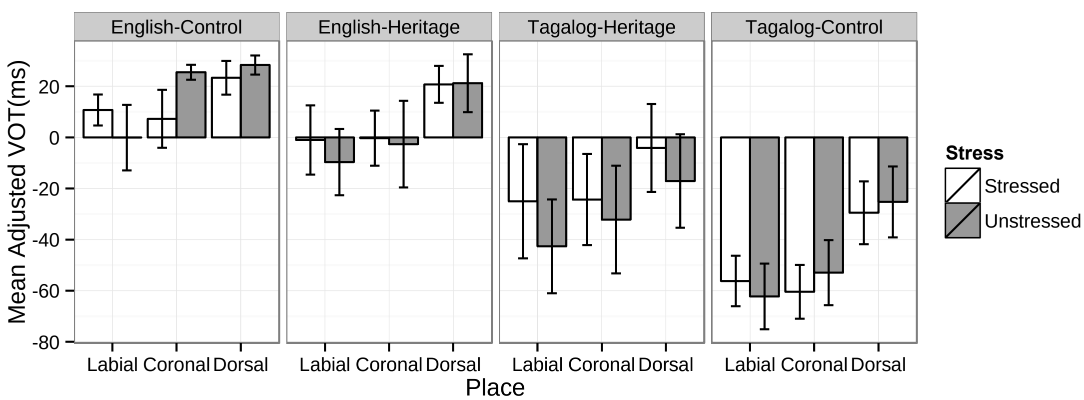


Table 6.

Linear mixed effects models of VOT in word-initial voiced stops. The reference level of each comparison is underlined.

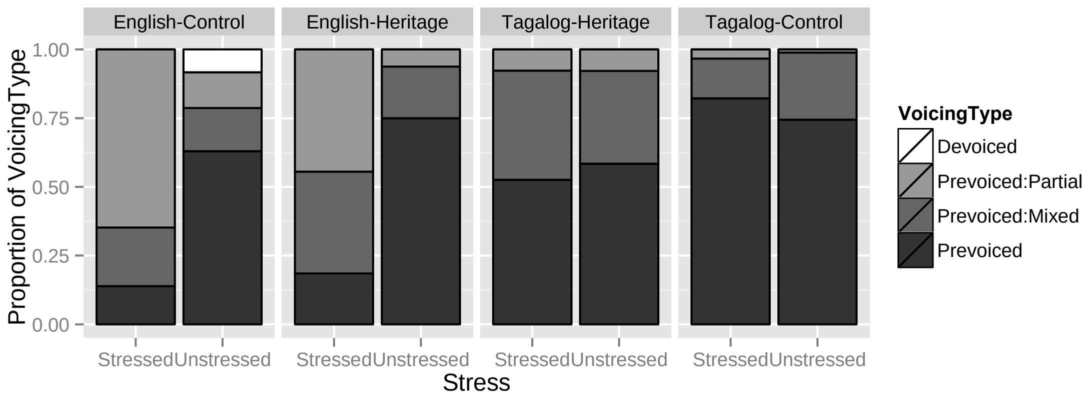
	Control: Tagalog vs. English		English: Heritage vs. Control		Tagalog: Control vs. Heritage	
	β	*p*	*Β*	*P*	*β*	*P*
Intercept	-15.9	<0.001 ***	10.3	0.003 **	-35.3	<0.001 ***
GROUP	-63.8	<0.001 ***	-11.2	0.074	-24.9	0.007 **
STRESS (unstressed vs. stressed)	3.2	0.307	0.6	0.882	-6.9	0.185
PLACE (coronal vs. labial)	6.5	0.193	7.5	0.087	2.5	0.707
PLACE (dorsal vs. coronal)	19.4	<0.001 ***	15.9	<0.001 ***	23.9	0.001 **
GROUP*STRESS	-2.6	0.678	-7.2	0.362	17.7	0.095 .
GROUP*PLACE (cor vs. lab)	-8.0	0.421	-6.9	0.415	0.0	0.999
GROUP*PLACE (dor vs. cor)	19.9	0.020 *	13.1	0.060	10.8	0.400
STRESS*PLACE (cor vs. lab)	20.7	0.029 *	16.8	0.129	15.2	0.180
STRESS*PLACE (dor vs. cor)	-8.2	0.269	-5.3	0.491	-5.1	0.712
GROUP*STRESS*PLACE (cor vs. lab)	-14.4	0.423	-24.1	0.269	-3.3	0.880
GROUP*STRESS*PLACE (dor vs. cor)	10.0	0.498	16.2	0.296	3.6	0.894

Significance codes: <0.001 '***'; <0.01 '**'; <0.05 '*' ; <0.1 '.'

Voiced Stops: Word-medial Position

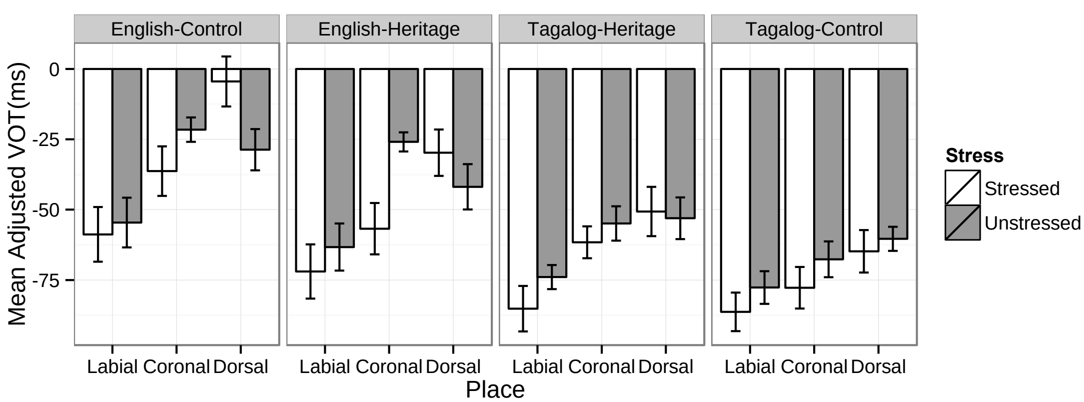
The breakdown of voicing types for word-medial voiced stops is summarized in Figure 10. The most notable difference across the data sets is found in the proportion of tokens with full closure voicing, with or without some devoicing at the stop release (prevoiced + prevoiced- Mixed), which is represented by the two darker shades of gray in the graphs of Figure 10. Tagalog control speakers produce 97.8% of medial voiced stops with full closure voicing (174/180). The rate is only 56.9% (123/216) for English control speakers. The heritage speakers produce their Tagalog stops with a higher rate of full closure voicing (143/155, 92.3%) than English stops (120/161, 74.5%), but these rates differ from those of their respective comparison groups, suggesting bidirectional influence between the two languages. Another striking difference between English and Tagalog is that stress has a strong effect in English but does not show a clear effect inTagalog, presumably due to a ceiling effect. These generalizations hold true for heritage speakers as well as the monolingual groups.

Figure 10. Proportion of Voicing Types for Word-medial Voiced Stops by Stress Condition and Language Group



The adjusted VOT values, summarized in Figure 11, reflect the four-way difference of voicing pattern observed in Figure 10; Tagalog control speakers produce the longest voicing, English control speakers produce the shortest and heritage speakers produce intermediate degrees of voicing for their two languages (English control: -34.1 ms; English heritage: -48.2 ms; Tagalog heritage: -62.8 ms; Tagalog control: -72.4 ms). However, we find that the adjusted VOT values fail to quantify the stress effect observed for English; English voiced stops in unstressed syllables are more likely to be fully voiced, but the adjusted VOT values indicate a shorter prevoicing duration in unstressed position.

Figure 11. Mean Adjusted VOT (ms) of Word-medial Voiced Stops (Error bars indicate 95% confidence intervals)



A closer inspection of the data shows that this discrepancy is due to the fact that stress lengthens the closure duration of voiced stops (post-tonic voiced stops are shorter than pre-tonic voiced stops) and this effect is particularly strong for English /d/, which frequently lenites to [ɾ].[23] Given that the duration of prevoicing is bound by the duration of stop closure, the absolute voicing duration fails to reflect the qualitative difference in voicing. For example, two fully voiced stops can differ greatly in their VOT values, depending on closure duration; in addition, a partially voiced stop with a long closure duration can have a longer interval of prevoicing than a fully voiced and lenited stop such as [ɾ]. Instead, we used a ratio of voicing relative to closure duration. Specifically, we calculated this ratio by subtracting the duration of voicing lag from the closure voicing duration and dividing it by the total stop constriction duration: (cv–h)/(cv+cl). Essentially, this is the same measure as adjusted VOT, except that it is normalized by closure duration and the sign is reversed. Figure 12 and Table 7 provide a summary graph and the statistical models' outputs.[24]

The model for the control group comparison shows a significant main effect of GROUP ($p < 0.001$), with a higher closure voicing ratio for Tagalog than for English. In other words, Tagalog voiced stops are more likely to be fully voiced than English voiced stops. There is a significant interaction of STRESS*GROUP and a post-hoc comparison shows that for English, stress decreases voicing ratios ($p < 0.001$), while there is no such significant effect of stress for Tagalog. This is expected given that for Tagalog, 97.8% of stops are produced with full voicing during closure and there is very little variation in their voicing ratio. Within-language comparisons show a significant difference ($p = 0.039$) between the English groups and a small but significant difference ($p = 0.0496$) between the Tagalog groups. No significant interaction of STRESS*GROUP is found in either model, indicating that STRESS effects in heritage speech do not differ significantly from their respective controls.

Figure 12. Mean Adjusted Closure Voicing Ratio of Word-medial Voiced Stops (Error bars indicate 95% confidence intervals)

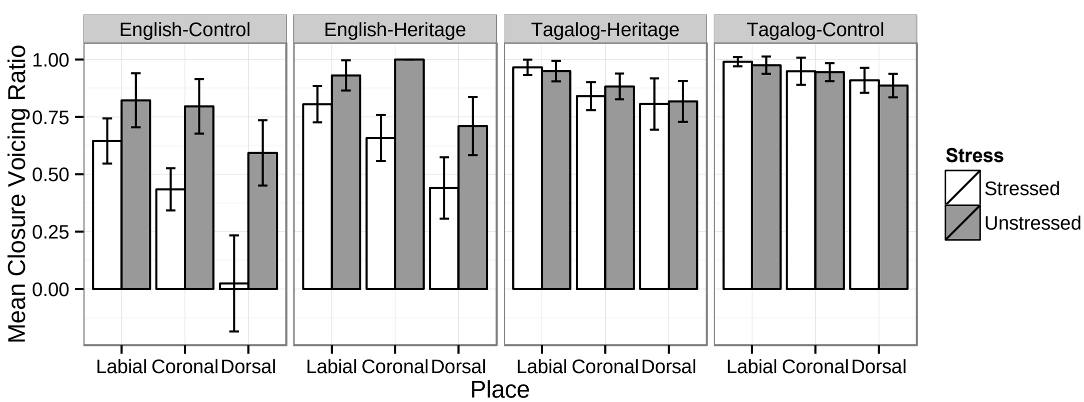


Table 7.

Linear mixed effects models of adjusted closure voicing ratio in word-medial voiced stops. The reference level of each comparison is underlined.

	Control: Tagalog vs. English		English: Heritage vs. Control		Tagalog: Control vs. Heritage	
	β	*p*	*β*	*p*	*β*	*p*
Intercept	0.748	<0.001 ***	0.655	<0.001 ***	0.909	<0.001 ***
GROUP	0.390	<0.001 ***	0.205	0.039 *	0.066	0.0496 *
STRESS (unstressed vs. stressed)	0.178	<0.001 ***	0.308	<0.001 ***	-0.003	0.832
PLACE (coronal vs. labial)	-0.077	0.023 *	-0.079	0.032 *	-0.067	0.003 **
PLACE (dorsal vs. coronal)	-0.177	0.001 **	-0.280	<0.001 ***	-0.052	0.144
GROUP*STRESS	-0.383	<0.001 ***	-0.123	0.184	-0.021	0.512
GROUP*PLACE (cor vs. lab)	0.083	0.206	0.079	0.270	0.063	0.136
GROUP*PLACE (dor vs. cor)	0.257	0.014 *	0.052	0.664	0.007	0.924
STRESS*PLACE (cor vs. lab)	0.098	0.068 .	0.200	0.001 **	0.037	0.275
STRESS*PLACE (dor vs. cor)	0.094	0.425	0.067	0.624	-0.03	0.367
GROUP*STRESS*PLACE (cor vs. lab)	-0.173	0.104	0.030	0.789	-0.052	0.447
GROUP*STRESS*PLACE (dor vs. cor)	-0.226	0.339	-0.279	0.316	0.023	0.732

Significance codes: <0.001 ‘***’; <0.01 ‘**’ ; <0.05 ‘*’ ; <0.1 ‘.’

Voiced Stops: Summary

As with voiceless stops, heritage speakers produced the voiced stops of their two languages distinctly; Tagalog voiced stops were produced with more prevoicing than English voiced stops. However, heritage speakers’ voiced stops differed from native norms and showed a systematic pull toward the voiced stops of the other language; that is, heritage speakers’ English voiced stops were more prevoiced (i.e., more Tagalog-like) and their Tagalog voiced stops were less prevoiced (i.e., more English-like) compared to their respective native comparisons. For English, the effect of stress (or lack thereof) in the native norm was matched in heritage speakers’ speech. On the other hand, for Tagalog, a divergent stress effect emerged in word-initial voiced stops; Tagalog control speakers showed no statistically significant stress effect, while heritage Tagalog speakers produced less prevoicing in stressed than in unstressed positions. The direction of this stress effect is unexpected for the voiced stops of a true voicing language. In fact, it mirrors the stress effect of English short-lag target voiced stops.

Discussion and Conclusions

Now we can take the findings from the voiced and voiceless stops together and address our main questions. First, we examined if and how cross-language influence affects the two stop types differently in the two languages of our heritage bilinguals. Table 8 summarizes whether the bilinguals’ production matched that of native speaker controls for each stop type in each word position. We can see a clear difference between voiced and voiceless stops. While bilinguals’ voiceless stops matched or nearly matched the native comparison groups’ voiceless stops, their voiced stops generally did not match native norms. This asymmetry held true not only in English, but also in Tagalog.

Table 8.

Cross-language influence in heritage speakers' English and Tagalog
(✓: no difference from the native pattern, (✓): marginally significant difference from the native pattern, ×: significant difference from the native pattern)

		English	**Tagalog**
Voiceless	*Initial*	✓	(✓)
	Medial	(✓)	(✓)
Voiced	*Initial*	(✓)	× (less prevoicing)
	Medial	× (more prevoicing)	× (less prevoicing)

Recall that the influence hierarchy generated by the phonological markedness based account (Table 1(a)) predicts stronger influence for voiced stops in English, but stronger influence for voiceless stops in Tagalog. The phonetic category based account (Table 1(b)), on the other hand, predicts stronger influence for voiced stops in both languages. Our findings, therefore, support the phonetic account over the phonological account. Recall, however, that the phonetic account makes a particular assumption about the category structure of cross-language stop consonants, namely, that the voiced stops of the two languages are categorized into a single phonetic category (i.e., {voiced}), while the voiceless stops of the two languages retain their distinct categories: {vl.unasp.} for Tagalog voiceless stops and {vl.asp.} for English voiceless stops. This asymmetry in cross-language sound structure is the reason for the asymmetrical influence pattern.

Now we turn to the effect of lexical stress on voicing realization, which is summarized in Table 9. For voiceless stops, heritage speakers not only produced comparable VOT values, but also produced the same stress effects as their monolingual counterparts, both for English and for Tagalog, in both initial and medial positions. Stress lengthens the VOT of English voiceless stops, which enhances the voicing contrast as well as the phonetic feature {vl.asp.} (i.e., {+spread glottis}, specifically). The stress effect on Tagalog voiceless stops is only significant in medial position and is more muted than in English. Stress shortens the VOT of Tagalog stops and enhances the phonetic feature of {vl.unasp.} (i.e., {-spread glottis}) at the expense of reducing the paradigmatic contrast against voiced stops. In other words, it seems that the voiceless stops of the two languages are robust against cross-language influence in terms of absolute VOT values as well as stress conditioned variation, and that speakers have established separate phonetic categories in their two languages: {vl.unasp.} for Tagalog voiceless stops and {vl.asp.} for English voiceless stops.

Table 9.

Stress Effects on Voicing Realization

		English-Control	English-Heritage	Tagalog-Heritage	Tagalog-Control
Voiceless	*initial*	longer VOT	longer VOT	no effect	no effect
	medial	longer VOT	longer VOT	shorter VOT	shorter VOT
Voiced	*initial*	no effect	no effect	***less prevoicing***	no effect
	medial	less prevoicing	less prevoicing	no effect	no effect

For voiced stops, on the other hand, the picture is more complicated. The English control group produced a statistically significant stress effect in medial position (i.e., stress reduced prevoicing), which enhances the {vl.unasp.} target of this short-lag stop category. The Tagalog voiced stops of the control group, on the other hand, did not show any statistically significant effect of stress. Recall that we found substantial cross-language influence between the voiced stops of the heritage speakers' two languages (see Table 8). According to the phonetic category based account of this influence (Table 1(a)), this selective divergence of voiced stops arises from the fact that English and Tagalog voiced stops are more likely to form a single phonetic category of {voiced} stops, as the {vl.unasp.} category is already taken up by the voiceless stops of Tagalog. Based on this view, we predict that the voiced stops of both languages should show the stress effect applicable to {voiced} stops, whereby stress should increase prevoicing. However, we found significant effects of stress in the opposite direction for both languages; in heritage speakers' speech, stress shortened the prevoicing of English medial voiced stops and of Tagalog initial voiced stops. This stress effect suggests that our bilingual speakers treat Tagalog voiced stops in initial position as {vl.unasp.}, or the same as English voiced stops.

In other words, we found evidence of an equivalence classification of English and Tagalog voiced stops in heritage speakers' speech, but retention of separate phonetic categories in their voiceless stops, all of which provide partial support for the phonetic category based account of voicing asymmetry in cross-language influence. However, the results for stress effects suggest that the voiced stops of both languages pattern as {vl.unasp.}, rather than as {voiced}, and this equivalence classification is contrary to the prediction that the voiced stops of English will be equated with the voiced stops of an L1 true voicing language due to the pre-emption of existing short-lag voiceless stops in the L1. One possible explanation for this discrepancy is that all our Tagalog heritage speakers are dominant in English, and even though Tagalog is their first language, the English category structure takes precedence in the reorganization of how the four stop types of the two languages are structured in the space of three phonetic categories. Therefore, the two voiced stop categories are equated as the English category (i.e., {vl.unasp.}) rather than the Tagalog category (i.e., {voiced}). Note, however, that despite the dominance of English in this population, Tagalog voiceless stops remain stable and are not equated with English short-lag voiceless stops ({vl.asp.}); that is, the existence of short-lag stops in English (the purported dominant language) does not force Tagalog voiceless stops to be equated with English voiceless stops. This asymmetry may be due to the distributional overlap of these stop categories; English voiced stops straddle both prevoiced and short-lag stop categories and overlap with Tagalog voiced stops, while voiceless stops are well separated in their distribution.[24] Another possibility is that the equivalence classification asymmetry between voiced and voiceless stops is an epiphenomenon of a more general asymmetry in category malleability between voiced and voiceless stops. This asymmetry remains to be explored.

To conclude, in this study we examined the production of English and Tagalog stops by heritage Tagalog speakers in Toronto. The key empirical findings suggest that these heritage speakers produced voiceless stops in their two languages nearly native-like in terms of overall VOT levels and stress effects, but their voiced stops exhibited more cross-language influence. Heritage speakers produced their English voiced stops with more prevoicing (i.e., more Tagalog-like) and their Tagalog voiced stops with less prevoicing (i.e., more English-like). Stress effects provided crucial evidence of voiced stops being treated as a merged category of short-lag stops, while voiceless stops retain independent category labels. The findings of this study provide partial support for the view that the cross-language influence is mediated by equivalence classification of similar categories over the view that influence is driven by phonological markedness.

ACKNOWLEDGEMENTS

We thank Jestine Abella and Yaruna Cooblal for their invaluable contribution to the study. We also thank Jessamyn Schertz, the audience at the 2015 Canadian Linguistic Association, Rajiv Rao and two anonymous reviewers for helpful comments and suggestions. This study was supported by the University of Toronto Excellence Awards to Jestine Abella and Rachel Soo, University of Toronto Scarborough Student Conference Travel Grants to Yaruna Cooblal, Sneha George, and Rachel Soo, and the Social Sciences and Humanities Research Council of Canada grant #435-2013-2092 to Yoonjung Kang.

REFERENCES

Allen, J., & Miller, J. (1999). Effects of syllable-initial voicing and speaking rate on the temporal characteristics of monosyllabic words. *The Journal of the Acoustical Society of America, 106*, 2031-2039.

Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2010). Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2. *Journal of Phonetics, 38*(4), 640-653.

Au, T. K. F., Oh, J. S., Knightly, L. M., Jun, S. A., & Romo, L. F. (2008). Salvaging a childhood language. *Journal of Memory and Language, 58*(4), 998-1011.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., & Grothendieck, G. (2015). lme4 package, version 1.1-10 [Computer software]. Available from https://cran.r-project.org/web/packages/lme4/

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.

Beckman, J., Helgason, P., McMurray, B., & Ringen, C. (2011). Rate effects on Swedish VOT: Evidence for phonological overspecification. *Journal of Phonetics, 39(1)*, 39-49.

Beckman, J., Jessen, M., & Ringen, C. (2013). Empirical evidence for laryngeal features: aspirating vs. true voice languages. *Journal of Linguistics, 29(2)*, 259-284.

Boersma, P., & Weenink, D. (2015). Praat: Doing phonetics by computer, version 5.4.19 [Computer program]. Available from http://www/praat/org/

Broselow, E., & Kang, Y. (2013). Phonology and speech. In J. Herschensohn & M. Young-Scholten (Eds.), *The Cambridge handbook of second language acquisition* (pp. 529-554). Cambridge, MA: Cambridge University Press.

Chang, C. B., Yao, Y., Haynes, E. F., & Rhodes, R. (2011). Production of phonetic and phonological contrast by heritage speakers of Mandarin. *Journal of the Acoustical Society of America, 129*(6), 3964-3980.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics, 27*, 207-229.

Cho, T., & McQueen, J. M. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics, 33*, 121-157.

Davidson, L. (2016). Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics, 54*, 35-50.

De Rosario-Martinez, H., Fox, J., & R Development Core Team. (2015). phia package, version 0.2-1 [Computer software]. Available from https://cran.r-project.org/web/packages/phia/

Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 232-272). Timonium, MD: York.

Flege, J. E., & Brown, W. S. J. (1982). The voicing contrast between English /p/ and /b/ as a function of stress and position-in-utterance. *Journal of Phonetics, 10*, 335-345.

Flege, J. E., & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics, 15*, 67-83.

Flege, J. E., & Eefting, W. (1988). Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *Journal of the Acoustical Society of America, 83*(2), 729-740.

Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A., & Hallé, P. (2008). Cross language phonetic influences on the speech of French-English bilinguals. *Journal of Phonetics, 36*(4), 649-663.

French, K. M. (1991). Secondary stress in Tagalog. *Oceanic Linguistics, 30*(2), 157-178.

Gonzalez, A. (1970). Acoustic correlates of accent, rhythm, and intonation in Tagalog. *Phonetica, 22*, 11-44.

Gonzalez, A. (1998). The language planning situation in the Philippines. *Journal of Multilingual and Multicultural Development, 19*(5), 486-525.

Honeybone, P. (2005). Diachronic evidence in segmental phonology: The case of laryngeal specifications. In M. van Oostendorp & J. van de Weijer (Eds.), *The internal organization of phonological segments* (pp. 319-354). Berlin, Germany: Mouton de Gruyter.

Iverson, G. K., & Salmons, J. C. (1995). Aspiration and laryngeal representation in German. *Phonology, 12*, 369-396.

Kang, K. H., & Guion, S. G. (2006). Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean-English bilinguals. *The Journal of the Acoustical Society of America, 119*(3), 1672-1683.

Kang, Y., & Nagy, N. (2016). The VOT merger in heritage Korean in Toronto. *Language Variation and Change, 28*(2), 249-272.

Keating, P. (1984). Phonetic and phonological representation of stop consonant voicing. *Language, 60*(2), 286-319.

Kessinger, R., & Blumstein, S. (1997). Effects of speaking rate on voice onset time in Thai, French and English. *Journal of Phonetics, 25*, 143-168.

Knightly, L. M., Jun, S. A., Oh, J. S., & Au, T. K. F. (2003). Production benefits of childhood overhearing. *Journal of the Acoustical Society of America, 114*(1), 465-474.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). lmerTest package, version 2.0-29 [Computer Software]. Available from http://cran.r-project.org/web/packages/lmerTest/

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word, 20*, 384-422.

Lisker, L., & Abramson, A. (1967). Some effects of context on Voice Onset Time in English stops. *Language and Speech, 10*, 1-28.

MacKay, I. R. A., Flege, J. E., Piske, T., & Schirru, C. (2001). Category restructuring during second-language speech acquisition. *The Journal of the Acoustical Society of America, 110*(1), 516-528.

MacLeod, A. A. N., & Stoel-Gammon, C. (2009). The use of voice onset time by early bilinguals to distinguish homorganic stops in Canadian English and Canadian French. *Applied Psycholinguistics, 30*, 53-77.

McCarthy, K. M., Evans, B. G., & Mahon, M. (2013). Acquiring a second language in an immigrant community: The production of Sylheti and English stops and vowels by London-Bengali speakers. *Journal of Phonetics, 41*(5), 344-358.

Nagy, N. (2015). A sociolinguistic view of null subjects and VOT in Toronto heritage languages. *Lingua, 164* (Part B), 309-327.

Newlin-Łukowicz, L. (2014). From interference to transfer in language contact: Variation in voice onset time. *Language Variation and Change, 26*(3), 359-385.

Oh, J. S., Jun, S. A, Knightly, L. M., & Au, T. K. F. (2003). Holding on to childhood language memory. *Cognition, 86*(3), B53-B64.

Polinsky, M., & Kagan, O. (2007). Heritage languages: In the 'wild' and in the classroom. *Language and Linguistics Compass, 1*(5), 368-395.

R Development Core Team. (2015). R: A language and environment for statistical computing, version 3.2.1 [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Available from https://http://www.R-project.org

Ramos, T. V. (1971). *Tagalog structures*. Honolulu, HI: University of Hawaii Press.

Rubino, C. R. G., & Llenado, M. G. T. (Eds.). (2002). *Tagalog-English/English-Tagalog (Philipino) standard dictionary.* Conshohocken, PA: Hippocrene Books.

Schachter, P., & Otanes, F. T. (1972). *Tagalog reference grammar*. Los Angeles, CA: University of California Press.

Schertz, J. (2012). Featural enhancement in clarification of misheard stops in Spanish. *Journal of the Acoustical Society of America, 132*(3), 1937.

Simon, E. (2009). Acquiring a new second language contrast: an analysis of the English laryngeal system of native speakers of Dutch. *Second Language Research, 25*(3), 377-408.

Simonet, M., Casillas, J. V., & Díaz, Y. (2014). The effects of stress/accent on VOT depend on language (English, Spanish) consonant (/d/,/t/) and linguistic experience (monolinguals, bilinguals). *Speech Prosody 7*, 202-206.

Smiljanic, R., & Bradlow, A. (2008). Stability of temporal contrasts across speaking styles in English and Croatian. *Journal of Phonetics, 36*, 91-113.

Statistics Canada. (2012). NHS Profile, 2011. Non-official languages spoken for Canada and the City of Toronto. National Household Survey. Ottawa, Canada: Author. Available from https://www12.statcan.gc.ca/nhs-enm/2011/dp-pd/prof/index.cfm? Lang=E

Sundara, M., Polka, L., & Baum, S. (2006). Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition, 9*(1), 97-114.

APPENDIX A
TAGALOG AND ENGLISH WORD LISTS

Tagalog

	Voiced		**Voiceless**	
	Initial Stress	Final Stress	Initial Stress	Final Stress
Initial	***b**ása* ('to read')	***b**asáʔ* ('wet')	***p**ása* ('to pass')	***p**asáʔ* ('bruise')
	***d**áyaʔ* ('fraud')	***d**alá* ('to bring')	***t**ása* ('cup')	***t**asá* ('to sharpen')
	***g**álang* ('respect')	***g**aláw* ('to move')	***k**áma* ('bed')	***k**amáy* ('hand')
Medial	*há**b**aʔ* ('long')	*la**b**á* ('laundry')	*má**p**a* ('map')	*da**p**áʔ* ('to trip')
	*bí**d**a* ('movie star')	*i**d**ád* ('age')	*lá**t**a* ('can')	*ma**t**á* ('eye')
	*lá**g**aʔ* ('to boil')	*la**g**áy* ('to place')	*ná**k**aw* ('to steal')	*ta**k**á* ('to wonder')

English

	Voiced		**Voiceless**	
	Initial Stress	Final Stress	Initial Stress	Final Stress
Initial	***b**ottom*	***b**uffoon*	***p**arty*	***p**ertain*
	***d**ollar*	***d**ecay*	***t**opic*	***t**aboo*
	***g**oblin*	***g**arage*	***c**arpet*	***c**aress*
Medial	*ha**b**it*	*a**b**out*	*o**p**us*	*u**p**on*
	*la**dd**er*	*a**d**opt*	*bo**tt**om*	*a**t**op*
	*da**gg**er*	*be**g**one*	*ra**ck**et*	*be**c**ause*

APPENDIX B
LANGUAGE BACKGROUND QUESTIONNAIRE

Age: ________ Gender: Male Female

1) Do you have any history of known hearing/speech problems or difficulties? ______
2) Have you studied Tagalog as a foreign language at a post-secondary institute?
   Yes No
   If yes:
   - How long? _____ year(s) _____ month(s)
   - Institution ___________________________
   - Course(s) ___________________________

5) Place of birth: Canada Other ____________________
   - If other, at what age did you come to Canada? __________ years old

6) Father's ethnic background: _________________________
7) Mother's ethnic background: ________________________
8) Circle all the family members who you have lived with for over 3 years that are native speakers of Tagalog.

Grandfather	Grandmother	Mother	Father
Brother(s)	Sister(s)	Spouse	Other(s): _________

9) What was your first language(s)? Check all that apply. (e.g., if your parents spoke to you in Tagalog before you were 5, Tagalog is your first language.)
   Tagalog English Other(s): ____________________
10) Have you ever lived/visited the Philippines? Y N
   If yes, (list all occurrences if more than one):
   - At what age: _____ Duration: ____ year(s) ____month(s)

11) Have you ever studied in the Philippines? Y N
   If yes:
   - School level (e.g., K-6, junior high, senior high, Post-secondary, summer, abroad etc…):______________
   - Duration: _____year(s) _____month(s)

For the following questions, please circle:
1 – never 2 – rarely 3 – sometimes 4 – often 5 – always

11) How often do you speak Tagalog with the following people?

Friends	N/A	1	2	3	4	5
Spouse	N/A	1	2	3	4	5
Filipino Classmates/Teacher	N/A	1	2	3	4	5
Grandparent(s)	N/A	1	2	3	4	5
Mother	N/A	1	2	3	4	5
Father	N/A	1	2	3	4	5
Sibling(s)	N/A	1	2	3	4	5

○ Relatives	N/A	1	2	3	4	5
○ Other(s): _______________	N/A	1	2	3	4	5

12) How often do the following people speak Tagalog to you?

○ Friends	N/A	1	2	3	4	5
○ Spouse	N/A	1	2	3	4	5
○ Filipino Classmates/Teacher	N/A	1	2	3	4	5
○ Grandparent(s)	N/A	1	2	3	4	5
○ Mother	N/A	1	2	3	4	5
○ Father	N/A	1	2	3	4	5
○ Sibling(s)	N/A	1	2	3	4	5
○ Relatives	N/A	1	2	3	4	5

Other(s): _______________ N/A 1 2 3 4 5 For the following questions, please circle:

1 – Strongly disagree 2 – Disagree 3 – Maybe 4 – Agree 5 – Strongly agree

13) I am confident communicating in:

○ Tagalog	1	2	3	4	5
○ English	1	2	3	4	5

14) I am confident in the following Tagalog language skills:

○ Speaking	1	2	3	4	5
○ Listening	1	2	3	4	5
○ Reading	1	2	3	4	5
○ Writing	1	2	3	4	5
○ Grammar	1	2	3	4	5
○ Polite style (Official Use – ex. News, official documents, etc.)	1	2	3	4	5
○ Conversation style	1	2	3	4	5

NOTES

1. See Nagy (2015) for an overview of different definitions of the term.

2. Keating (1984) further observes that aspirating languages tend to show positional variation in phonetic category mapping and in English, voiced stops are realized as {voiced} and voiceless stops are realized as {vl.unasp} in word-medial post-tonic position. In this paper, we will take a slightly more abstract definition of these phonetic categories and assume that English voiced and voiceless stops have a phonetic target of {vl.unasp.} and {vl.asp.}, respectively, in all positions, but the length of voicing lag or prevoicing is modulated by word position and stress. This more abstract definition of phonetic targets provides a better explanation of the stress conditioned variation in word-medial position, as will be discussed below.

3. Some of these studies do not refer to their speakers as heritage speakers, but we include them in our review, as the profile of their speakers fits that of heritage speakers.

4. Beyond the literature on heritage languages, a similar asymmetry is reported for Dutch speaking L2 learners of English (Simon, 2009) and early bilinguals who do not fit the description of heritage speakers (Flege & Eefting, 1987; MacLeod & Stoel-Gammon, 2009; Sundara, Polka, & Baum, 2006).

5. Note that equivalence classification does not prevent phonetic learning and that speakers can develop language-dependent phonetic realization rules for the voiced stops of two languages (Flege, 1995; MacKay, Flege, Piske, & Schirru, 2001); the equivalence classification, nevertheless, should hinder the attainment of native-like production and perception.

6. In Keating's (1984) model, English word-medial voiceless stops vary between two phonetic categories conditioned by stress, {vl.asp.} and {vl.unasp}. In other words, stress controls the VOT of voiceless stops to the extent of changing their phonetic categories from {vl.asp} to {vl.unasp} in post-tonic position.

7. Under this view, the ambiguous stress effect on English word-medial voiced stops supports the view that even in word-medial position, where voiced stops are often fully voiced, the phonetic target remains {vl.unasp.} rather than {voiced}. If the phonetic target for medial voiced stops is {voiced}, as assumed in Keating (1984)'s model, we would incorrectly predict that English voiced stops show stress-condition enhancement toward longer prevoicing in medial position.

8. It is also noteworthy that while English voiced stops are generally prevoiced in intervocalic position, stress effects indicate that the phonetic target is not {voiced} but {vl.unasp.}.

9. See Gonzalez (1998) for a comprehensive overview of the complex sociolinguistic situation of the Philippines and a background on the relationship between Tagalog and Filipino. We thank an anonymous reviewer for referring us to this article.

10. Tagalog also has a glottal stop, which surfaces contrastively only in word-final position. The glottal stop occurs predictably at the onset of a vowel-initial word.

11. Two additional female heritage speakers participated in the study, but they had a markedly low proficiency in Tagalog, as judged by the experimenter, who is a heritage Tagalog speaker. Their data are excluded from our analysis.

12. One exception was a speaker whose father was a Korean-Filipino.

13. One speaker was born in the Philippines and came to Canada at the age of one, while another speaker was born in Saudi Arabia and came to Canada at the age of six.

14. One exception was a speaker whose father spoke Spanish.

15. The Tagalog control group's English production is not analyzed in this paper.

16. Here, we only included those cases of closure voicing where voicing duration was longer than the average duration of a single glottal pulse for the speaker.

17. One female native speaker produced all tokens of intervocalic /k/ as [x], while one male native speaker produced post-tonic intervocalic /k/ as [x]. One female heritage speaker produced /k/ in one token of English *because* as [x] but did not spirintize any of her Tagalog /k/ tokens.

18. MacKay, Flege, Piske and Schirru (2001) also report that this type of partial prevoicing (*ceased* prevoicing in their term) is found for heritage Italian speakers and monolingual English speakers but never for native Italian speakers.

19. For initial voiced stops, the case of interrupted voicing (prevoiced-partial) is relatively few in number and the adjusted VOT and traditional VOT measures give comparable results in statistical tests. However, the difference is more substantial for word medial voiced stops, where practically all stops have voicing starting from the onset of stop closure. Hence, the traditional VOT measure reflects the closure duration rather than the voicing duration. For consistency, we used the adjusted VOT measure instead of the regular VOT measure for initial stops as well. However, we cannot measure adjusted closure voicing ratio for initial voiced stops because we cannot identify the onset of closure.

20. This speaker produced all trochaic words with medial voiced stop targets with a creaky second vowel.

21. Note that a number of native Tagalog /k/ tokens were eliminated from the VOT analysis due to spirintization and this may have resulted in an overall lower VOT level for the control Tagalog group compared to the heritage Tagalog.

22. The effect of stress on voiceless stop duration is the opposite of the effect found for voiced stops with the exception of English /t/; post-tonic voiceless stops are longer than pre-tonic voiceless stops in both languages in our data.

23. We include voicing lag in this calculation, as exploratory analyses showed that English and Tagalog differ in the voicing lag found in mixed stops, with English showing longer voicing lag than Tagalog.

24. As an anonymous reviewer points out, English voiceless stops vary between long-lag and short-lag stops and as a result, Tagalog and English voiceless stops also show significant overlap. However, the nature of the variation and the resulting overlap in voiceless stops is different from the one found for voiced stops; the long-lag versus short-lag alternation in English voiceless stops is conditioned by stress, while some aspects of short-lag versus voicing lead alternation in voiced stops is less predictable from phonological contexts.