International Journal of Speech-Language Pathology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/iasl20

Phonemic and pitch variability in bilingual preschoolers: A comparison of Jamaican Creole and English

Amna Samir Mira, Mary Elizabeth Wilkens, Karla Nadine Washington & Victoria Sue McKenna

To cite this article: Amna Samir Mira, Mary Elizabeth Wilkens, Karla Nadine Washington & Victoria Sue McKenna (20 Feb 2024): Phonemic and pitch variability in bilingual preschoolers: A comparison of Jamaican Creole and English, International Journal of Speech-Language Pathology, DOI: 10.1080/17549507.2023.2287426

To link to this article: https://doi.org/10.1080/17549507.2023.2287426

International Journal of Speech-Language Pathology, 2024; Early Online: 1–16

Phonemic and pitch variability in bilingual preschoolers: A comparison of Jamaican Creole and English

AMNA SAMIR MIRA[1,2], MARY ELIZABETH WILKENS[1],
KARLA NADINE WASHINGTON[1,3,4] & VICTORIA SUE MCKENNA[1,5,6]

[1]Department of Communication Sciences and Disorders, University of Cincinnati, Cincinnati, OH, USA, [2]College of Applied Medical Sciences, King Saud Bin Abdulaziz University for Health Sciences, Jeddah, Saudi Arabia, [3]Department of Speech-Language Pathology, University of Toronto, Toronto, Canada, [4]Department of Communicative Sciences and Disorders, New York University, New York, NY, USA, [5]Department of Otolaryngology-Head & Neck Surgery, University of Cincinnati, Cincinnati, OH, USA, [6]Department of Biomedical Engineering, University of Cincinnati, Cincinnati, OH, USA

Abstract

Purpose: The purpose of this study was to investigate the cross-linguistic influences of Jamaican Creole (JC) and English on phonemic and vocal development in bilingual JC-English-speaking preschoolers.

Method: Sixteen typically developing children (12 females, $M = 4$ years; 4 months) completed the Diagnostic Evaluation of Articulation in Phonology Word Inconsistency Assessment subtest in JC and in English. Acoustic measures of voice onset time (VOT), VOT variability (VOT *SD*), mean fundamental frequency (f_o), and f_o variability (f_o *SD*) were extracted from each target word. Prevoicing was noted. Mixed models and regression models were analysed to understand the patterns of acoustic measures in each language, and the relationship between phonemic and vocal variability, respectively.

Result: Analyses showed a significant effect of language on f_o *SD*, wherein *SD* was greater in English than JC. JC spoken (percentage) was a significant positive predictor of VOT *SD* for voiced (short lag) productions. There was no relationship between phonemic and vocal variability measures.

Conclusion: Greater f_o *SD* in English may be due to linguistic f_o differences and speaking environment. Variability for voiced VOT is likely due to the continued maturation of vocal and articulatory control when children are developing adult-like productions, though longitudinal studies are needed.

Keywords: *acoustics; bilingualism; vocal variability; phonemic variability; VOT; speech development*

Introduction

The advanced accessibility to mobility and immigration has led to an increased multilingual population in the USA. Based on the USA Census Bureau (2020), more than 21% of people older than 5 years of age speak languages other than English at home, making bilingualism common rather than the exception with many children speaking more than one language daily. For example, 30% of children served by the Head Start program are bilingual (Office of Head Start, 2016) and 23% of school-aged children are bilingual (Federal Interagency Forum on Child and Family Statistics, 2019). The heterogeneity of bilingual populations along with the significant variability in emerging linguistic skills contribute to the complexity of their speech and language development (American Speech-Language-Hearing Association [ASHA], n.d.). These factors have given rise to an increased interest in characterising the speech production profile of bilingual children, particularly production variability, to enable culturally responsive practices amongst speech-language pathologists (SLPs; Abu El Adas et al., 2021).

Speech development in bilingual children is a complex phenomenon that requires additional research for a more accurate diagnosis of speech sound disorders (SSD). This need was highlighted by Henrich et al. (2010) by way of their acronym “WEIRD”, which represents Western, Educated, Industrialised, Rich, and Democracies. The acronym reflects that

---

Correspondence: Amna Samir Mira, Department of Communication Sciences and Disorders, University of Cincinnati, 3225 Eden Avenue, Cincinnati, OH 45267, USA. E-mail: miraas@mail.uc.edu

Presentations: Portions of this work have been presented as a poster presentation to the Boston Speech Motor Control Symposium, June 2023.

ISSN 1754-9507 print/ISSN 1754-9515 online
Published by Taylor & Francis
DOI: 10.1080/17549507.2023.2287426

most research, including that about speech development, is generalised from a homogenous and monolingual orientation. Because bilingual children are underrepresented in the research literature, less is known about how to distinguish between speech differences and SSD. Consequently, bilingual children are at higher risk of being misdiagnosed when compared to their monolingual peers (Castilla-Earls et al., 2020; Irizarry-Pérez et al., 2023).

The proportion of SLPs who work with bilingual children has increased over the last two decades, with now 8.3% of registered SLPs in the USA being multilingual service providers (ASHA, 2022). As part of their scope of practice, SLPs need to provide culturally sensitive services to ensure that appropriate decisions are made about disorder vs. development for children on their caseload (ASHA, 2017). However, the cultural and linguistic mismatch between SLPs and their bilingual clientele creates the potential for service provision that does not appropriately meet the needs of these speakers. The lack of research-based tools for clinical management of SSD in bilingual populations further contributes to a growing need for evidence to guide decision-making practices (Washington et al., 2023). This concern is pressing in the context of understudied language combinations, such as our pairing of Jamaican Creole (JC) with English. To address the need for research to guide practice, our study investigated acoustic measures of speech and voice in JC-English bilingual preschoolers to identify cross-linguistic effects in children during their early speech development.

JC and English

JC is an English-lexifier language, meaning that JC is an independent language although it has shared linguistic foundations with its input languages of English and West African (Cassidy & Le Page, 1980; Devonish & Harry, 2008). Jamaica is a bilingual nation where most people speak two languages: JC and English, depending on the sociolinguistic context. English is used in formal contexts such as education and JC is the social-informal language used within the community (Devonish & Carpenter, 2007). JC-English is a common pairing used by Jamaicans in their home country and their diasporic communities, such as in the USA and the UK (Mair, 2003). JC-English bilinguals are mostly simultaneous bilinguals who were exposed to both languages since birth (Meade, 2001; Washington et al., 2017, 2019).

JC-English speakers use both languages daily, which contributes to the need to code-switch between the two languages. The simultaneous use of the two languages has created a spectrum of language varieties ranging between JC and English (Devonish & Harry, 2008; Washington et al., 2017). The existence of this JC-to-English continuum has resulted in the inconsistent use of both languages and systematically varying levels of speech production interaction between languages (Devonish & Harry, 2008). This means that speakers use larger or smaller percentages of JC and English based on their communication partner (e.g. in school vs. at home).

JC and English share some phonological features, but they both reserve portions of their own unique phonological inventories. Some noted phonological differences include the smaller phonemic vowel and consonant sets in JC than in English. For example, English has 15 vowels and 24 consonants compared with 12 and 22 for JC, respectively (Devonish & Harry, 2008). Phonological differences in both languages, and Jamaicans' simultaneous use across the linguistic continuum of JC and English, results in a range of typical phonological varieties in their speech productions (see León et al., 2022, and Washington et al., 2017, for a discussion). Previous research reported that JC-English bilingual children produced higher within-word phonemic variability in JC than in English, although they presented a similar amount of variability when compared to their monolingual peers (Abu El Adas et al., 2021). The authors purported that increased variability in JC might be related to language proficiency and amount of exposure; however, this hypothesis was not explored. The use of acoustic methods was recommended for future research to examine the role of proficiency and exposure on production variability.

Culturally adapted speech protocols

To enable culturally responsive practices, researchers examining JC-English bilingual speech productions have adapted well-known assessment tools to ensure that linguistic features of the native and ambient languages are reflected in the phonetic inventory under evaluation (León et al., 2022; Washington et al., 2017). The culturally adapted protocols refer to evidence-based scoring criteria tailored to JC-English bilinguals, which provide clinicians guidance on acceptable forms of speech productions that vary from the "standard" targets (Washington et al., 2017). For example, the Standard American English (SAE) production of the word "pig" is /pɪg/, whereas acceptable versions of the word in the Jamaican context include response options such as /pɪg, pɪgə, pɪgi, pɪgɪ, pɪgz/ or [pi:g]. Another example of different pronunciations in JC is the multiple accepted forms of vowels such as in the word "watch", where both /waʧ/ and /wɑʧ/ are acceptable (Washington et al., 2017). Not only are these productions judged as being "accurate" in a modified protocol that accounts for linguistic variations, but the difference between JC and English productions could be measured objectively using acoustic word or vowel durations.

León et al. (2022) investigated speech characteristics of JC-English bilingual preschoolers by comparing their speech productions with adults of the same linguistic background. Speech productions were objectively compared using acoustic measures of voice onset time (VOT), word duration, and vowel duration. Results revealed that adults and children produced a variety of acoustic and perceptual variations that were different from the "standard" (i.e. SAE) target response, reflecting the necessity to use a culturally-adapted protocol to capture the phonological variations in JC-English bilinguals (León et al., 2022).

Utility of acoustic measures for understanding phonemic and vocal development

Acoustic measures have been shown to reflect developmental changes in the underlying anatomical and physiological systems involved in speech and voice production (Vorperian & Kent, 2007), and have been of interest to bilingualism researchers. Acoustic measures of vowel duration (Coy & Watson, 2020; Jacewicz & Fox, 2015; León et al., 2022), vowel formants (Bosch & Ramon-Casas, 2011; Coy & Watson, 2020), and VOT (León et al., 2022) provide rich information about cross-language differences as well as subtle allophonic variations. To date, however, there have been few explorations into acoustic measures that consider the cross-effects of articulatory and vocal development in languages with shared linguistic foundations. This has resulted in missing, yet important, information about codeveloping physiological systems that must coordinate with one another to produce target phonemes effectively and consistently in both languages. Given that the trajectory of vocal motor development occurs over a longer period of time than articulatory development (i.e. through puberty for full maturational vocal changes in late adolescence), vocal development is critical to understanding the transition from early childhood to adult productions (Hamilton, 2023).

Acoustic measures of vocal development

Speech and vocal development begin at birth when infants start vocalising, though initial vocal productions in infants are universal and not language-specific (Oller, 2000). As children age, their speech productions progressively attune to the ambient language/s (Blake & Boysson-Bardies, 1992). Not only does the complexity of the child's verbal output increase during this time, but the anatomy (vocal folds, vocal tract, and related structures), physiology, and motor control are all simultaneously changing. Consequently, children learn and adjust motor timing and sequences of speech sounds in a changing body system, while also learning the phonemic and phonotactic rules of their language/s (Callan et al., 2000). Although this progression and attunement have been well described and quantified in monolingual speakers, studies on vocal development, specifically, in bilingual speakers are lacking.

One measure commonly examined in vocal development research is mean fundamental frequency (f_o), which is the corresponding measure of a perceptual speech parametre, pitch. Mean f_o is higher in infants (336.9 Hz; Rothgänger, 2003) and children (244.8 Hz; Gelfer & Denor, 2014), lowering with the increase in laryngeal size that comes with age and hormonal changes (e.g. testosterone in males; Evans et al., 2008). The growing size of the larynx corresponds with vocal fold ligament changes and lamina propria differentiation (i.e. the epithelial layer of the vocal folds responsible for vibration), occurring from the age of 1 year through adolescence to be fully developed (Hartnick et al., 2005). Throughout this time, the ability to control mean f_o and f_o range (i.e. pitch range) varies by age, in which children aged 6–7 years do not exhibit the same degree of control over f_o as those aged 8–11 years (Wuyts et al., 2003). Furthermore, the ability to quickly and consistently abduct and adduct the vocal folds during laryngeal diadochoknetics (e.g. /i-i-i/) shows that younger children exhibit greater variability in the timing and consistency of voicing onsets when compared to older children (Modolo et al., 2011).

The variability of f_o, defined as f_o *SD*, has been identified as a marker of vocal motor development and control (Ohde, 1985). Lee et al. (1999) examined f_o and f_o *SD* across more than 400 monolingual children from the ages of 5–18 years, and compared their maturational trajectories to established voice parametres in adults aged 25–50 years. Children reached adult-like ranges for f_o and f_o *SD* around age 15 years, following the completion of puberty. The variability of f_o decreased over the course of development; however, a 50% reduction of within-subject f_o *SD* was observed in the 5–8 years range, indicating a rapid change in earlier childhood and a non-linear trajectory over development. Thus, f_o *SD* can be used as a measure of vocal motor development in children and, in particular, information on its stabilisation in younger children is important to better understand speech development.

In addition to anatomical and physiological developmental factors that influence measures of f_o and f_o *SD*, f_o is also known to be impacted by language and sociocultural factors. That is, there are dialect-specific f_o values and f_o ranges reflecting different regional communities (Jacewicz & Fox, 2015). For example, Coy and Watson (2020) evaluated children aged 8–11 years speaking Jamaican English (JE) as well as other dialects of English, such as Birmingham (British) English and Pittsburgh (American) English. Analysis of mean f_o within corner vowel productions revealed that Jamaican children had significantly higher mean f_o when speaking JE than other English varieties. Moreover, JE speakers had larger vowel spaces and longer vowel durations (likely contributing to the higher mean f_o found), contributing to information about the sociolinguistic uses of the language in comparison to the other dialects. Authors concluded that JE is a distinct variety of English and differences should be taken into consideration during clinical and educational practices (Coy & Watson, 2020).

In bilingual speakers, similar patterns emerge when comparing f_o, f_o range, and f_o *SD*. Several studies have established differences in the way f_o is used across language pairings of adult speakers of German-English (Mennen et al., 2012), English-French (Pépiot & Arnold, 2021), Russian-English (Altenberg & Ferrand, 2006), Korean-English (Lee & Sidtis, 2017), Dutch-English (Theelen, 2017), and Mandarin-English (Lee & Sidtis, 2017), which are thought to be influenced by both phonological and phonetic context of the language as well as cultural differences (Mennen et al., 2012). There are relatively few evaluations of f_o parametres in bilingual children (Ng et al., 2010), but still, preliminary results show language-specific differences as well. That is, Ng et al. (2010) compared pitch parametres in Cantonese to that of English in bilingual children and adolescents aged 5–15 years. The investigators determined that speaking f_o and pitch range were both lower in Cantonese compared to English productions, suspected to be due to the tonal nature of Cantonese. Additionally, speaking f_o and pitch range tended to be lower in children aged over 10 years, suggesting a lowering in pitch with vocal maturation. To gain deeper insights into the effects of development within each language, further analyses are warranted since language-specific demands on f_o are likely to differ across different languages and cultures.

Acoustic measures of phonemic development

VOT is a phonemic-related acoustic measure that reflects both phonemic and vocal motor development. VOT is a characteristic of plosive consonants that contributes to perceptually identifying different phonemes across some languages (Kellogg & Chang, 2023; Lisker, 1978), such as distinguishing voicing in English (e.g. /p/ vs. /b/; Lisker & Abramson, 1967). VOT has been shown to be impacted by place of articulation (POA; McCarthy et al., 2014), word phonotactic structure, age, sex, and language (Swartz, 1992). Voiceless plosives (e.g. /p, t, k/) have longer VOT than voiced plosives and are categorised as a "long voicing lag" with an average duration of 74 ms in JC-English bilingual adult speakers (León et al., 2023).

Short lag VOT for a voiced plosive is reported as a positive value and defined as the duration between the plosive air burst and the voicing onset for the following phoneme. Short lag voiced VOT is typically closer to 20 ms in adult speakers of SAE (McKenna et al., 2020). Although not as common in SAE, speakers may also use a voicing lead (sometimes referred to as "prevoicing"), in which the voicing occurs prior to the plosive burst and is measured from the voicing onset to the beginning of the following phoneme (Adi et al., 2016). In JC, adults tend to produce voiced plosives with prevoicing, which is represented as a negative VOT value and reported to be approximately −40 ms (León et al., 2023).

Vocal control is important for VOT phonemic distinctions, as VOT requires precise articulatory and temporal coordination between laryngeal and oral articulators (Löfqvist, 1980). For example, when producing unaspirated voiceless plosives, the vocal folds must first abduct to allow air to flow through the glottis to the constriction point of the articulators in the oral cavity. Then, the constriction release coordinates with adduction and vibration of the vocal folds for the following voiced phoneme (Löfqvist, 1980). On the other hand, when producing unaspirated voiced plosives, the vocal folds are mostly adducted before the release of the constriction point of the articulators (Dixit, 1989). The timing and coordination of these laryngeal movements, in addition to correct articulatory movements (including velopharyngeal closure and adequate lip and/or tongue constriction), are essential for accurate VOT distinctions.

Monolingual children begin to develop this coordination and refine their productions throughout their childhood (Kewley-Port & Preston, 1974). In a study by Khattab (2004), the researcher compared VOT productions in monolingual English and Arabic children (aged 5–10 years) and adults. VOT duration and variability in English monolingual children decreased significantly with age, indicating a gradual acquisition process to precise and consistent articulatory and vocal targets. The Arabic monolinguals exhibited prevoicing behaviour, although they did not resemble the adult form, indicating they were still developing that fine-tuned distinction of the adult prevoicing form.

Prevoicing requires more advanced motor control and timing than short voice lag, as it has been consistently difficult for children to achieve across several languages, such as Italian (Bortolini et al., 1995) and Arabic (Khattab, 2002, 2004). In the case of bilingual JC-English preschoolers, children exhibited inconsistent prevoicing patterns for voiced plosives in the initial position, leading to an inability to meet language demands in both languages (León et al., 2023). That is, JC-English bilingual preschoolers had VOT values of 14 and 17 ms for JC and English respectively, whereas adult productions were significantly lower at −40 and −37 ms (León et al., 2023). As such, it appears that VOT development in bilingual speakers is influenced by a range of factors, including age, maturation, and language-specific features.

Researchers have also identified cross-linguistic effects in bilingual children's VOT productions, in which one language may influence productions in another. A study on Sylheti-English (British) bilingual preschoolers revealed significantly shorter voiced VOT productions in the English language (that were comparable to Sylheti voiced plosives) than monolingual English-speaking children prior to children starting school. This longitudinal study then showed that bilingual children's VOT measures were later equivalent to their monolingual peers once they spent time at school, suggesting refinement of VOT with more language exposure and use (McCarthy et al., 2014). Language exposure can have a significant impact on the development of phonemic distinctions in bilingual children, influencing the subsequent development and acoustic measures in children exposed to varying degrees and timing of different languages (Bosch & Ramon-Casas, 2011; Khattab, 2004; Zen, 2019). Questions remain regarding the role of language input in bilingual phonemic distinctions during language development and how best to characterise bilingual vocal and phonemic development over childhood.

The relationship between vocal and phonemic variability

A neural theory of speech acquisition, the Direction into Velocity of Articulators (DIVA) model, reported a negative relationship between variability and phonetic accuracy, in that variability of speech production decreases as acquisition competency increases for monolingual speakers (Tourville & Guenther, 2011). As such, it is purported that this principle holds true for vocal variability and that, moreover, phonemic and vocal variability may be related to one another and impacted by age.

To investigate the relationships between VOT variability and vocal variability over a critical time period of acquisition, Heller Murray and Chao (2021) examined VOT and f_o from monolingual children, aged 2, 4–6, and 8 years. The researchers expected to find a positive relationship between measures of phonemic and vocal variability due to their overlapping motor control components. Contrary to their expectations, the researchers reported a significant negative relationship between the two variables, with larger VOT variability being correlated with smaller amounts of vocal variability (as measured by coefficient of variation of f_o within corner vowels) when unadjusted for age. The researchers proposed that children with larger articulatory variation may need more constrained vocal motor targets to compensate for their articulatory variability. They further suggested that future studies should examine the impact of vocal and articulatory variability on vocal-articulatory coordination using a developmental approach to disambiguate the interplay between these two systems (Heller Murray & Chao, 2021).

A similar knowledge gap persists for bilingual children, where understanding the codevelopmental trajectories for the vocal and phonemic systems remains unknown. While other studies investigated the vocal system (Heller Murray & Chao, 2021) and differences of vocal and phonemic systems produced among languages spoken by bilingual children during their development (Khattab, 2002, 2004; León et al., 2022, 2023; Ng et al., 2010), there are still no studies to our knowledge that examined the interplay between vocal and phonemic systems in bilingual children.

Research aims and hypotheses

Despite previous research on the differences between JC and English, less is known about the codevelopment of these two languages and even less about the interaction between the developing vocal system and phonemic inventories for each language (Abu El Adas et al., 2021). Understanding how two phonemic language inventories interact while children are acquiring both languages offers crucial information that provides clinical significance towards appropriately diagnosing bilingual children with suspected SSD. The following aims and corresponding hypotheses were addressed:

> Aim 1: Examine language-specific patterns of phonemic and vocal acoustic properties in JC-English bilingual children.
>
> Hypothesis: We hypothesised an impact of language on phonemic and vocal acoustic measures and their variability, indicating language-specific differences as seen in previous literature (Pépiot & Arnold, 2021; Theelen, 2017).

First, we expected mean f_o values to be significantly impacted by language and mediated by contextual factors of language exposure and use. This was proposed due to the results of a previous study on Dutch-English bilingual speakers, in which adults had higher mean f_o when speaking Dutch than when speaking English (Theelen, 2017). The author suggested that this result was consistent with previous research that found different languages have different mean f_o due to language-specific characteristics. Specifically, we expected mean f_o to be lower in English given that Jamaican children had reportedly higher f_o in JE when compared to other dialects of English (Coy & Watson, 2020). Second, we expected to find greater vocal variability (f_o *SD*) in English compared to JC, also influenced by the factors of language use and exposure. This result would be consistent with previous work by Pépiot and Arnold (2021), where English-French bilingual speakers presented greater vocal variability in English than in French due to linguistic differences. This could be possibly due to allophonic variations in the vowel space and durations that influence f_o, and are likely due to sociolinguistic differences in the way English is used.

Third, we expected language-specific phonemic productions that rely on vocal motor control (VOT, VOT *SD*, prevoicing) to vary between languages for plosives. This was hypothesised because other studies of bilingual preschoolers presented different VOT values in their two spoken languages, with prevoicing being more common in their dominant language, Arabic (Khattab, 2004). If more negative voiced VOT values and greater voiced VOT *SD* are found in JC compared to English, this would indicate that children may have more difficulty in consistently producing VOT that resembles adult-like presentations (León et al., 2023). This would point to possible developmental effects on coordinating articulatory and vocal systems. Fourth, because prevoicing has been shown to be a later developing voicing target (Khattab, 2004; León et al., 2023), we expected that there would be a difference in the proportional amount that children were able to produce prevoicing across languages, with larger amounts evidenced in JC compared to English.

> Aim 2: Explore the relationship between phonemic and vocal variability in bilingual JC-English speaking children.
>
> Hypothesis: We hypothesised that a positive relationship exists between phonemic and vocal variability because mean f_o and VOT are both dependent upon vocal motor control.

A positive relationship would indicate that voice-related phonemes are influenced by vocal maturation and that children who exhibit increased f_o variability would also exhibit increased VOT variability. This hypothesis was based on previous theoretical work proposing a relationship between variation and speech mastery, in which a reduction in variation occurred as phonemic accuracy increased (Tourville & Guenther, 2011). Moreover, previous work has shown that abnormal changes in the vocal system resulted in greater VOT variability in monolingual adults (McKenna et al., 2020), supporting that variability is evidence of an unstable system.

Method

Participants

The Institutional Review Board of the University of Cincinnati and the Medical Ethics Board of the Faculty of Medicine, University of the West Indies Mona Campus, Kingston, Jamaica approved the research project (Jamaican Creole Language Project). Support and permissions for conducting research were also provided by school board officials for participating schools and by the Early Childhood Commission, Jamaica. Data were collected as part of a larger research project, the Jamaican Creole Language Project, in Kingston, Jamaica. All children's parents consented to their children's participation and all children verbally assented. Originally, datasets from a total of 23 children were available for this study; however, due to background noise at the time of data collection, reliable acoustics for the purposes of this research could only be extracted from 16 JC-English bilingual preschoolers. The final subset of participants included 12 females (75%) and 4 males (25%), aged 3;6 to 4;11 with an average of 4;4 (i.e. 4 years, 4 months).

Table I. Participant descriptions and performance (ages shown in years; months).

Variable	Characteristics
Age range (years; months)	3;6–4;11
Sex	
Male	4
Female	12
PTONI	
SS range	91–>149
Qualitative description	Average–very superior
M (*SD*)	120.5 (13.7)
Oral motor skills	WNL
ICS	
English *M* (*SD*)	4.5 (0.62)
JC *M* (*SD*)	4.5 (0.61)
Language exposure	
English range (average)	20–80% (55%)
JC range (average)	20–80% (45%)
Language use	
English range (average)	40–80% (66%)
JC range (average)	20–60% (34%)

Note. PTONI = Primary Test of Nonverbal Intelligence. SS = standard scores, which are reported for nonverbal IQ as measured by the PTONI. WNL = Within normal limits. ICS = Intelligibility in Context Scale (McLeod et al., 2012). JC = Jamaican Creole.

All participants passed a binaural hearing screening at 25 dB HL for 1 kHz, 2 kHz, and 4 kHz. Additionally, their parents reported that they had typical development with no history of neurological or developmental disorders. Participants passed the Primary Test of Nonverbal Intelligence (PTONI), a standardised test that provides quantitative nonverbal measures of reasoning abilities in young children across a variety of linguistic backgrounds (Ehrler & McGhee, 2008). Children included in our study achieved an "average" score or higher on the PTONI, i.e. an average standard score of 120 (above average) and a range of 91 (average) to >149 (very superior).

Parents completed the Intelligibility in Context Scale (ICS), which is a 7-item, 5-point parent questionnaire that measures a child's functional intelligibility across different interlocutors (McLeod et al., 2012; Washington et al., 2017). The ICS was completed for both JC and English in sessions counterbalanced by language. Results were similar in both languages for all children, indicating that all participants were functionally intelligible in both languages. Results ranged from 3.42 (sometimes to usually intelligible) to 5 (always intelligible) with an average score of 4.5 (usually to always intelligible). Parents also provided information regarding the percentage of language (JC and English) exposure and language use in their daily interactions at home. Please see Table I for all demographic and scoring information for each child participant.

DEAP

The Diagnostic Evaluation of Articulation and Phonology (DEAP; Dodd et al., 2006) is a standardised norm-referenced test that assesses phonological processes for ages between 3;0 and 8;11. For the purposes of this study, the Word Inconsistency Assessment (WIA) subtest of the DEAP was administered to preschoolers using the hard copy print version. This subtest aims to determine production variations of phonemes across three trials. It contains 25 words, which were elicited using a picture-naming task that is repeated three times in full order. The DEAP, including DEAP-WIA, was originally normed on SAE speakers but has been previously used with Jamaican children (Washington et al., 2017). Participants in our study completed the DEAP-WIA on two occasions, once with an SAE SLP assessor and once with a JC-English SLP assessor, in counterbalanced sessions. The DEAP battery includes an Oral Motor Screening, which was also administered. All children achieved the age-based criterion on this subtest.

Protocol

DEAP WIA

The DEAP-WIA was used to elicit speech for acoustical analyses. Children completed the entire test twice, counterbalanced, in which they were prompted in English by an SAE-speaking SLP and prompted in JC by a native JC-speaking SLP. The assessor used hierarchical prompting if the child was unable to produce the target word. A non-targeted or no response resulted in semantic hints, followed by binary forced-choice cuing with the target response said first, and finally, a direct repetition of the target was elicited. Children in our sample required minimal to no prompting, as they spontaneously produced target words during the naming task >95% of the time.

Acoustic data collection

Audio recordings were collected from all participants in relatively quiet rooms in school settings in Kingston, Jamaica. Recordings were completed using ZOOM H6N portable recorders with four microphones attached; MOVO LV4-C XLR unidirectional cardioid lavalier microphone attached to the fitted vest on participants, MOVO LV4-O Omnidirectional directed to the adult assessor, and two LV4-O2 Omnidirectional with noise isolation feature directed to two sources of background noise in the room (e.g. one directed to the air conditioner and the other directed to the window facing the courtyard). There was a consistent 6-feet distance between the adult assessor and the participant to avoid microphone interactions. Samples were digitised with a sampling rate of 22 kHz and 24-bit encoding.

Speech stimuli

Although the DEAP-WIA includes 25 words, our acoustic analysis targeted a list of phonemes (/p, b, t, d, k, g/) in specific speech contexts. Phonemes that were targeted for the VOT extractions were in the initial word position, followed by a vowel or the consonant-vowel form. Eight words were available for VOT extraction, including *parrot*, *boat*, *butterfly*, *teeth*, *tongue*, *dinosaur*, *kangaroo*, and *girl*. Therefore, words contained all possible plosives (e.g. /p, b, t, d, k, g/) in the JC and English languages. Phonemes targeted in this set were categorised by voicing (voiced or voiceless) and POA (bilabial, alveolar, and velar).

The words used to extract measures of mean f_o and its *SD* included two corner vowels, /i/ and /u/. Both phonemes are included in JC and English phonemic inventories (Devonish & Harry, 2008). Five words met this inclusion criterion: *teeth*, *zebra*, *cleaner*, *vacuum*, and *kangaroo*. Mean f_o measures were extracted from steady-state segments within the targeted sustained vowels. Any instances of different vowel productions were excluded from analysis (e.g. /ɛ / for /i/). Of note, the corner vowel /ɑ/ (in the word "watch") was not targeted in our analysis due to the multiple accepted forms of pronunciations in JC adapted scoring form (i.e. [waʧ] vs. [wɑʧ]; Washington et al., 2017). See Table II for all speech stimuli and phonemic transcriptions.

Acoustic extraction

Acoustic measures were extracted using the software Praat (Version 6.2.15; Boersma, 2001). Data were extracted by authors A.M. and M.W., as both were trained on VOT and vowel extraction processes used for this study. This training was facilitated by K.N.W.

Table II. Stimuli and phonemic transcriptions.

Measure	Variable	Stimuli	Phonemic transcription*
Mean f_o, f_o *SD*	Vowel /i/	Cleaner	/kl**i**nər/
		Teeth	/t**i**θ/
		Zebra	/z**i**brə/
	Vowel /u/	Kangaroo	/kæŋgər**u**/
		Vacuum	/vækj**u**m/
VOT, VOT *SD*	Voiced plosives	Butterfly	/**b**ʌtərflaɪ/
		Boat	/**b**oʊt/
		Dinosaur	/**d**aɪnəsɔr/
		Girl	/gɜrl/
	Voiceless plosives	Parrot	/**p**ærət/
		Teeth	/**t**iθ/
		Tongue	/**t**ʌŋ/
		Kangaroo	/**k**æŋgəru/

Note. f_o = fundamental frequency; VOT = voice onset time.
* Target sound bolded.

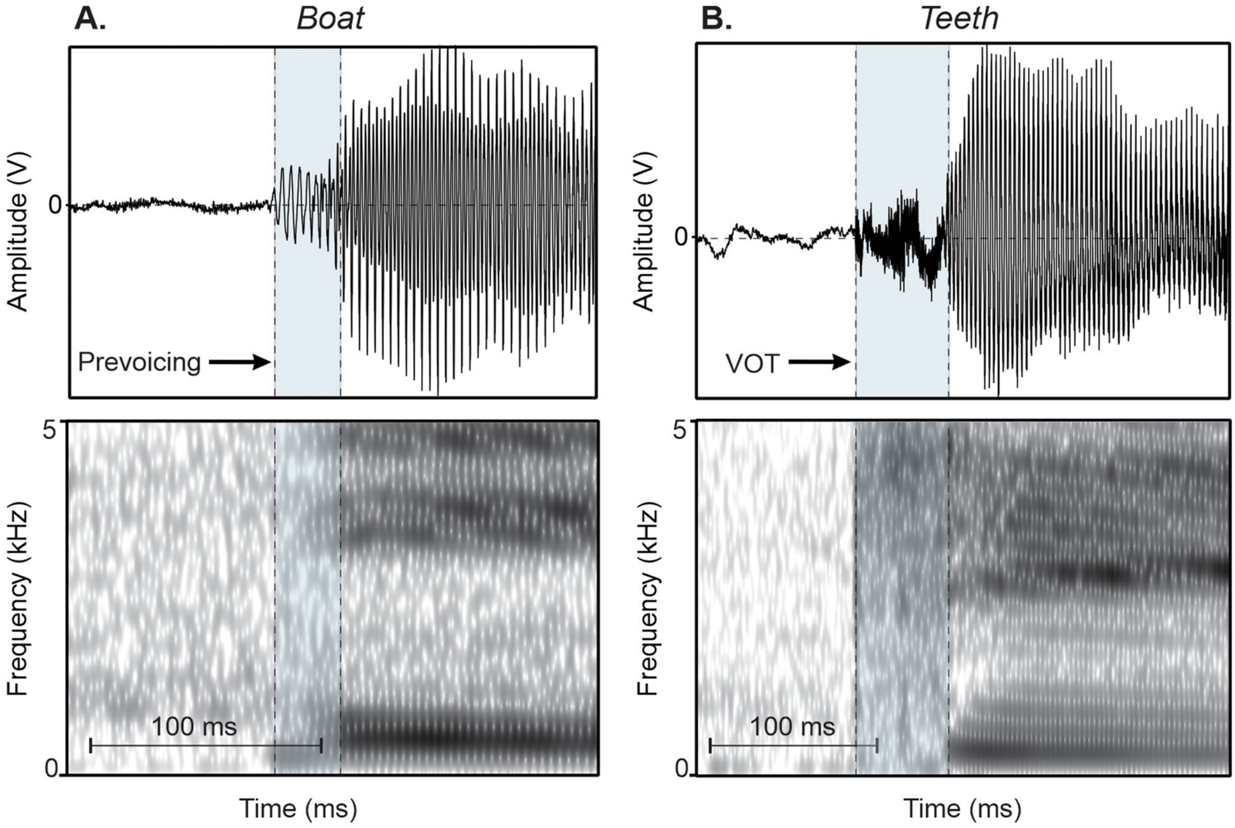

Figure 1. Acoustic measures extracted from the same child. (A) Example of a prevoicing segment that was extracted from the word *boat* with a negative value of -28 ms. (B) Example of a long lag VOT segment that was extracted from the word *teeth* with a positive value of 67 ms.

(a bilingual JC-English SLP researcher) and V.S.M. Prior to the data extraction, inter-rater reliability was established between raters on randomly selected samples from the original 23 children. Inter-rater reliability was determined via an intraclass correlation coefficient (ICC) on measures of VOT and mean f_o for productions in JC and English. Inter-rater reliability was excellent between the two raters at an average ICC = .97. These raters then each extracted 50% of the data (i.e. eight subjects each) for the data in this study.

Mean f_o and f_o SD extraction protocol

The corner vowels /i/ and /u/ were targeted in five designated words, as described above. Once again, the acoustic waveform and spectrogram were used to determine vowel location. The vowel steady-state of each word was analysed to identify vowel mean f_o and within word f_o *SD*. The steady-state was extracted to reduce the possibility of co-articulation and/or voicing transitions into and out of other phonemes. Mean f_o and f_o *SD* were extracted using the voice report function in Praat, and then were averaged across the three productions. There were four missing cases for mean f_o (i.e. 2.5% of all possible mean f_o values) and eight missing cases of f_o *SD* (i.e. 5% of all possible mean f_o *SD*).

VOT extraction protocol

VOT was extracted from the eight words identified and described previously for a total of 768 analysed tokens (16 participants × 8 targets × 3 repetitions × 2 languages). The authors obtained this measure by examining the acoustic waveform and corresponding spectrogram (see Figure 1). Using Praat's pulse function, f_o detection was completed (range 90–500 Hz). VOT was measured from the plosive burst to the voicing onset of the vowel segment (Lisker & Abramson, 1964), or prevoicing was measured from the onset of voicing (prior to the plosive burst) to the beginning of the vowel (Adi et al., 2016). VOT values were considered 0 ms if the release of the burst coincided immediately with the onset of voicing, and VOT was considered a positive value if the release of the burst and the initial voicing wave were delayed more than 0 ms. If the preschooler began voicing prior to the burst release, it was measured in a negative value and considered as a voicing lead, or prevoicing (Stoehr et al., 2018).

Voiceless VOT values were averaged across the three productions for each word and the VOT variation across the three words was calculated via *SD*. If a word was only available for extraction one time, that value was reported as the average VOT for that word, but *SD* could not be calculated because there was no variability to report. There were no missing values for voiceless mean VOT and voiceless VOT *SD* was only missing in one case (i.e. < 1% of all possible *SD* values).

Voiced VOT values were first divided by prevoicing (negative values) and short lag (positive values) productions, based on previous work that recommended separate analyses for each subgroup due to different data distributions (Herd, 2020). Next, prevoiced and short lag VOT were averaged separately for each speaker in each language. Each language had a total of 12 possible tokens for averaging (4 target words × 3 productions each), and a *SD* was also calculated from these as well. There were no cases of missing mean VOT (short lag) but some children did not exhibit prevoicing, resulting in three missing values for mean VOT (prevoiced), representing 9% of possible datapoints. Two were missing in JC and one in English. Similarly, there were no missing values for short lag VOT *SD*, but because some children did not produce prevoicing or only produced one prevoiced token in a language, there were eight (25%) missing prevoiced VOT *SD* values in the analysis, evenly distributed between JC and English languages.

Finally, the frequency of prevoicing was calculated as a percentage of prevoiced tokens divided by the total number of voiced VOT productions for each language. For example, if a participant prevoiced six times out of 12 total productions, their prevoicing rate was reported as 50%. In some cases not all 12 tokens could be extracted, so the proportion was based on the total number of successfully extracted productions for each participant. Missing productions were primarily due to background noise in the recordings that affected the sound quality and, subsequently, the confidence in the extraction.

Reliability

Authors A.M. and M.W. extracted data from four randomly selected participants, in which they were blinded to their previous extractions. Each rater extracted data from two participants from their original eight participants for intrarater reliability measures, as well as two participants from the other extractor's assignment for inter-rater reliability. Intrarater reliability was calculated via a Pearson's correlation coefficient, yielding good-to-excellent reliability (Koo & Li, 2016) for author A.M. with an average $r = .92$ (range = .88–.95) across the four study measures. Author M.W. showed moderate-to-excellent intrarater reliability with an average $r = .85$ (range = .71–.97) across all four measures. Inter-rater reliability was deemed good-to-excellent with ICC values of .93 (VOT), .80 (VOT *SD*), .95 (mean f_o), and .85 (f_o *SD*). Inter-rater reliability analyses were completed in R using package irr.

Statistical analysis

To identify the impact of language on acoustic measures, each acoustic measure was analysed in separate mixed-effects linear models. Voiceless, voiced (short lag), and voiced (prevoiced) plosives were analysed separately for mean VOT and VOT *SD*. The fixed effects in the voiceless plosive model included language (JC, English) and POA (bilabial, alveolar, velar). POA was included in this model because it is a known factor that influences VOT in English and JC (León et al., 2023). Language (JC, English) was a fixed factor for the voiced (short lag) and voiced (prevoiced) models as well; however, POA could not be added as a factor due to the division of voiced plosives into two subgroups and the subsequent need to average over all available productions. In the models for mean f_o and f_o *SD*, the analyses included language (JC, English) and vowel (/i, u/), because vowels have been found to have intrinsically different f_o (Jacewicz & Fox, 2015). Two-way interactions of fixed effects were not investigated in any model as they were not the focus of this study. Covariates of JC use and JC exposure (each measured as a percentage) were added to each model to investigate their influence on the other variables. For all models, "participant" was input as a random effect to account for random variation inherent in each speaker and was calculated as a random intercept. Random effect variance was estimated using a restricted maximum likelihood algorithm and fixed effects were calculated using the Kenward-Roger approximation. Significance was set to $p < .05$ for each variable. Prior to analysis, acoustic variables were assessed for normality using the Anderson-Darling test at the $p < .05$ level. Variables of VOT *SD* for voiceless plosives and f_o *SD* did not meet normality assumptions and required log transformations prior to further analysis.

To investigate whether the proportion of prevoicing for voiced VOT productions varied by language, a mixed-effects model with the same factors of language (JC, English), JC exposure, JC use, and participant (random) was calculated. A significance level of $p < .05$ was applied to this analysis.

Next, simple linear regression models were calculated to understand the relationship between phonemic and vocal variability within each language. To create these models, voiced (short lag, prevoiced) and voiceless VOT *SD* were averaged and analysed separately as predictors of f_o *SD* for each language. That is, six models were created: three for English and three for JC, with voiced (short lag) VOT *SD*, voiced (prevoiced) VOT *SD*, and voiceless VOT *SD* as predictors of f_o *SD* for each language. First, VOT *SD* were averaged separately for all prevoiced, short lag, and voiceless productions for JC and English. Next, f_o *SD* was averaged within each phoneme for each language (i.e. separately for /i/ and /u/) so that results were not skewed towards a phoneme that was produced more frequently than another. The averaged /i/ and /u/ were then averaged together into a final single value of f_o *SD* per participant, per language. Variables were assessed for normality (Anderson-Darling, $p < .05$), finding that variables of voiceless VOT *SD* (JC) and

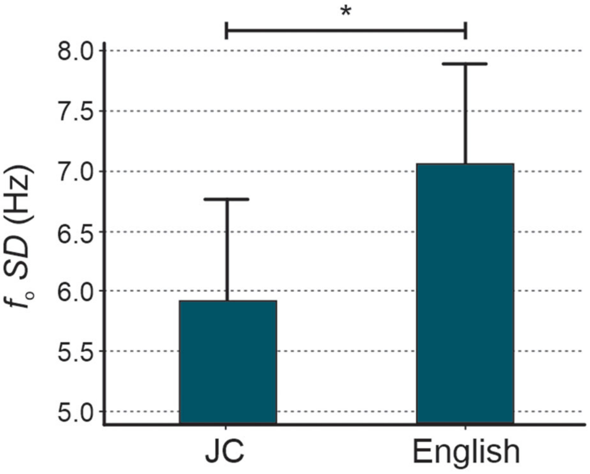

Figure 2. Mean and 95% confidence interval for fundamental frequency variability (f_o *SD*) by language. *= significant at $p < .5$.

Table III. Statistical results for mean fundamental frequency (f_o) and fundamental frequency variability (f_o *SD*).

Model	Variable	Coefficient	*SE* coefficient	*t* value	*p* value
Mean f_o	Language^a				
	JC	−2.005	2.669	−0.751	.454
	Vowel^b				
	/i/	11.045	2.727	4.051	<.001*
	JC exposure	−0.515	0.484	−1.271	.226
	JC use	−0.294	0.778	−0.378	.712
f_o *SD*	Language^a				
	JC	−0.047	0.019	−2.443	.016*
	Vowel^b				
	/i/	−0.009	0.020	−0.435	.664
	JC exposure	−0.001	0.001	−0.694	.500
	JC use	−0.002	0.002	−0.743	.471

Note. *SE* = standard error; *SD* = standard deviation; f_o = fundamental frequency; JC = Jamaican Creole.
^{a}Reference comparison is English.
^{b}Reference comparison is /u/.
*Significant at $p < .05$.

Table IV. Mean (*SD*) of each variable for Jamaican Creole (JC) and English.

Variable	JC	English
Mean f_o (Hz)	269.4 (48.5)	274.1 (52.8)
f_o *SD* (Hz)	5.9 (3.7)	7.1 (3.7)
Mean VOT (ms)		
Voiceless	64.1 (36.4)	63.7 (37.5)
Voiced (short lag)	16.4 (4.1)	18.4 (5.7)
Voiced (prevoiced)	−74.6 (38.4)	−61.9 (37.1)
VOT SD (ms)		
Voiceless	20.3 (16.2)	25.2 (21.3)
Voiced (short lag)	11.0 (4.2)	13.1 (6.5)
Voiced (prevoiced)	26.7 (11.3)	38.3 (28.0)
Prevoicing (%)	29.6 (19.1)	25.6 (17.9)

Note. *SD* = standard deviation; JC = Jamaican Creole; f_o = fundamental frequency; Hz = hertz; VOT = voice onset time; ms = millisecond.

f_o *SD* (JC) required log transformations to meet normality assumptions. Finally, the relationships among voiceless VOT *SD*, voiced (short lag, prevoiced) VOT *SD*, and f_o *SD* could be investigated within language. Significance was set to $p < .05$ and the coefficient of determination (r^2) was reported for each model. All statistical analyses were conducted using Minitab Statistical software (Ver. 20.3).

Result

Mean f_o and f_o SD

Mean f_o was not significantly different between JC and English ($p = .454$). The factor of vowel was significant ($p < .001$), in which the /i/ vowel showed higher f_o ($M = 280$ Hz) compared to /u/ ($M = 258$ Hz). JC exposure and use were not significant covariates for the mean f_o model.

The results for f_o *SD* showed a significant impact of language ($p = .016$), in which there was significantly greater f_o *SD* for English ($M = 7.06$ Hz) compared to JC ($M = 5.91$ Hz; Figure 2). However, vowel ($p = .664$), JC exposure ($p = .500$), and JC use ($p = .471$) were not significant predictors in this model. Please see Table III for more statistical information on the mean f_o and f_o *SD* models.

Mean VOT and VOT SD

Voiced (short lag) mean VOT ($p = .255$), voiced (prevoiced) mean VOT ($p = .378$), and voiceless mean VOT ($p = .951$) were not significantly different between the two languages in any model. For mean VOT values, please see Table IV. POA was not a significant factor for voiceless productions, and JC exposure as well as JC use were not significant predictors of mean VOT for any model (all $p > .05$).

Similar to mean VOT results, VOT *SD* was not affected by language for voiced (short lag; $p = .263$), voiced (prevoiced; $p = .242$), or voiceless ($p = .330$) productions. VOT *SD* was not impacted by POA for voiceless productions ($p > .05$). There was no impact of JC exposure or use on VOT *SD* for voiced (prevoiced) or voiceless models, but JC use was a significant, positive predictor of voiced (short lag) VOT *SD* ($p = .020$). Please see Table V for more information on each statistical model.

The proportions of prevoicing for voiced plosives were compared between JC and English, finding no significant difference ($p = .359$). The proportion of JC productions with prevoicing was 29.6%, whereas the proportion of English productions with prevoicing was 25.6%.

Relationship between phonemic and vocal variability by language

The relationships between voiced/voiceless VOT *SD* and f_o *SD* were investigated using simple linear regression models for each language. There were no significant results (all $p > .05$), meaning that there was no relationship between phonemic and vocal variability measures for either language. A range of no-to-weak relationships were found (r^2 = 1–6%). See Table VI for all regression model results.

Discussion

A dearth of evidence exists regarding bilingual acoustic measures during the language development of

Table V. Statistical results for mixed-effect models for mean voice onset time (VOT), voice onset time variability (VOT *SD*), and percentage of prevoicing.

Model	Variable	Coefficient	*SE* coefficient	*t* value	*p* value
Voiceless mean VOT	Language[a]				
	JC	0.164	2.694	0.061	.951
	POA[b]				
	Alveolar	4.873	3.592	1.356	.178
	Bilabial	−5.056	4.213	−1.200	.233
	JC exposure	−0.467	0.298	−1.568	.141
	JC use	−0.204	0.455	−0.448	.662
Voiced (short lag) mean VOT	Language[a]				
	JC	−1.000	0.846	−1.183	.255
	JC exposure	−0.015	0.045	−0.333	.745
	JC use	0.118	0.068	1.729	.107
Voiced (prevoiced) mean VOT	Language[a]				
	JC	−6.076	6.683	−0.909	.378
	JC exposure	−0.001	0.406	−0.002	.999
	JC use	−0.040	0.631	−0.063	.951
Voiceless VOT *SD*	Language[a]				
	JC	−0.031	0.032	−0.978	.330
	POA[b]				
	Alveolar	−0.001	0.042	−0.002	.998
	Bilabial	0.056	0.049	1.141	.256
	JC exposure	−0.001	0.002	−0.249	.807
	JC use	−0.002	0.003	−0.843	.415
Voiced (short lag) VOT *SD*	Language[a]				
	JC	−1.024	0.896	−1.143	.263
	JC exposure	−0.063	0.046	−1.359	.185
	JC use	0.174	0.070	2.474	.020*
Voiced (prevoiced) VOT *SD*	Language[a]				
	JC	−5.892	4.757	−1.238	.242
	JC exposure	−0.149	0.235	−0.633	.544
	JC use	−0.155	0.416	−0.373	.715
Percentage of prevoicing	Language[a]				
	JC	1.513	2.494	0.607	.553
	JC exposure	0.121	0.205	0.593	.563
	JC use	−0.324	0.312	−1.037	.319

Note. SE = standard error; VOT = voice onset time; JC = Jamaican Creole; POA = place of articulation; *SD* = standard deviation.
[a]Reference comparison is English.
[b]Reference comparison is velar.
*Significant at $p < .05$.

Table VI. Simple linear regression models examining the relationship between phonemic and vocal variability for each language. All models have the dependent outcome variable of fundamental frequency variability (f_o *SD*).

Language	Predictor variable	Coefficient	*SE* coefficient	*t* value	*p* value	r^2 (%)
JC	Voiced (short lag) VOT *SD*	−10.2	10.4	−0.97	.346	6.35
	Voiced (prevoiced) VOT *SD*	1.26	4.88	0.26	.802	0.66
	Voiceless VOT *SD*	−0.09	0.28	−0.30	.765	0.66
English	Voiced (short lag) VOT *SD*	25.3	69.2	0.37	.720	0.95
	Voiced (prevoiced) VOT *SD*	−14.2	18.0	−0.79	.450	5.82
	Voiceless VOT *SD*	−24.8	42.5	−0.58	.569	2.38

Note. SE = standard error; JC = Jamaican Creole; VOT = voice onset time; *SD* = standard deviation; r^2 = coefficient of determination.

bilingual children, with a greater disparity in how vocal development interacts with phoneme refinement. Our research examined voice acoustic measures (VOT, VOT *SD*, mean f_o, and f_o *SD*) in an understudied population, JC-English-speaking bilingual preschoolers. Based on the existing literature, we hypothesised a significant impact of language on phonemic and vocal variability measures as well as a significant positive relationship between the variability measures. Our results showed a significant impact of language on f_o *SD* with greater variability in English when compared to JC. We also determined that short lag VOT *SD* for voiced plosives was positively related to JC use. There were, however, no other acoustic measures significantly impacted by language, and no relationship was observed between phonemic and pitch variability.

Vocal variability

Vocal variability was measured by f_o variation (f_o *SD*) in our study and our results demonstrated greater f_o *SD* in English when compared to JC, indicating an impact of language on f_o variability. This result was expected and similar to previous literature. One study that was conducted by Sosa and Bunta (2019) showed that Spanish-English bilingual children exhibited increased f_o variability in their second language, English, when compared to their native language, Spanish. The researchers suggested that vocal variability could be influenced by language proficiency and the amount of exposure to each language (Sosa & Bunta, 2019). Our model included JC exposure and use as possible factors influencing f_o *SD*, but neither were significant predictors in the model.

In another study by Pépiot and Arnold (2021), English-French speakers presented greater f_o variation in English when compared to French, although English was their first language. These results could indicate linguistic features of English influencing the way f_o is being used within the language. However, our study did not directly evaluate how phonological context, word position, and stress patterns influenced f_o *SD*. That is, f_o *SD* was extracted from both monosyllabic and multisyllabic words in different stressed contexts (e.g. /i/ from *teeth* and /u/ from *kangaroo*). Indeed, we did find differences in mean f_o across vowels in which /i/ showed higher f_o compared to /u/, perhaps due to the variation in word position. Differences in f_o across vowels was also noted in previous research, where authors reported that /i/ had a higher f_o when compared to /u/ in different American dialects (Jacewicz & Fox, 2015). Future research could evaluate the impact of stress, word position, and inherent vowel f_o within languages to understand their influence on f_o variation.

Phonemic development

Unexpectedly, mean VOT and VOT *SD* were not impacted by language for voiced (short lag, prevoiced) or voiceless plosives. When comparing our results to previously reported data on bilingual adults (León et al., 2022, 2023), we observed that voiceless mean VOT resembled adults' productions in both languages. That is, children's VOT productions for JC and English in the present study were 64 ms and 64 ms, and adult productions in the study by Leon et al. were 74 ms and 74 ms, respectively (León et al., 2023). On the other hand, voiced mean VOT for the children in our study, averaged across both short lag and prevoiced productions, was 0 ms for both languages (i.e. $M = -0.5$ ms for JC; $M = -1.1$ ms for English), which neither matches adult bilingual productions in JC ($M = -40$ ms; León et al., 2023), nor English ($M = -37$ ms; León et al., 2023).

These VOT findings point towards a maturational effect on VOT because of the consistency found across both languages. Studies have shown that children develop phonemic proficiency gradually and do not necessarily master it by preschool, which can be evidenced by longer VOT productions in preschoolers when compared to adult productions (Khattab, 2004). Stoehr et al. (2018) reported that Dutch-German bilingual children (aged 3;7–5;11) produced voiced plosives similarly in their two languages, and these productions are not monolingual-like in either language. Another study by Khattab (2002) was conducted on VOT production across English and Arabic in bilingual and monolingual children aged 5–10 years. Results showed that bilingual speakers produced VOT values that were intermediate between monolingual Arabic and monolingual English speakers, indicating cross-linguistic influence. These productions did not resemble monolingual speakers' productions although they were still acquired gradually for each of their languages towards adult patterns.

Interestingly, short lag VOT *SD* was significantly related to JC use at home. That is, as JC use increased the variability of short lag productions also increased. Previous research has shown that adult bilingual JC-English speakers do not typically produce short lag VOT and instead prevoice almost all voiced VOT segments (León et al., 2023). The variability in children's short lag VOT may have been due to their attempts of reaching the adult target, which only successfully occurred ~27% of the time across both languages. When children produced prevoicing, they averaged −75 ms (JC) and −62 ms (English), which were similar to the adult form. Nevertheless, the pattern through which JC use influences the variability of targeted voice production throughout the development of children's phonemic systems is unclear, requiring further investigations.

It is well established that prevoicing, in particular, is difficult to master by the preschool years (Khattab, 2002; León et al., 2023). Researchers have suggested that this may be due to the increased complexity of muscle activity needed to maintain voicing during closure, requiring advanced neuromuscular control (Bortolini et al., 1995; Kewley-Port & Preston, 1974). We statistically examined the percentages of prevoicing in each language, finding no differences in the frequency of prevoicing across them, further bolstering the likelihood of maturational effects that are not based on language. A comparison of mean VOT and VOT *SD* between adults and children would be beneficial to further understand how mastery (defined as reduced variability with increased target accuracy) develops over time. Our work motivates a need for future research using a longitudinal design to track developmental changes in bilingual preschoolers as their vocal and articulatory systems change and become adult-like.

Variables of JC use and exposure

We further investigated how JC exposure and use may have influenced phonemic and vocal outcomes. We found JC use to be a positive predictor of short lag VOT *SD*; however, our results found no impact of exposure or use on any other acoustic measure. The lack of impact was unexpected as exposure and use are known to influence speech development for bilingual children (Coy & Watson, 2020; Khattab, 2004; León et al., 2023; Pépiot & Arnold, 2021). For example, a study conducted by Levy and Hanulíková (2019) showed that children who experience higher variable input exhibited increased variability in their productions of vowels. Although JC-English bilingual speakers in Jamaica usually acquire the languages simultaneously from birth (Meade, 2001), that does not necessarily mean that they are equally used or that input is the same.

In Jamaica, adults usually encourage the use of English to improve English linguistic abilities in children's early language development in support of their children's academic advancement. This practice is not uncommon amongst bilingual parents who demonstrate an interest in educational advancement and, as such, encourage the use of the majority language (McCarthy et al., 2014; Stoehr et al., 2018). Our study only evaluated parent-reported language exposure and use in the home setting; however, English use and exposure in the school setting could have been a factor since all children are required to speak English at school. All data were collected in the school setting, which may have influenced acoustic parametres (i.e. f_o *SD*, which was greater in English compared to JC), due to speaking environment. Individual differences in language use and exposure across different speaking environments for simultaneous bilinguals is a further area of inquiry, as it may prove beneficial to factor into SSD assessments for bilingual children.

Relationship between phonemic and vocal variability by language

We hypothesised a potential positive relationship between phonemic and f_o variability based on the overlapping contributions of vocal motor control to our acoustic measures. Specifically, we expected that children with greater f_o variation would also exhibit greater variability in VOT. As explained in the DIVA model (Tourville & Guenther, 2011), as accuracy increases with age, variability decreases in speech production. Hence, we anticipated that children with more variability in one speech system would also demonstrate variability across their other developing systems. However, our results did not support this hypothesis, as we did not find any statistically significant relationship between vocal and phonemic variability.

Interestingly, a study by Heller Murray and Chao (2021) reported a negative relationship between vocal variability (measured using the coefficient of variation of f_o) and vocal-articulatory control. Their findings revealed that increased VOT variability was associated with decreased vocal variability. It is important to note that methodological differences may partially explain the contrasting results with our results. In our study, we analysed shorter vowel segments from whole words in bilingual preschoolers, whereas Heller Murray and Chao (2021) focused on sustained vowels in monolingual preschoolers. The higher level of linguistic complexity employed in our study, in addition to a different linguistic population, may have revealed different results.

Taken together, our findings suggest that the vocal and phonemic systems may follow distinct developmental trajectories in bilingual preschoolers. Future research should further explore the interplay between vocal and phonemic variability, considering different methodological approaches and linguistic contexts to gain a comprehensive understanding of the developmental processes in bilingual speech motor control development.

Limitations and future directions

This research is not without its limitations, which should be addressed in future studies. First, the small sample size and limited population of preschoolers in Kingston, Jamaica restricts the generalisability of the results. Second, regarding the data collection environment, the school-based setting led to incidental background noise that interfered with audio recordings and resulted in the exclusion of seven participants. Nonetheless, we believe that data collection in the field for this population is representative of a naturalistic setting encountered by most SLPs during their speech sound assessments. Third, it is possible that repeated productions might have been tiring or even boring to children and could have influenced some speech productions (e.g. making them more monotone). Fourth, the phonological contexts were not controlled for and that could have impacted the results, especially f_o that is influenced by stressed and unstressed vowels. Finally, our data analyses were completed at the word level, which may not be representative of measures from continuous speech and more complex linguistic contexts. To overcome many of these limitations, larger studies across various speech contexts are needed. Further, studies should include English-speaking monolingual peers and include longer acquisition time frames to fully understand maturational speech effects in bilingual children.

Conclusion

Understanding typical patterns of speech development is crucial for improving the diagnostic process of SSDs in bilingual children. Our results revealed that JC-English preschoolers tend to exhibit mostly similar acoustic measures in JC and English, except for an increased f_o variability in English, which may be attributed to linguistic differences. Notably, voiced VOT did not consistently resemble the adult form in either language, suggesting maturational effects rather than a language-specific pattern. These findings align with previous research that children continue to refine the fine-grained motor development necessary for speech throughout the preschool years. Further cross-sectional research investigating the same linguistic pairing of JC-English is necessary to systematically explore the phonemic and vocal patterns observed here. Such research will provide valuable insights to be utilised in clinical settings, and contribute to advancements in automated speech recognition tools, enhancing their applicability in daily life for bilingual speakers.

Acknowledgements

The authors acknowledge Leslie E. Kokotek and other members of the Education Abroad Team to Jamaica for their efforts in data collection in Jamaica. We would like to acknowledge all school personnel who participated in the Jamaican Creole Language Project, preschoolers, and their families for their efforts and collaborations. Lastly, we thank Professors Laura and Richard Kretschmer and the Jamaican Language Unit for their ongoing support of our efforts.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Saudi Arabian Cultural Mission funding (A.S.M.); The American Speech-Language-Hearing Association (ASHA) Advancing Academic-Research Careers (AARC) Award (V.S.M.); The Jamaican Creole Language Project Endowment Fund (K.N.W); National Institutes on Deafness and Other Communication Disorders [R21 DC018170-01A1] (K.N.W).

References

Abu El Adas, S., Washington, K. N., Sosa, A., Harel, D., & McAllister, T. (2021). Variability across repeated productions in bilingual children speaking Jamaican Creole and English. *International Journal of Speech-Language Pathology*, *22*(6), 648–659. https://doi.org/10.1080/17549507.2020.1843712

Adi, Y., Keshet, J., Dmitrieva, O., & Goldrick, M. (2016). *Automatic measurement of voice onset time and prevoicing using recurrent neural networks* [Paper presentation]. Interspeech 2016, 3152–3155. https://doi.org/10.21437/Interspeech.2016-893

Altenberg, E. P., & Ferrand, C. T. (2006). Fundamental frequency in monolingual English, bilingual English/Russian, and bilingual English/Cantonese young adult women. *Journal of Voice: Official Journal of the Voice Foundation*, *20*(1), 89–96. https://doi.org/10.1016/j.jvoice.2005.01.005

American Speech-Language-Hearing Association. (n.d). *Bilingual service delivery (practice portal)*. https://www.asha.org/Practice-Portal/Professional-Issues/Bilingual-Service-Delivery/

American Speech–Language–Hearing Association. (2022). *2021 demographic profile of ASHA members providing multilingual services*. https://www.asha.org/siteassets/surveys/demographic-profile-bilingual-spanish-service-members.pdf

American Speech-Language-Hearing Association. (2017). *Issues in ethics: Cultural and linguistic competence*. https://www.asha.org/practice/ethics/cultural-and-linguistic-competence/#sec1.1

Blake, J., & Boysson-Bardies, B. D. (1992). Patterns in babbling: A cross-linguistic study. *Journal of Child Language*, *19*(1), 51–74. https://doi.org/10.1017/S0305000900013623

Boersma, P. (2001). Praat, a system for doing phonetics by computer. (6.2.15). *Glot International*, *5*(9/10), 341–345.

Bortolini, U., Zmarich, C., Fior, R., & Bonifacio, S. (1995). Word-initial voicing in the productions of stops in normal and preterm Italian infants. *International Journal of Pediatric Otorhinolaryngology*, *31*(2–3), 191–206. https://doi.org/10.1016/0165-5876(94)01091-B

Bosch, L., & Ramon-Casas, M. (2011). Variability in vowel production by bilingual speakers: Can input properties hinder the early stabilization of contrastive categories? *Journal of Phonetics*, *39*(4), 514–526. https://doi.org/10.1016/j.wocn.2011.02.001

Callan, D. E., Kent, R. D., Guenther, F. H., & Vorperian, H. K. (2000). An auditory-feedback-based neural network model of speech production that is robust to developmental changes in the size and shape of the articulatory system. *Journal of Speech, Language, and Hearing Research: JSLHR*, *43*(3), 721–736. https://doi.org/10.1044/jslhr.4303.721

Cassidy, F. G., & Le Page, R. B. (1980). *Dictionary of Jamaican English* (2d ed.). Cambridge University Press.

Castilla-Earls, A., Bedore, L., Rojas, R., Fabiano-Smith, L., Pruitt-Lord, S., Restrepo, M. A., & Peña, E. (2020). Beyond scores: Using converging evidence to determine speech and language services eligibility for dual language learners. *American Journal of Speech-Language Pathology*, *29*(3), 1116–1132. https://doi.org/10.1044/2020_AJSLP-19-00179

Coy, A., & Watson, S. (2020). Acoustic similarity of inner and outer circle varieties of child-produced English vowels. *Journal of Speech, Language, and Hearing Research*, *63*(3), 722–737. https://doi.org/10.1044/2019_JSLHR-19-00179

Devonish, H., & Carpenter, K. (2007). Towards full bilingualism in education: The Jamaican bilingual primary education project. *Social and Economic Studies*, *56*(1/2), 277–303.

Devonish, H., & Harry, O. G. (2008). Jamaican Creole and Jamaican English: Phonology. In B. Kortmann & E. W. Schneider (Eds.), *The Americas and the Caribbean* (pp. 256–289). Mouton de Gruyter. https://doi.org/10.1515/9783110208405.1.256

Dixit, R. P. (1989). Glottal gestures in Hindi plosives. *Journal of Phonetics*, *17*(3), 213–237. https://doi.org/10.1016/S0095-4470(19)30431-0

Dodd, B., Hua, Z., Crosbie, S., Holm, A., & Ozanne, A. (2006). *Diagnostic evaluation of articulation and phonology*. Harcourt Assessment.

Ehrler, D. J., & McGhee, R. L. (2008). *Primary test of nonverbal intelligence*. Pro-Ed.

Evans, S., Neave, N., Wakelin, D., & Hamilton, C. (2008). The relationship between testosterone and vocal frequencies in human males. *Physiology & Behavior*, *93*(4–5), 783–788. https://doi.org/10.1016/j.physbeh.2007.11.033

Federal Interagency Forum on Child and Family Statistics. (2019). *America's children: Key national indicators of well-being, 2019*. U.S. Government Printing Office. https://www.childstats.gov/pdf/ac2019/ac_19.pdf

Gelfer, M. P., & Denor, S. L. (2014). Speaking fundamental frequency and individual variability in Caucasian and African American school-age children. *American Journal of Speech-Language Pathology*, *23*(3), 395–406. https://doi.org/10.1044/2014_AJSLP-13-0016

Hamilton, N. J. I. (2023). The life-cycle and restoration of the human vocal fold. *Laryngoscope Investigative Otolaryngology*, *8*(1), 168–176. https://doi.org/10.1002/lio2.993

Hartnick, C. J., Rehbar, R., & Prasad, V. (2005). Development and maturation of the pediatric human vocal fold lamina propria. *The Laryngoscope*, *115*(1), 4–15. https://doi.org/10.1097/01.mlg.0000150685.54893.e9

Heller Murray, E. S., & Chao, A. (2021). The relationships among vocal variability, vocal-articulatory coordination, and dysphonia in children. *Journal of Voice: Official Journal of the Voice Foundation*, *37*(6), 969.e43–969.e49. https://doi.org/10.1016/j.jvoice.2021.06.008

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*(7302), 29–29. https://doi.org/10.1038/466029a

Herd, W. (2020). Sociophonetic voice onset time variation in Mississippi English. *Journal of the Acoustical Society of America*, *147*(1), 596–605. https://doi.org/10.1121/10.0000545

Irizarry-Pérez, C. D., Fabiano-Smith, L., & Martinez-Fisher, A. (2023). A complexity approach to promoting within- and cross-language generalization in bilingual children with phonological delays. *Language, Speech, and Hearing Services in Schools*, *54*(3), 799–814. https://doi.org/10.1044/2023_LSHSS-22-00128

Jacewicz, E., & Fox, R. A. (2015). Intrinsic fundamental frequency of vowels is moderated by regional dialect. *Journal of* *the Acoustical Society of America*, *138*(4), EL405–EL410. https://doi.org/10.1121/1.4934178

Kellogg, J., & Chang, C. B. (2023). Exploring the onset of phonetic drift in voice onset time perception. *Languages*, *8*(1), 78. https://doi.org/10.3390/languages8010078

Kewley-Port, D., & Preston, M. S. (1974). Early apical stop production: A voice onset time analysis. *Journal of Phonetics*, *2*(3), 195–210. https://doi.org/10.1016/S0095-4470(19)31270-7

Khattab, G. (2002). VOT production in English and Arabic bilingual and monolingual children. In D. B. Parkinson & E. Benmamoun (Eds.), *Current issues in linguistic theory* (Vol. 230, pp. 1–37). John Benjamins Publishing Company. https://doi.org/10.1075/cilt.230.03kha

Khattab, G. (2004). Developmental and communicative factors affecting VOT production in English and Arabic bilingual and monolingual speakers. *Journal of the Acoustical Society of America*, *115*(5_Supplement), 2503–2503. https://doi.org/10.1121/1.4783040

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lee, B., & Sidtis, D. V. L. (2017). The bilingual voice: Vocal characteristics when speaking two languages across speech tasks. *Speech, Language and Hearing*, *20*(3), 174–185. https://doi.org/10.1080/2050571X.2016.1273572

Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America*, *105*(3), 1455–1468. https://doi.org/10.1121/1.426686

León, M., Washington, K. N., McKenna, V. S., Crowe, K., & Fritz, K. (2022). Linguistically informed acoustic and perceptual analysis of bilingual children's speech productions: An exploratory study in the Jamaican Context. *Journal of Speech, Language, and Hearing Research: Jslhr*, *65*(7), 2490–2509. https://doi.org/10.1044/2022_JSLHR-21-00386

León, M., Washington, K. N., McKenna, V. S., Crowe, K., Fritz, K., & Boyce, S. (2023). Characterizing speech sound productions in bilingual speakers of Jamaican Creole and English: Application of durational acoustic methods. *Journal of Speech, Language, and Hearing Research: JSLHR*, *66*(1), 61–83. https://doi.org/10.1044/2022_JSLHR-22-00304

Levy, H., & Hanulíková, A. (2019). Variation in children's vowel production: Effects of language exposure and lexical frequency. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *10*(1), 9. https://doi.org/10.5334/labphon.131

Lisker, L. (1978). On learning a new contrast. In M. A. Jazayery, E. C. Polomé, & W. Winter (Eds.), *Linguistic and Literary Studies, Vol 1, General and Theoretical Linguistics* (pp. 201–216). De Gruyter Mouton. https://doi.org/10.1515/9783110808995.201

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *WORD*, *20*(3), 384–422. https://doi.org/10.1080/00437956.1964.11659830

Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech*, *10*(1), 1–28. https://doi.org/10.1177/002383096701000101

Löfqvist, A. (1980). Interarticulator programming in stop production. *Journal of Phonetics*, *8*(4), 475–490. https://doi.org/10.1016/S0095-4470(19)31502-5

Mair, C. (2003). Language, code, and symbol: The changing roles of Jamaican Creole in diaspora communities. *AAA: Arbeiten Aus Anglistik Und Amerikanistik*, *28*(2), 231–248. http://www.jstor.org/stable/43025702

McCarthy, K. M., Mahon, M., Rosen, S., & Evans, B. G. (2014). Speech perception and production by sequential bilingual children: A longitudinal study of voice onset time acquisition. *Child Development*, *85*(5), 1965–1980. https://doi.org/10.1111/cdev.12275

McKenna, V. S., Hylkema, J. A., Tardif, M. C., & Stepp, C. E. (2020). Voice onset time in individuals with hyperfunctional voice disorders: Evidence for disordered vocal motor control. *Journal of Speech, Language, and Hearing Research: JSLHR*, *63*(2), 405–420. https://doi.org/10.1044/2019_JSLHR-19-00135

McLeod, S., Harrison, L. J., & McCormack, J. (2012). The intelligibility in context scale: Validity and reliability of a subjective rating measure. *Journal of Speech, Language, and Hearing Research: JSLHR*, *55*(2), 648–656. https://doi.org/10.1044/1092-4388(2011/10-0130)

Meade, R. R. (2001). *Acquisition of Jamaican phonology*. Foris Publication.

Mennen, I., Schaeffler, F., & Docherty, G. (2012). Cross-language differences in fundamental frequency range: A comparison of English and German. *Journal of the Acoustical Society of America*, *131*(3), 2249–2260. https://doi.org/10.1121/1.3681950

Modolo, D. J., Berretin-Felix, G., Genaro, K. F., & Brasolotto, A. G. (2011). Oral and vocal fold Diadochokinesis in children. *Folia Phoniatrica et Logopaedica: Official Organ of the International Association of Logopedics and Phoniatrics (IALP))*, *63*(1), 1–8. https://doi.org/10.1159/000319728

Ng, M. L., Hsueh, G., & Sam Leung, C.-S. (2010). Voice pitch characteristics of Cantonese and English produced by Cantonese-English bilingual children. *International Journal of Speech-Language Pathology*, *12*(3), 230–236. https://doi.org/10.3109/17549501003721080

Office of Head Start. (2016). Services snapshot (2015–2016). National all programs. https://eclkc.ohs.acf.hhs.gov/sites/default/files/pdf/service-snapshot-HS-2015-2016.pdf

Ohde, R. N. (1985). Fundamental frequency correlates of stop consonant voicing and vowel quality in the speech of preadolescent children. *Journal of the Acoustical Society of America*, *78*(5), 1554–1561. https://doi.org/10.1121/1.392791

Oller, D. K. (2000). *The emergence of the speech capacity*. Lawrence Erlbaum Associates.

Pépiot, E., & Arnold, A. (2021). Cross-gender differences in English/French bilingual speakers: A multiparametric study. *Perceptual and Motor Skills*, *128*(1), 153–177. https://doi.org/10.1177/0031512520973514

Rothgänger, H. (2003). Analysis of the sounds of the child in the first year of age and a comparison to the language. *Early Human Development*, *75*(1–2), 55–69. https://doi.org/10.1016/j.earlhumdev.2003.09.003

Sosa, A. V., & Bunta, F. (2019). Speech production accuracy and variability in monolingual and bilingual children with cochlear implants: A comparison to their peers with normal hearing. *Journal of Speech, Language, and Hearing Research: JSLHR*, *62*(8), 2601–2616. https://doi.org/10.1044/2019_JSLHR-S-18-0263

Stoehr, A., Benders, T., Van Hell, J. G., & Fikkert, P. (2018). Heritage language exposure impacts voice onset time of Dutch–German simultaneous bilingual preschoolers. *Bilingualism: Language and Cognition*, *21*(3), 598–617. https://doi.org/10.1017/S1366728917000116

Swartz, B. L. (1992). Gender difference in voice onset time. *Perceptual and Motor Skills*, *75*(3), 983–992. https://doi.org/10.2466/pms.1992.75.3.983

Theelen, M. (2017). Fundamental frequency differences including language effects. *Junctions: Graduate Journal of the Humanities*, *2*(1), 9. https://doi.org/10.33391/jgjh.25

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, *26*(7), 952–981. https://doi.org/10.1080/01690960903498424

United States Census Bureau. (2020). https://data.census.gov/table?tid=ACSDP5Y2019.DP02&hidePreview=true

Vorperian, H. K., & Kent, R. D. (2007). Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research: JSLHR*, *50*(6), 1510–1545. https://doi.org/10.1044/1092-4388(2007/104)

Washington, K. N., Fritz, K., Crowe, K., Kelly, B., & Wright Karem, R. (2019). Bilingual preschoolers' spontaneous productions: Considering Jamaican Creole and English. *Language, Speech, and Hearing Services in Schools*, *50*(2), 179–195. https://doi.org/10.1044/2018_LSHSS-18-0072

Washington, K. N., McDonald, M. M., McLeod, S., Crowe, K., & Devonish, H. (2017). Validation of the intelligibility in context scale for Jamaican Creole-speaking preschoolers. *American Journal of Speech-Language Pathology*, *26*(3), 750–761. https://doi.org/10.1044/2016_AJSLP-15-0103

Washington, K. N., Karem Wright, R., Kokotek, L., & León, M. (2023). Supporting culturally responsive practices with preschoolers: Guidance from methods in the Jamaican context. *Journal of Speech, Language, and Hearing Research: JSLHR*, 1–23. https://doi.org/10.1044/2023_JSLHR-23-00106

Wuyts, F. L., Heylen, L., Rooman, R., Mertens, F., Van De Heyning, P. H., Caju, M. D., & De Bodt, M. (2003). Effects of age, sex, and disorder on voice range profile characteristics of 230 children. *Annals of Otology, Rhinology, and Laryngology*, *112*(6), 540–548. https://doi.org/10.1177/000348940311200611

Zen, E. L. (2019). On the acquisition of English voiceless stop VOT by Indonesian-English bilinguals: Evidence of input frequency. *k@ta*, *20*(2), 45–52. https://doi.org/10.9744/kata.20.2.45-52