ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Multilingual Speech Acquisition by Vietnamese-English–Speaking Children and Adult Family Members

Sharynne McLeod,[a] Sarah Verdon,[a] Kate Margetson,[a] Van H. Tran,[a] Cen Wang,[a] Ben Phạm,[a,b] Lily To,[a] and Kylie Huynh[a]

[a] Charles Sturt University, Bathurst, New South Wales, Australia [b] Hanoi National University of Education, Viet Nam

ARTICLE INFO

Article History:
Received December 19, 2021
Revision received December 29, 2022
Accepted March 21, 2023

Editor-in-Chief: Bharath Chandrasekaran
Editor: Kate Bunton

https://doi.org/10.1044/2023_JSLHR-21-00669

ABSTRACT

Purpose: This article presents a large-scale example of culturally responsive assessment and analysis of multilingual Vietnamese-English–speaking children and their family members using the VietSpeech Protocol involving (a) examining all spoken languages, (b) comparing ambient phonology produced by family members, (c) including dialectal variants in the definition of accuracy, and (d) clustering participants with similar language experience.
Method: The VietSpeech participants (N = 154) comprised 69 children (2;0–8;10 [years;months]) and 85 adult family members with Vietnamese heritage living in Australia. Speech was sampled using the Vietnamese Speech Assessment (Vietnamese) and the Diagnostic Evaluation of Articulation and Phonology (English).
Results: Children's Vietnamese consonant accuracy was significantly higher when dialectal variants were accepted (percentage of consonants correct–dialect [PCC-D]: M = 87.76, SD = 8.18), compared to when only Standard Vietnamese was accepted as the correct production (percentage of consonants correct–standard [PCC-S]: M = 70.34, SD = 8.78), Cohen's d = 3.55 (large effect). Vietnamese voiced plosives, nasals, semivowels, vowels, and tones were more often correct than voiceless plosives and fricatives. Children's Standard Australian English consonant accuracy (PCC-S) was 82.51 (SD = 15.57). English plosives, nasals, glides, and vowels were more often correct than fricatives and affricates. Vietnamese word-initial consonants had lower accuracy than word-final consonants, whereas English consonant accuracy was rarely influenced by word position. Consonant accuracy and intelligibility were highest for children with high proficiency in both Vietnamese and English. Children's consonant productions were most similar to their mothers' than other adults or siblings' productions. Adults' Vietnamese consonants, vowels, and tones were more likely to match Vietnamese targets than their children's productions.
Conclusions: Children's speech acquisition was influenced by cross-linguistic, dialectal, maturational, language experience, and environmental (ambient phonology) factors. Adults' pronunciation was influenced by dialectal and cross-linguistic factors. This study highlights the importance of including all spoken languages, adult family members, dialectal variants, and language proficiency to inform differential diagnosis of speech sound disorders and identify clinical markers in multilingual populations.
Supplemental Material: https://doi.org/10.23641/asha.23290055

Correspondence to Sharynne McLeod: smcleod@csu.edu.au. ***Disclosure:*** *Ben Phạm and Sharynne McLeod are co-authors of the Vietnamese Speech Assessment. Sharynne McLeod is co-author of the Intelligibility in Context Scale. All other authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

Children's speech acquisition has been a focus of research and practice since the beginning of the speech-language pathology profession (e.g., Poole, 1934). There has been extensive cross-linguistic research that has considered monolingual acquisition of English and many

other languages (Crowe & McLeod, 2020; McLeod & Crowe, 2018). Internationally, this information about monolingual speech acquisition underpins speech-language pathologists' (SLPs') assessment, analysis, diagnosis, intervention, and service delivery practices. SLPs report lacking confidence in the assessment of multilingual children (Kritikos, 2003; Williams & McLeod, 2012) and over- and underdiagnose communication disorders in multilingual children (Bedore & Peña, 2008; Stow & Dodd, 2005) due to difficulties differentiating between typical multilingual speech and multilingual children with speech sound disorder (SSD). Comprehensive, culturally responsive data are needed to identify clinical markers for SSD among multilingual children to supplement resources that support SLPs' assessment and diagnosis (Hopf et al., 2021; McLeod, Baker, et al., 2017). Data are required from children speaking diverse language combinations so that SLPs avoid classifying typically developing multilingual/multidialectal children as disordered (Farrugia-Bernard, 2018), or not identifying children who are in genuine need of intervention (Margetson, McLeod, Verdon, & Tran, 2023).

More than half of the people in the world are multilingual (Grosjean, 2021); that is, they speak, sign, understand, read, or write more than one language. Being multilingual is important for identity, family connections, social and cultural cohesion, executive brain function, academic and occupational advantage, and national prosperity in an increasingly globalized world (Bialystok, 2011; Blake et al., 2018; Clyne, 2008; Klein et al., 2016; Kroll & Dussias, 2017; McLeod et al., 2016; Oh & Fuligni, 2010). Children benefit from a level of competency in their languages before the advantages of multilingualism can be realized (Goldfeld et al., 2014; Nguyen et al., 2001). Many nations are not fully using their rich multilingual resources because many children are not maintaining their home languages, especially after they commence formal schooling (Verdon, McLeod, & Winsler, 2014). Therefore, it is important to support linguistic multicompetence among multilingual children and provide speech-language pathology services when necessary (McLeod, Verdon, et al., 2022).

Historically, a monolingual Western (English) mindset has underpinned a large amount of the world's research (Márquez & Porras, 2020; Meneghini & Packer, 2007; Stenius et al., 2017), including speech-language pathology research and practice (Easton & Verdon, 2021; Ellis et al., 2021; MacLeod & Demers, 2023). This century, SLPs have paid greater attention toward multilingual speech acquisition in research and practice (Gildersleeve-Neumann & Wright, 2010; Goldstein & Iglesias, 2001; Hambly et al., 2013; Hopf et al., 2021; Li & To, 2017). For example, Supplemental Material S1 lists 32 studies of multilingual speech acquisition identified by members of the International Expert Panel on Multilingual Children's Speech. While inclusion of multilingual participants is important, most of these studies only described acquisition of one language (typically English—not both/all languages) and did not take into account children's proficiency/exposure to each language, ambient phonology, or dialectal variants. Studies of bilingual Spanish-English–speaking children within the United States have paved the way for a paradigm shift in SLPs' research and practice with multilingual children (e.g., Fabiano-Smith & Barlow, 2010; Fabiano-Smith & Goldstein, 2010; Goldstein & Washington, 2001). For example, Fabiano-Smith and Barlow (2010) studied both languages spoken by 24 children and found evidence for bidirectional cross-linguistic transfer, with children acquiring consonant inventories for both languages in the same time as monolingual English-speaking children acquired one. Furthermore, Ruiz-Felter et al. (2016) demonstrated that consideration of language experience (input–output) was a better predictor of consonant accuracy than age of first exposure to English.

Historically, SLPs' judgments of "correct" speech productions (e.g., on test forms) have matched the standard dialect of a country or region. However, the role of SLPs is to consider children's communicative competence rather than how closely their speech adheres to the standardized norm. The standard dialect provides socially constructed "correct" and "error" speech sound productions. Classifying "errors" as all speech productions that do not match the standardized norm can lead to misdiagnosis, discrimination, and marginalization of speakers of nonstandard forms (Easton & Verdon, 2021; Farrugia-Bernard, 2018) and does not take into consideration individuals' milieu or societal evolution of languages and dialects (cf. To et al., 2015). SLPs need to use two culturally responsive definitions of "correct" speech sound productions: (a) standard productions and (b) productions that consider cross-linguistic, dialectal, environmental, and maturational influences. If a production does not fit into either definition of "correct," then SLPs can be more confident that a clinical condition may underlie a multilingual child's ability to communicate, diagnose an SSD, and provide clinical intervention.

Theoretical Frameworks

An in-depth understanding of multilingual children's typical speech acquisition enables SLPs to know what maturation patterns to expect, to inform diagnosis (specifically distinguishing between speech difference versus SSD), to predict between-language interaction, and to classify severity (Goldstein & Iglesias, 2001; Sabri & Fabiano-Smith, 2018; Stoehr et al., 2022). Theoretically, knowledge about how multilingual children's languages interact can inform models of speech and language acquisition. The current research applied three theoretical

frameworks: the emergence approach (Davis & Bedore, 2013), the theory of linguistic multicompetence (Cook, 2016), and the Culturally Responsive Teamwork Framework (Hopf et al., 2021). The emergence approach espouses, "a child's acquisition of phonology is seen as a product of her physical and social interaction capacities, supported by input from adult models about ambient language sound patterns" (n.p.). The emergence approach considers the multiple physical, cognitive, and social processes involved in children's acquisition of phonological features in their ambient language(s). From an emergence theory perspective, speech acquisition occurs through an integrative process of perceiving, organizing, storing, and retrieving phonological information from diverse communicative models as children move toward intelligible speech. Linguistic multicompetence is "the overall system of a mind or a community that uses more than one language" (Cook, 2016, p. 2) and recognizes that all languages within a person's repertoire are interrelated. Specifically, (a) multicompetence concerns the knowledge of more than one language and relationships between languages, (b) multicompetent speakers should not be compared to monolingual native speaker norms, and (c) multicompetent speakers perceive and encode information differently from monolingual speakers (Cook, 2016).

The four interacting levels of the Culturally Responsive Teamwork Framework (Hopf et al., 2021) also informed the current research. Specifically, (a) intrapersonal practices involved the research team undergoing critical self-examination and continuing professional development, (b) interpersonal practices involved "culturally responsive collaborations with clients, colleagues and communities" (p. 1953), (c) intraprofessional practices involved "drawing on resources from across the SLP profession in order to increase the diversity of an individual SLP's practice to meet the diverse needs of their clients" (p. 1954), and (d) interprofessional practices involved to "acknowledge and respect the cultural and linguistic diversity of each other…act with cultural sensitivity and safety and to upskill in multicultural collaboration" (p. 1955).

VietSpeech Protocol

The current research is a large-scale example of a culturally responsive assessment of multilingual children supporting a paradigm shift within speech-language pathology by employing the VietSpeech Protocol (see the Appendix) involving (a) examining both languages spoken within the children's cultural community, (b) considering ambient phonology by assessing children and adult family members in both languages, (c) defining accuracy using both the standard form of the language (e.g., percentage of consonants correct–standard [PCC-S]) and the dialectal variants of the language (e.g., percentage of consonants correct–dialect [PCC-D]), and (d) analyzing statistically identified clusters of children and adults based on their proficiency in both languages. This paradigm shift is informed by the emergence approach of speech acquisition (Davis & Bedore, 2013), the theory of linguistic multicompetence (Cook, 2016), and the Culturally Responsive Teamwork Framework (Hopf et al., 2021). The current research focused on the speech acquisition of Vietnamese-English–speaking children, a language pair that is linguistically distant and rarely studied, despite Vietnamese and English being in the top 20 languages spoken in the world. The research involved examination of both Vietnamese and English speech samples, and ambient phonology was considered by comparing the speech of children and their adult family members. Accuracy was defined using the Standard and dialectal forms of Vietnamese and English. Three clusters of children and adults were statistically identified based on their proficiency in both languages (Crowther et al., 2021). This landmark study of Vietnamese and English speech production aimed to demonstrate the influence of cross-linguistic, dialectal, maturational, language experience, and environmental (ambient phonology) factors on children's speech acquisition to inform understanding of typical development and SLPs' differential diagnosis of SSD in multilingual children.

Vietnamese

Vietnamese is one of the top 20 most commonly spoken languages across the world (Eberhard et al., 2023). As the official language in Vietnam, Vietnamese is spoken by the majority of the 96 million people across the country (General Statistics Office of Viet Nam, 2020). Due to the Vietnamese diaspora, Vietnamese is the third most commonly used home language in Australia after English (Australian Bureau of Statistics, 2021a), the fourth most commonly spoken home language other than English in the United States (Camarota & Zeigler, 2014; Ryan, 2013), one of the top 25 languages spoken in Canada (Statistics Canada, 2012), and Vietnamese people are the third largest foreigner group in the Czech Republic (Czech Statistics Office, 2011).

Vietnamese has many features that differ from English (see Table 1). Vietnamese is a tonal monosyllabic language. All Vietnamese syllables contain a main vowel (V) and tone (T). Optional components in syllables are the initial consonant (C), final consonant (C), medial semivowel (w), and final semivowel (w). Therefore, the Vietnamese syllable structure is $(C_{0-1})(w_{0-1})V(C_{0-1}/w_{0-1})T$ (Cao, 2006; Đoàn, 2003; Hoàng, 2004; Kirby, 2011; Phạm & McLeod, 2016). Standard Vietnamese refers to the Vietnamese dialect with standard pronunciation (*chuẩn phát âm*) regulated by

Table 1. Comparison between Standard Australian English and Standard Vietnamese consonants.

Manner	Language	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Glottal
Plosive	English	p b			t d				k ɡ	
	Vietnamese	(p) b			t ʈʰ d		ʈ	c	k [kᵖ]	(ʔ)
Nasal	English	m			n				ŋ	
	Vietnamese	m			n			ɲ	ŋ [ŋᵐ]	
Trill	English									
	Vietnamese				(r)					
Fricative	English		f v	θ ð	s z	ʃ ʒ				h
	Vietnamese		f v		s z		ʂ ʐ		x ɣ	h
Affricate	English							ʧ ʤ		
	Vietnamese				[ʦ]					
Approximant semivowel	English	w			ɹ			j		
	Vietnamese	w						j		
Lateral approximant	English				l					
	Vietnamese				l					

Note. In Vietnamese, dialectal variants are presented in square brackets, and contested symbols are in rounded brackets. Data at the left side of the cells are voiceless consonants, and those at the right side or center of the cells are voiced.

the Vietnamese government to be used in education, politics, the military, economy, society, culture, science, and the arts. Standard Vietnamese has 23 syllable-initial consonants /(p), b, ʈʰ, t, d, ʈ, c, k, (ʔ), m, n, ɲ, ŋ, f, v, s, z, ʂ, ʐ, x, ɣ, h, l/, two semivowels /w, j/, and six syllable-final consonants /p, t, k, m, n, ŋ/ (Đoàn, 2003; Phạm & McLeod, 2016; see Table 1; throughout this article, variants are presented in square brackets, and contested symbols are presented in round brackets). Standard Vietnamese has 14 vowels, including nine long singleton vowels /i, e, ɛ, ɯ, u, o, ɔ, ɤ, ɑ/, two short singleton vowels /ă, ɤ̆/, and three diphthongs /ie, uo, ɯɤ/ (Cao, 2006; Hoàng, 2004). Standard Vietnamese has six tones, namely, (1) level (*thanh không dấu* or *thanh ngang*), (2) falling (*thanh huyền*), (3) creaky (*thanh ngã*), (4) dipping–rising (*thanh hỏi*), (5) rising (*thanh sắc*), and (6) constricted (*thanh năng*); Cao, 2006; Đỗ & Lê, 2005; Đoàn, 2003; Lã et al., 2011; Nguyễn & Edmondson, 1998; Pham, 2003; Phạm & McLeod, 2016, 2019; Tang & Barlow, 2006). Two additional tones (7) and (8), which can be phonologically considered as allophones of tones (5) rising and (6) constricted, are present in syllable-final voiceless stop positions (e.g., *bất cập*; Phạm & McLeod, 2016).

Prominent spoken dialects of Vietnamese (*phương ngữ*) vary in vocabulary and phonology and include Northern, Central, and Southern Vietnamese, although up to nine dialects have been documented (Đinh & Nguyễn, 1998; Đoàn, 2003; Hoàng, 2004; Huỳnh, 2014; Nguyễn, 1997). Some consonants, vowels, and tones vary between Standard, Northern, Central, and Southern Vietnamese (see Phạm & McLeod, 2016, and Supplemental Material S2). There are 23 initial consonants in the Standard and Central dialects, 21 in the Southern dialect, and 20 in the Northern dialect. These 13 consonants /b, ʈʰ, t, d, m, n, ɲ, ŋ, f, s, x, ɣ, l/ have similar pronunciations for Standard, Northern, Central, and Southern Vietnamese. There is dialectal variability in the production of the additional seven to 10 consonants. For example, "r" is pronounced as the retroflex /ʐ/ in Standard and Central dialects, as the voiced alveolar fricative /z/ and the trilled /r/ in the Northern dialect, and as the voiced velar fricative /ɣ/ in the Southern dialect. Similarly, "v" is pronounced as the voiced bilabial fricative /v/ in the Standard, Northern, and Central dialects; as the voiced palatal approximant /j/ in the Southern dialect; and as the voiceless labiodental fricative /f/ in the lower Northern dialect (Ninh Binh dialect; Phạm & McLeod, 2016).

English

While English is one of the most commonly spoken languages in the world, many speakers use English as a second or subsequent language (Eberhard et al., 2023). In English, the main vowel (V) is compulsory, whereas initial and final consonants (C) are optional. English allows up to three consonants within a word-initial consonant cluster and up to four consonants within a word-final consonant cluster. Therefore, the English syllable shape can be written as $(C_{0–3})V(C_{0–4})$. English contains many polysyllabic words. English dialects contain similar consonants but differ in the number and type of vowels and diphthongs.

Consonants in Standard Australian English are similar to many other English dialects (see Table 1). Australian English contains 23 syllable-initial consonants /p, b, t, d, k, g, m, n, f, v, θ, ð, s, z, ʃ, ʒ, h, ʧ, ʤ, ɹ, j, l, w/ and 20 syllable-final consonants /p, b, t, d, k, g, m, n, ŋ, f, v, θ, ð, s, z, ʃ, ʒ, ʧ, ʤ, l/. Of note, Standard Australian English does not include syllable-final /ɹ/ (cf. U.S. English and Scottish English). Like other English dialects, Standard Australian English has the syllable shape $C_{(0–3)}VC_{(0–4)}$ and has many syllable-initial and syllable-final consonant clusters and polysyllabic words.

Comparison Between Vietnamese and English

Table 1 demonstrates the shared consonants between Vietnamese and English /(p), b, t, d, k, m, n, ŋ, f, v, s, z, h, w, j, l/ and the unshared consonants for Vietnamese /ʈʰ, c, ʈ, [kᵖ], (ʔ), ɲ, [ŋᵐ], (r), ʂ, ʐ, x, ɣ, [ʦ]/ and English /g, θ, ð, ʃ, ʒ, ʧ, ʤ, ɹ/. Vietnamese includes semivowels, typically has monosyllabic words, and is known as a syllable-timed language. In contrast, Australian English has approximants, has consonant clusters, has many bisyllabic and polysyllabic words (e.g., *hippopotamus* has five syllables), and is known as a stress-timed language. While Vietnamese uses tones to differentiate meaning, English does not.

VietSpeech Project

Australia is a diverse multicultural country, with 51.5% of the population having one or more parent born overseas and 22.3% speaking one or more of 300 languages other than English at home (Australian Bureau of Statistics, 2021b), with Vietnamese as the third most commonly spoken language after English (Australian Bureau of Statistics, 2021a). Like many other English-speaking nations, Australian SLPs are largely monolingual English speakers (Verdon, McLeod, & McDonald, 2014; Nancarrow et al., 2023) and need additional support to differentially diagnose multilingual children. This article presents results from a large-scale research project titled *VietSpeech: Vietnamese–Australian Children's Speech and Language Competence* funded by the Australian Research Council Discovery Program (DP180102848). The overarching aim of

the VietSpeech Project was to explore Vietnamese–English language proficiency, use, and maintenance in Australia. The objectives of the VietSpeech Project were to "(1) support Vietnamese–Australian children and their families to maintain their home language, (2) enhance their speech skills in Vietnamese and English, and (3) equip English-speaking professionals to support multilingual children's speech acquisition." There were four VietSpeech studies, and the current article presents data from VietSpeech Study 2. In VietSpeech Study 1, a total of 271 adults with Vietnamese heritage living in Australia completed a questionnaire describing their linguistic multicompetence, language use, proficiency, and home language maintenance (McLeod et al., 2019). Three adult cluster profiles were statistically identified based on their language proficiency: Vietnamese proficient (31.3%), similar proficiency in Vietnamese and English (52.1%), and English proficient (16.6%; Wang et al., 2021). Child-, family-, and community-related variables associated with home language maintenance were identified (Tran, McLeod, et al., 2021; Tran, Wang, et al., 2021), and Vietnamese–Australian family language policies/rules were described (Tran, Verdon, & McLeod, 2022; Tran, Verdon, McLeod, & Wang, 2022). In VietSpeech Study 2 (the current article), bilingual speech acquisition was described for 154 Australian Vietnamese-English–speaking children and adults. The VietSpeech Multilingual Assessment Protocol (see the Appendix) was described and used in the current article, and the VietSpeech Multilingual Transcription Protocol was described by Margetson, McLeod, Verdon, and Tran (2023). Three case studies have been written about children from Study 2 that examine the impact of cross-linguistic transfer, ambient phonology, and developmental maturation: (a) a three-generation case study (McLeod, Margetson, et al., 2022), (b) a longitudinal case study (Margetson et al., 2023a), and (c) a comparison between two pairs of siblings with different language proficiencies (Margetson et al., 2023b). In VietSpeech Study 3, the expertise of professionals across the world was sourced to develop an evidence-based Vietnamese–English speech program to support children's speech competence and language maintenance (Verdon, McLeod, et al., 2021). In VietSpeech Study 4, the feasibility and efficacy of SuperSpeech, an online Vietnamese–English speech group program were evaluated, and there was some evidence that the program supported children's Vietnamese and English speech and language maintenance (McLeod, Verdon, et al., 2022; Tran, Verdon, McLeod, & Wang, 2022; Verdon, Tran, et al., 2021).

Aims

The overarching aim of the current article based on data from VietSpeech Study 2 was to describe Vietnamese–Australian children's and adults' Vietnamese and Australian English speech production through the application of the emergence approach to speech acquisition (Davis & Bedore, 2013), the theory of linguistic multicompetence (Cook, 2016), and the Culturally Responsive Teamwork Framework (Hopf et al., 2021). The aim was to identify the impact of cross-linguistic, dialectal, maturational, language experience and environmental (ambient phonology) factors on children's speech acquisition to support SLPs' differential diagnosis of SSD in multilingual children. Specific aims were as follows:

1. to identify the impact of bilingual language experience on speech production by comparing statistically identified cluster profiles of children and adults;
2. to compare children's intelligibility for Vietnamese and Australian English;
3. to compare children's production of Vietnamese consonants, semivowels, vowels, and tones and Australian English consonants and vowels and determine the impact of Standard and dialectal definitions of accuracy;
4. to identify consonant realizations and phonological patterns used by children as they acquire Vietnamese and Australian English;
5. to compare adult family members' production of Vietnamese consonants, semivowels, vowels, and tones and Australian English consonants and vowels and determine the impact of Standard and dialectal definitions of accuracy; and
6. to compare children's and adult family members' accuracy of consonants, vowels, and tones and identify the frequency children and family members' productions matched.

Method

Ethical Approval

Ethical approval was received from the Charles Sturt University Human Research Ethics Committee (Approval No. H18084). The adult participants provided written consent to participate after receiving information about the study in Vietnamese and English. Child participants' parent(s) provided written consent for their children to participate, and then each child was asked to provide assent after hearing a child-friendly description of the research in their preferred language (either English or Vietnamese).

Participants

Total Sample

VietSpeech Study 2 had a total sample of 154 participants within 53 families: 69 children and 85 adults who lived in the same households as the children. To be included in the child sample, participants had to have Vietnamese heritage, live in Australia, be 2–8 years old (so they were acquiring speech), have an adult family member who was willing to participate, and consent to be involved. All of the participants resided in Sydney, Australia, at the time of data collection (the city with the highest concentration of people who speak Vietnamese; Australian Bureau of Statistics, 2016). Originally, the sample size was intended to be larger with a narrower age range; however, participant recruitment was impacted due to strict COVID-19 lockdown restrictions in Australia.

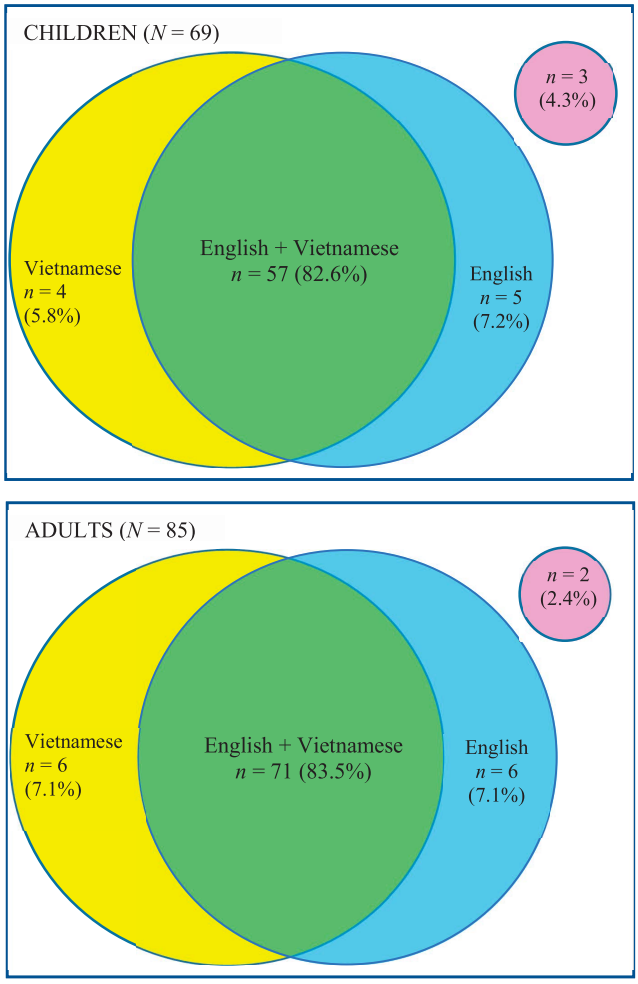
Child and Adult Participants

Each of the 154 participants was invited to complete a single-word speech assessment in Vietnamese and English; however, not everyone completed all of the tasks due to lack of competence with Vietnamese or English, time constraints, or technology issues (rare). The overlap between the child and adult participants who completed the Vietnamese and English speech assessments is illustrated in Figure 1. Therefore, the participants' demographic data are described for the Vietnamese (V) and English (E) participant samples.

Child Participants

Of the 69 children, 61 completed the Vietnamese Speech Assessment (VSA; Phạm, Le, & McLeod, 2016) and 62 completed the English speech assessment (Diagnostic Evaluation of Articulation and Phonology [DEAP];

Figure 1. Derivation of participants for VietSpeech Study 2. The total sample contained 69 children and 85 adults, the Vietnamese speech sample contained 61 children and 77 adults, the English speech sample contained 62 children and 77 adults, and the outlying circles represent three children and two adults who were assessed but were unable to complete either the Vietnamese or English speech assessments.



Dodd et al., 2002[1]; see Figure 1 and Table 2). There were a similar number of male (V 49.18%, E 45.16%) and female (V 50.82%, E 54.84%) participants. The Vietnamese sample ranged in age from 2;2 to 8;10 (years;months; M = 68.00 months, SD = 21.09, range: 26–106), and the English sample ranged in age from 2;0 to 8;10 (M = 68.26 months, SD = 20.85, range: 24–106). Most attended school (V 75.00%, E 74.19%); the others were too young. The participants lived in a range of suburbs from the most disadvantaged (1st decile) to most advantaged (10th decile) according to the Index of Relative Socioeconomic Advantage and Disadvantage (IRSAD; Australian Bureau of Statistics, 2018). The decile number is based on individual postcodes' ranking within Australia. The mean IRSAD decile of the 53 families was 6.72 (SD = 3.27), the median was 8, and the mode was 9 (25th percentile = 3, 50th percentile = 8, 75th percentile = 9). Most child participants had lived in Australia throughout their lives; however, some had lived in Vietnam (V 21.31%, E 19.35%) or another country (V 16.39%, E 16.13%) for more than 1 year.

Both Vietnamese and English were spoken in the children's households. Most of the children were reported to speak Vietnamese (V 85.25%, E 82.26%) and English (V 78.89%, E 82.26%) and hear Vietnamese (V 95.08%, E 93.55%) and English (V 88.52%, E 90.32%). The children's speech and language skills were described by the parents using the Vietnamese translation of the Inventory to Assess Language Knowledge (ITALK; Peña et al., 2018; see Table 2). The child participants' scores were higher for English than Vietnamese for vocabulary, speech, sentence length, grammar, and receptive language (see Table 2). Children were included in the study regardless of their developmental status (cf. Peña et al., 2006); however, most of the children had typical developmental status according to parents' responses to questions from the Parents' Evaluation of Developmental Status (PEDS; Glascoe, 2000). Some parents had concerns about their children's talking and making speech sounds (V 27.87%, E 29.03%) and understanding (V 11.48%, E 12.90%), and some reported that their child had received an assessment/intervention from an SLP (V 22.95%, E 25.81%; see Table 2). The majority were identified by their parents as having no "developmental delay" (V 93.33%, E 93.44%), no concerns about hearing (V 96.72%, E 96.77%), not having frequent ear infections (V 96.67%, E 99.38%), and/or no ongoing medical problems (V 95.08%, E 95.16%; valid number of responses for each question vary). The developmental status of the sample was similar to population-level studies of Australian children using similar questions (cf. McLeod & Harrison, 2009).

Adult Participants

Every child participant had at least one adult in their family who also completed the English speech assessment (see Table 3). Of the 85 adults in the total sample, there were 71 adult participants who completed both the Vietnamese and English speech assessments, six who only completed the VSA, and another six who only completed the English sample (see Figure 1). There were more women (V 74.03%, E 70.13%) than men (V 25.97%, E 29.87%). Most were the child participants' mothers (V 61.04%, E 62.34%), but others were their fathers (V 19.48%, E 22.08%), grandmothers (V 11.69%, E 5.19%), grandfathers (V 3.90%, E 3.90%), aunts (V 2.60%, E 2.60%), uncles (V 2.60%, E 2.60%), or adult siblings (V 1.30%, E 1.30%). The adults ranged in age from 22 to 80 years (M = 42.38 years, SD = 11.74) in the Vietnamese sample and from 22 to 93 years (M = 44.38 years, SD = 13.38) in the English sample (see Table 3). Most were born in Vietnam (V 92.21%, E 85.71%) and described themselves as first-generation migrants (V 92.11%, E 89.47%). Most spoke Vietnamese as their first language (V 97.40%, E 89.61%). They reported speaking a Northern dialect (63.64%), a Central dialect (5.19%), a Southern dialect (29.87%), or were not sure; none of the adult participants reported that they spoke Standard Vietnamese (see Table 3). While none of the adults in the Vietnamese sample spoke English as their first language, 5.19% in the English sample spoke English as their first language. A few spoke another first language (V 2.60%, E 5.19%; i.e., Cantonese, Thai, and Teo Chew). Socioeconomic status (SES) indicators included adult participants' education and occupation; many participants had either a postgraduate degree/diploma (V 37.66%, E 36.36%) or bachelor's degree (V 32.47%, E 32.47%), and many had a professional job (V 32.47%, E 31.17%; see Table 3). Their SES profile of a mean IRSAD decile of 6.72 was similar to the profile of the 271 Vietnamese–Australian participants from VietSpeech Study 1 who completed an online questionnaire (McLeod et al., 2019) and had a mean IRSAD decile of 6.52, even though different participant groups were sampled.

Instruments

Speech Assessments

Vietnamese Speech Assessment. The VSA (Phạm, Le, & McLeod, 2016) is a picture-naming task comprising 77 single words. Each Vietnamese consonant, semivowel, vowel,

[1]There were 67 children who commenced the English speech assessment and 62 who were in the final sample. Three were unable to complete it, and two were excluded because they were identified as outliers due to extremely low percentage of standard consonants correct (PCC-S = 23.67% and 38.16%).

Table 2. Demographic and communication information for child participants in the total sample (N = 69), Vietnamese speech sample (n = 61), and English speech sample (n = 62).

Variable		Child total sample (N = 69)					Child Vietnamese speech sample (n = 61)					Child English speech sample (n = 62)				
M (*SD*) *n* (%)	Total	High Proficiency (n = 26)	Low Proficiency (n = 21)	English Proficient (n = 22)	Total	Valid	High Proficiency (n = 26)	Low Proficiency (n = 14)	English Proficient (n = 21)	Total	Valid	High Proficiency (n = 24)	Low Proficiency (n = 16)	English Proficient (n = 22)	Total	Valid
Age (months)		71.19 (19.18)	53.29 (20.62)	69.68 (22.33)	65.26 (21.87)	69	71.19 (19.18)	56.29 (22.72)	71.86 (20.36)	68.00 (21.09)	61	73.96 (16.84)	57.75 (21.48)	69.68 (22.33)	68.26 (20.85)	62
Age (range)		26–101	30–96	24–106	24–106		26–101	30–96	35–106	26–106		43–101	34–96	24–106	24–106	
Sex	Male	14 (53.85%)	10 (47.62%)	9 (40.91%)	33 (47.83%)	69	14 (53.85%)	7 (50.00%)	9 (42.86%)	30 (49.18%)	61	12 (50.00%)	7 (43.75%)	9 (40.91%)	28 (45.16%)	62
	Female	12 (46.15%)	11 (52.38%)	13 (59.09%)	36 (52.17%)		12 (46.15%)	7 (50.00%)	12 (57.14%)	31 (50.82%)		12 (50.00%)	9 (53.85%)	13 (59.09%)	34 (54.84%)	
Attend school	Yes	20 (76.92%)	9 (45.00%)	17 (77.27%)	46 (67.65%)	68	20 (76.92%)	8 (61.54%)	17 (80.95%)	45 (75.00%)	60	20 (83.33%)	9 (56.25%)	17 (77.27%)	46 (74.19%)	62
	No	6 (23.08%)	11 (55.00%)	5 (22.73%)	22 (32.35%)		6 (23.08%)	5 (38.46%)	4 (19.05%)	15 (25.00%)		4 (16.67%)	7 (43.75%)	5 (22.73%)	16 (25.81%)	
Lived in Vietnam for more than 1 year	Yes	4 (15.38%)	6 (28.57%)	3 (13.64%)	13 (18.84%)	69	4 (15.38%)	6 (42.86%)	3 (14.29%)	13 (21.31%)	61	4 (16.67%)	5 (31.25%)	3 (13.64%)	12 (19.35%)	62
	No	22 (84.62%)	15 (71.43%)	19 (86.36%)	56 (81.16%)		22 (84.62%)	8 (57.14%)	18 (85.71%)	48 (78.69%)		20 (83.33%)	11 (68.75%)	19 (86.36%)	50 (80.65%)	
Lived in another country for more than 1 year	Yes	3 (11.54%)	5 (23.81%)	2 (9.09%)	10 (14.49%)	69	3 (11.54%)	5 (35.71%)	2 (9.52%)	10 (16.39%)	61	3 (12.50%)	5 (31.25%)	2 (9.09%)	10 (16.13%)	62
	No	23 (88.46%)	16 (76.19%)	20 (90.91%)	59 (85.51%)		23 (88.46%)	9 (64.29%)	19 (90.48%)	51 (83.61%)		21 (87.50%)	11 (68.75%)	20 (90.91%)	52 (83.87%)	
Speak English at home	Yes	16 (61.54%)	16 (76.19%)	22 (100.00%)	54 (78.26%)	69	16 (61.54%)	11 (78.57%)	21 (100.00%)	48 (78.69%)	61	16 (66.67%)	13 (81.25%)	22 (100.00%)	51 (82.26%)	62
	No	10 (38.46%)	5 (23.81%)	0 (0.00%)	15 (21.74%)		10 (38.46%)	3 (21.43%)	0 (0.00%)	13 (21.31%)		8 (33.33%)	3 (18.75%)	0 (0.00%)	11 (17.74%)	
Hear English at home	Yes	22 (84.62%)	18 (85.71%)	22 (100.00%)	62 (89.86%)	69	22 (84.62%)	13 (81.25%)	22 (100.00%)	54 (88.52%)	61	21 (87.50%)	13 (81.25%)	22 (100.00%)	56 (90.32%)	62
	No	4 (15.38%)	3 (14.29%)	0 (0.00%)	7 (10.14%)		4 (15.38%)	3 (18.75%)	0 (0.00%)	7 (11.48%)		3 (12.50%)	3 (18.75%)	0 (0.00%)	6 (9.68%)	
Speak Vietnamese at home	Yes	26 (100.00%)	13 (61.90%)	16 (72.73%)	55 (79.71%)	69	26 (100.00%)	11 (68.75%)	15 (71.43%)	52 (85.25%)	61	24 (100.00%)	11 (68.75%)	16 (72.73%)	51 (82.26%)	62
	No	0 (0.00%)	8 (38.10%)	6 (27.27%)	14 (20.29%)		0 (0.00%)	3 (21.43%)	6 (28.57%)	9 (14.75%)		0 (0.00%)	5 (31.25%)	6 (27.27%)	11 (17.74%)	
Hear Vietnamese at home	Yes	26 (100.00%)	19 (90.48%)	20 (90.91%)	65 (94.20%)	69	26 (100.00%)	13 (92.86%)	19 (90.48%)	58 (95.08%)	61	24 (100.00%)	14 (87.50%)	20 (90.91%)	58 (93.55%)	62
	No	0 (0.00%)	2 (9.52%)	2 (9.09%)	4 (5.80%)		0 (0.00%)	1 (7.14%)	2 (9.52%)	3 (4.92%)		0 (0.00%)	2 (12.50%)	2 (9.09%)	4 (6.45%)	
Concerns about talking and speech sounds[a]	Yes	5 (19.23%)	11 (52.38%)	6 (27.27%)	22 (31.88%)	69	5 (19.23%)	6 (42.86%)	6 (28.57%)	17 (27.87%)	61	4 (16.67%)	8 (50.00%)	6 (27.27%)	18 (29.03%)	62
	A little	3 (11.54%)	5 (23.81%)	5 (22.73%)	13 (18.84%)		3 (11.54%)	5 (35.71%)	5 (23.81%)	13 (21.31%)		3 (12.50%)	4 (25.00%)	5 (22.73%)	12 (19.35%)	
	No	18 (69.23%)	5 (23.81%)	11 (50.00%)	34 (49.28%)		18 (69.23%)	3 (21.43%)	10 (47.62%)	31 (50.82%)		17 (70.83%)	4 (25.00%)	11 (50.00%)	32 (51.61%)	
Concerns about understanding what you say[a]	Yes	2 (7.69%)	5 (23.81%)	3 (13.64%)	10 (14.49%)	69	2 (7.69%)	2 (14.29%)	3 (14.29%)	7 (11.48%)	61	1 (4.17%)	4 (25.00%)	3 (13.64%)	8 (12.90%)	62
	A little	4 (15.38%)	5 (23.81%)	2 (9.09%)	11 (15.94%)		4 (15.38%)	3 (21.43%)	2 (9.52%)	9 (14.75%)		4 (16.67%)	3 (18.75%)	2 (9.09%)	9 (14.52%)	
	No	20 (76.92%)	11 (52.38%)	17 (77.27%)	48 (69.57%)		20 (76.92%)	9 (64.29%)	16 (76.19%)	45 (73.77%)		19 (79.17%)	9 (56.25%)	17 (77.27%)	45 (72.58%)	

(*table continues*)

Table 2. *(Continued).*

Variable		Child total sample (*N* = 69)					Child Vietnamese speech sample (*n* = 61)					Child English speech sample (*n* = 62)				
M (*SD*) *n* (%)	Total	High Proficiency (*n* = 26)	Low Proficiency (*n* = 21)	English Proficient (*n* = 22)	Total	Valid	High Proficiency (*n* = 26)	Low Proficiency (*n* = 14)	English Proficient (*n* = 21)	Total	Valid	High Proficiency (*n* = 24)	Low Proficiency (*n* = 16)	English Proficient (*n* = 22)	Total	Valid
Vocabulary (English)[b]	/5	3.50 (1.14)	2.90 (0.89)	4.36 (0.73)	3.60 (1.10)	67	3.50 (1.14)	2.86 (0.95)	4.33 (0.73)	3.64 (1.11)	59	3.50 (1.14)	2.88 (0.96)	4.36 (0.73)	3.65 (1.12)	62
Vocabulary (Vietnamese)[b]	/5	3.69 (0.97)	2.88 (1.59)	1.90 (0.94)	2.89 (1.37)	63	3.69 (0.97)	3.14 (1.51)	1.95 (0.94)	2.98 (1.33)	60	3.67 (1.01)	3.00 (1.63)	1.90 (0.94)	2.88 (1.38)	58
Speech (English)[b]	/5	4.10 (1.09)	3.19 (0.98)	4.23 (1.02)	3.84 (1.12)	64	4.10 (1.09)	3.14 (0.86)	4.19 (1.03)	3.89 (1.09)	56	4.10 (1.09)	3.44 (0.81)	4.23 (1.02)	3.97 (1.03)	59
Speech (Vietnamese)[b]	/5	3.92 (1.35)	3.35 (1.32)	3.30 (0.80)	3.57 (1.21)	63	3.92 (1.35)	3.57 (1.28)	3.32 (0.83)	3.64 (1.20)	59	4.08 (1.25)	3.61 (1.04)	3.30 (0.80)	3.70 (1.10)	57
Sentence length (English)[b]	/5	4.64 (0.73)	3.52 (1.33)	4.77 (0.53)	4.32 (1.06)	65	4.64 (0.73)	3.71 (1.14)	4.86 (0.36)	4.49 (0.87)	57	4.64 (0.73)	3.81 (1.17)	4.77 (0.53)	4.47 (0.89)	60
Sentence length (Vietnamese)[b]	/5	4.40 (0.91)	3.07 (1.79)	2.35 (1.31)	3.38 (1.57)	60	4.40 (0.91)	3.58 (1.62)	2.42 (1.30)	3.55 (1.49)	56	4.42 (0.93)	3.25 (1.71)	2.35 (1.31)	3.43 (1.55)	56
Grammar (English)[b]	/5	3.95 (0.86)	2.90 (1.07)	3.73 (0.88)	3.54 (1.03)	63	3.95 (0.86)	3.15 (0.99)	3.76 (0.89)	3.69 (0.94)	55	3.95 (0.86)	3.07 (0.88)	3.73 (0.88)	3.64 (0.93)	58
Grammar (Vietnamese)[b]	/5	3.46 (1.24)	3.07 (1.22)	2.84 (1.21)	3.17 (1.24)	60	3.46 (1.24)	3.42 (1.08)	2.89 (1.23)	3.27 (1.21)	56	3.54 (1.18)	3.25 (1.14)	2.84 (1.21)	3.24 (1.20)	55
Receptive language (English)[b]	/5	4.46 (0.78)	3.38 (0.97)	4.41 (1.05)	4.10 (1.05)	67	4.46 (0.78)	3.36 (0.93)	4.38 (1.07)	4.17 (1.02)	59	4.46 (0.78)	3.44 (0.89)	4.41 (1.05)	4.18 (1.00)	62
Receptive language (Vietnamese)[b]	/5	3.65 (1.20)	3.20 (1.37)	3.57 (0.93)	3.52 (1.16)	62	3.65 (1.20)	3.50 (1.31)	3.55 (0.94)	3.59 (1.12)	58	3.88 (0.95)	3.25 (1.29)	3.57 (0.93)	3.63 (1.03)	57

Note. Child Cluster 1: High Proficiency = children's reported ability to speak and understand both Vietnamese and English was high; Child Cluster 2: Low Proficiency = children's reported ability to speak and understand both Vietnamese and English was low; Child Cluster 3: English Proficient = children's reported ability to speak and understand English was higher (well–very well) than their ability to speak and understand Vietnamese.

[a]Parents' Evaluation of Developmental Status (Glascoe, 2000). [b]Inventory to Assess Language Knowledge (Peña et al., 2018).

Table 3. Demographic information for adult participants in the total sample (N = 85), Vietnamese speech sample (n = 77), and English speech sample (n = 77).

Variable		Adult total sample (N = 85)					Adult Vietnamese speech sample (n = 77)					Adult English speech sample (n = 77)				
M (*SD*) *n* (%)		Vietnamese Proficient (n = 40)	Similar Proficiency (n = 36)	English Proficient (n = 9)	Total	Valid	Vietnamese Proficient (n = 39)	Similar Proficiency (n = 36)	English Proficient (n = 2)	Total	Valid	Vietnamese Proficient (n = 33)	Similar Proficiency (n = 36)	English Proficient (n = 8)	Total	Valid
Age (years)		47.10 (13.91)	37.53 (5.99)	51.00 (19.86)	43.42 (13.02)	84	47.16 (14.09)	37.53 (5.99)	39.00 (0.00)	42.38 (11.74)	76	44.38 (13.38)	37.53 (5.99)	52.13 (20.92)	41.94 (12.43)	76
Sex	Male	8 (20.00%)	12 (33.33%)	5 (55.56%)	25 (29.41%)	85	7 (17.95%)	12 (33.33%)	1 (50.00%)	20 (25.97%)	77	7 (21.21%)	12 (33.33%)	4 (50.00%)	23 (29.87%)	77
	Female	32 (80.00%)	24 (66.67%)	4 (44.44%)	60 (70.59%)		32 (82.05%)	24 (66.67%)	1 (50.00%)	57 (74.03%)		26 (78.79%)	24 (66.67%)	4 (50.00%)	54 (70.13%)	
Place of birth	Australia	1 (2.50%)	1 (2.78%)	4 (44.44%)	6 (7.06%)	85	1 (2.56%)	1 (2.78%)	1 (50.00%)	3 (3.90%)	77	1 (3.03%)	1 (2.78%)	4 (50.00%)	6 (7.79%)	77
	Vietnam	38 (95.00%)	33 (91.67%)	3 (33.33%)	74 (87.06%)		37 (94.87%)	33 (91.67%)	1 (50.00%)	71 (92.21%)		31 (93.94%)	33 (91.67%)	2 (25.00%)	66 (85.71%)	
	Other	1 (2.50%)	2 (5.56%)	2 (22.22%)	5 (5.88%)		1 (2.56%)	2 (5.56%)	0 (0.00%)	3 (3.90%)		1 (3.03%)	2 (5.56%)	2 (25.00%)	5 (6.49%)	
Years in Australia		11.23 (8.19)	12.57 (6.66)	42.22 (27.18)	15.08 (14.44)	85	10.95 (8.11)	12.57 (6.66)	34.00 (7.07)	12.31 (8.21)	77	11.79 (8.09)	12.57 (6.66)	43.50 (28.77)	15.45 (14.73)	77
First/main language	English	0 (0.00%)	0 (0.00%)	4 (44.44%)	4 (4.71%)	85	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	77	0 (0.00%)	0 (0.00%)	4 (50.00%)	4 (5.19%)	77
	Vietnamese	40 (100.00%)	35 (97.22%)	2 (22.22%)	77 (90.59%)		39 (100.00%)	35 (97.22%)	1 (50.00%)	75 (97.40%)		33 (100.00%)	35 (97.22%)	1 (12.50%)	69 (89.61%)	
	Other	0 (0.00%)	1 (2.78%)	3 (33.33%)	4 (4.71%)		0 (0.00%)	1 (2.78%)	1 (50.00%)	2 (2.60%)		0 (0.00%)	1 (2.78%)	3 (37.50%)	4 (5.19%)	
Vietnamese dialect	Standard	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	85	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	77	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	77
	Northern	24 (60.00%)	25 (69.44%)	0 (0.00%)	49 (57.65%)		24 (61.54%)	25 (69.44%)	0 (0.00%)	49 (63.64%)		18 (54.55%)	25 (69.44%)	0 (0.00%)	43 (55.84%)	
	Central	3 (7.50%)	1 (2.78%)	0 (0.00%)	4 (4.71%)		3 (7.69%)	1 (2.78%)	0 (0.00%)	4 (5.19%)		3 (9.09%)	1 (2.78%)	0 (0.00%)	4 (5.19%)	
	Southern	12 (30.00%)	10 (27.78%)	3 (33.33%)	25 (29.41%)		11 (28.21%)	10 (27.78%)	2 (100.00%)	23 (29.87%)		11 (33.33%)	10 (27.78%)	2 (25.00%)	23 (29.87%)	
	Not sure	1 (2.50%)	0 (0.00%)	0 (0.00%)	1 (1.18%)		1 (2.56%)	0 (0.00%)	0 (0.00%)	1 (1.30%)		1 (3.03%)	0 (0.00%)	0 (0.00%)	1 (1.30%)	
	Not applicable	0 (0.00%)	0 (0.00%)	6 (66.67%)	6 (7.06%)		0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)		0 (0.00%)	0 (0.00%)	6 (75.00%)	6 (7.79%)	
Generation of immigration	First	38 (97.44%)	33 (91.67%)	2 (25.00%)	73 (87.95%)	83	37 (97.37%)	33 (91.67%)	0 (0.00%)	70 (92.11%)	76	33 (100.00%)	33 (91.67%)	2 (28.57%)	68 (89.47%)	76
	1.5	0 (0.00%)	1 (2.78%)	2 (25.00%)	3 (3.61%)		0 (0.00%)	1 (2.78%)	1 (50.00%)	2 (2.63%)		0 (0.00%)	1 (2.78%)	1 (14.29%)	2 (2.63%)	
	Second	0 (0.00%)	1 (2.78%)	2 (25.00%)	3 (3.61%)		0 (0.00%)	1 (2.78%)	1 (50.00%)	2 (2.63%)		0 (0.00%)	1 (2.78%)	2 (28.57%)	3 (3.95%)	
	Third	0 (0.00%)	0 (0.00%)	1 (12.50%)	1 (1.20%)		0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)		0 (0.00%)	0 (0.00%)	1 (14.29%)	1 (1.32%)	
	Other	1 (2.56%)	1 (2.78%)	1 (12.50%)	3 (3.61%)		1 (2.63%)	1 (2.78%)	0 (0.00%)	2 (2.63%)		0 (0.00%)	1 (2.78%)	1 (14.29%)	2 (2.63%)	
Education (highest level)	High school/ diploma	18 (45.00%)	6 (16.67%)	2 (22.22%)	26 (30.59%)	85	17 (43.59%)	6 (16.67%)	0 (0.00%)	23 (29.87%)	77	13 (39.39%)	6 (16.67%)	2 (25.00%)	21 (27.27%)	77
	Bachelor's degree	16 (40.00%)	8 (22.22%)	3 (33.33%)	27 (31.76%)		16 (41.03%)	8 (22.22%)	1 (50.00%)	25 (32.47%)		14 (42.42%)	8 (22.22%)	3 (37.50%)	25 (32.47%)	
	Graduate diploma/ postgraduate degree	6 (15.00%)	22 (61.11%)	4 (44.44%)	32 (37.65%)		6 (15.38%)	22 (61.11%)	1 (50.00%)	29 (37.66%)		6 (18.18%)	22 (61.11%)	3 (37.50%)	31 (40.26%)	

(*table continues*)

Table 3. *(Continued).*

Variable		Adult total sample (*N* = 85)					Adult Vietnamese speech sample (*n* = 77)					Adult English speech sample (*n* = 77)				
M (*SD*) *n* (%)		Vietnamese Proficient (*n* = 40)	Similar Proficiency (*n* = 36)	English Proficient (*n* = 9)	Total	Valid	Vietnamese Proficient (*n* = 39)	Similar Proficiency (*n* = 36)	English Proficient (*n* = 2)	Total	Valid	Vietnamese Proficient (*n* = 33)	Similar Proficiency (*n* = 36)	English Proficient (*n* = 8)	Total	Valid
Occupation[a]	Manager	3 (7.50%)	4 (11.11%)	2 (22.22%)	9 (10.59%)	85	3 (7.69%)	4 (11.11%)	1 (50.00%)	8 (10.39%)		2 (6.06%)	4 (11.11%)	2 (25.00%)	8 (10.39%)	77
	Professional	6 (15.00%)	18 (50.00%)	1 (11.11%)	25 (29.41%)		6 (15.38%)	18 (50.00%)	1 (50.00%)	25 (32.47%)		5 (15.15%)	18 (50.00%)	1 (12.50%)	24 (31.17%)	
	Technician or trade	2 (5.00%)	3 (8.33%)	1 (11.11%)	6 (7.06%)		2 (5.13%)	3 (8.33%)	0 (0.00%)	5 (6.49%)		1 (3.03%)	3 (8.33%)	1 (12.50%)	5 (6.49%)	
	Community and personal service	4 (10.00%)	2 (5.56%)	1 (11.11%)	7 (8.24%)		4 (10.26%)	2 (5.56%)	0 (0.00%)	6 (7.79%)		4 (12.12%)	2 (5.56%)	1 (12.50%)	7 (9.09%)	
	Clerical/ administrative	4 (10.00%)	4 (11.11%)	0 (0.00%)	8 (9.41%)		4 (10.26%)	4 (11.11%)	0 (0.00%)	8 (10.39%)	77	4 (12.12%)	4 (11.11%)	0 (0.00%)	8 (10.39%)	
	Sales	1 (2.50%)	1 (2.78%)	0 (0.00%)	2 (2.35%)		1 (2.56%)	1 (2.78%)	0 (0.00%)	2 (2.60%)		1 (3.03%)	1 (2.78%)	0 (0.00%)	2 (2.60%)	
	Machinery	1 (2.50%)	1 (2.78%)	0 (0.00%)	2 (2.35%)		0 (0.00%)	1 (2.78%)	0 (0.00%)	1 (1.30%)		0 (0.00%)	1 (2.78%)	0 (0.00%)	1 (1.30%)	
	Laborer	6 (15.00%)	0 (0.00%)	0 (0.00%)	6 (7.06%)		6 (15.38%)	0 (0.00%)	0 (0.00%)	6 (7.79%)		5 (15.15%)	0 (0.00%)	0 (0.00%)	5 (6.49%)	
	Other	13 (32.50%)	3 (8.33%)	4 (44.44%)	20 (23.53%)		13 (33.33%)	3 (8.33%)	0 (0.00%)	16 (20.78%)		11 (33.33%)	3 (8.33%)	3 (37.50%)	17 (22.08%)	

Note. Adult Cluster 1: Vietnamese Proficient = adults' reported ability to speak and understand Vietnamese was higher than their ability to speak and understand English; Adult Cluster 2: Similar Proficiency = adults' reported ability to speak and understand Vietnamese and English was similarly high; Adult Cluster 3: English Proficient = adults' reported ability to speak and understand English was higher than their ability to speak and understand Vietnamese.

[a]Occupation classification is from the Australian Bureau of Statistics.

and tone is elicited in at least three stimulus items (except /p/ due to lack of suitable words familiar to children). The stimuli are presented in the order of the International Phonetic Alphabet (IPA) for the syllable-initial consonant and illustrated by colorful drawings. The score sheet includes acceptable productions for Standard, Northern, Central, and Southern Vietnamese using IPA transcription. The VSA underwent validation for test content, representativeness, and relevance of the test items (Phạm, McLeod, & Le, 2016) and has been validated with 195 monolingual Vietnamese children in Northern Vietnam (Phạm & McLeod, 2019) and 132 monolingual Vietnamese children in Southern Vietnam (Le et al., 2022).

English Speech Assessment. The Articulation and Phonology subtests of the DEAP (Dodd et al., 2002) were used to assess the children and adults' English. The target transcription on the Australian/U.K. score form was used as the definition of Standard Australian English for consonants. Two additional words were added (*go, chicken*) so that every word-initial English consonant was sampled twice; therefore, a total of 82 English words were elicited and analyzed. The DEAP (Dodd et al., 2002) was selected because the normative sample includes Australian children and it has been widely used in Australian English research and clinical practice (e.g., McLeod, Baker, et al., 2017).

Intelligibility. The children's intelligibility was assessed using the seven-item parent report Intelligibility in Context Scale (ICS; McLeod et al., 2012). The ICS was available for completion in English and Vietnamese. The ICS has been validated in over 14 languages within 18 studies of over 4,235 children (McLeod, 2020), including monolingual children in Vietnam speaking Northern Vietnamese (Phạm et al., 2017) and Southern Vietnamese (Le et al., 2022). The ICS was validated in Australia with 120 Australian English–speaking children (McLeod et al., 2012) and normed with 803 monolingual and bilingual Australian English–speaking children (McLeod et al., 2015). Mean ICS scores are not affected by the number of languages spoken by the children or whether the ICS was completed in English or in another language (McLeod, 2020; McLeod et al., 2015).

Parent Questionnaires

Parents completed questionnaires in either English or Vietnamese, addressing demographic information, children's development, factors influencing speech acquisition (e.g., age, sex, language profile, speech-language-hearing status, and disabilities), parental concern about their children's development, language proficiency, and family language policy. The questionnaires included questions created by the researchers after reviewing the literature (see McLeod et al., 2019; Tran, Wang, et al., 2021) and questions from previously published scales that were included after gaining permission from the authors. Previously published questionnaires included the PEDS (Glascoe, 2000) and the ITALK (Peña et al., 2018). To determine the participants' language proficiency, participants were asked to rate how well they/their child(ren) speak and understand English and Vietnamese in four separate questions on a 5-point scale (1 = *not at all*, 2 = *not well*, 3 = *average*, 4 = *well*, 5 = *very well*) using items adapted from the Australian census (Australian Bureau of Statistics, 2011) and the studies of Blake et al. (2018, 2021). A mean score based on their rating of their ability to speak and understand Vietnamese and English was calculated to represent their Vietnamese and English proficiency. A higher mean score indicated higher language proficiency. The mean score for the children's and adults' English and Vietnamese language proficiency was standardized and used to inform the statistical hierarchical cluster analysis of the participants, replicating analyses used in VietSpeech Study 1 (cf. Wang et al., 2021). Standardizing language proficiency variables help provide a clear picture of the relative strengths and weaknesses of participants' English and Vietnamese language proficiency within the participant sample (raw scores are found in Table 7 and standardized scores are found in Table 8 and Supplemental Material S3).

Procedure

Vietnamese–Australian families responded to social media posts and e-mails about the VietSpeech research and received information and consent forms written in either English or Vietnamese. Snowball sampling was encouraged. Adult family members completed the consent form giving permission to participate in the research and a demographic and language history questionnaire. One family member (typically the mother) provided consent for their child(ren) to participate and completed an additional questionnaire about their child. The questionnaires were available in English and Vietnamese; if they were completed in Vietnamese, the responses were translated into English by a bilingual VietSpeech team member.

The VietSpeech research team consisted of (a) proficient bilingual speakers in Vietnamese and English who were living in both countries (Vietnam and Australia), had a range of professions (SLP, linguist, and educator), and different migration generations (first and second generations since migration to Australia from Vietnam) as well as (b) speech-language pathology researchers who were predominantly monolingual speakers of English born in Australia who had visited and lived in Vietnam and worked with Vietnamese-speaking people as SLPs.

The child and adult participants were assessed in their homes by a VietSpeech team member who was proficient in Vietnamese and at least one other team member. Conversations undertaken in the participants' homes were

undertaken in the language preferred by each participant, usually bilingually in Vietnamese and English. Vietnamese and English speech production of the child(ren) and adult family member(s) was assessed using the VSA (Phạm, Le, & McLeod, 2016), the DEAP (Dodd et al., 2002), and the ICS (McLeod et al., 2012). Participants selected which assessment they started with (VSA or DEAP). Adults' assessments were completed after the child(ren)'s assessments to avoid the child(ren) hearing a model of the target words. If there was more than one child being tested in the household (e.g., siblings), the second child did not observe the first child's assessment.

During the speech assessments, the children were shown the stimulus books containing pictures, and the adults read the word list. A four-level cuing hierarchy script (see Table 4) was used with the children: (1) spontaneous "Đây là con…? [What's this]?", (2) cue "What has a trunk?", (3) binary choice "Is it an [elephant] or a [WORD]?", and (4) imitation "*Voi*. Con nhắc lại! [Say elephant]." During the assessment of Vietnamese, parents provided the imitated model of the word. If the children did not respond to Level 2 or 3 cues within the first several words, then a two-level cuing hierarchy script was used: (1) spontaneous and (4) imitation. On average, there were more spontaneous productions of English words (M = 77.75%, SD = 18.25) than Vietnamese words (M = 26.72%, SD = 23.67%; see Table 4). The number of Vietnamese words produced spontaneously differed according to cluster: High Proficiency cluster (M = 36.91%, SD = 25.86%), followed by the Low Proficiency cluster (M = 22.45%, SD = 19.55%) and then the English Proficient cluster (M = 16.95%, SD = 18.66%). The number of English words produced spontaneously also differed according to cluster: English Proficient cluster (M = 83.48%, SD = 12.76%), followed by the High Proficiency cluster (M = 80.34%, SD = 17.09%) and then the Low Proficiency cluster (M = 66.01%, SD = 21.76%; see Table 4).

The children's speech assessments took approximately 15 min to 1 hr, depending on their attention and the cueing hierarchy required. Assessments were audio-recorded using a Zoom H1 handy recorder and, if consent was given, video-recorded using a Zoom Q4n handy video recorder. Most families provided consent. Ambient noise was checked with an aim for 35 dBA or less.

Transcription

There were four members of the VietSpeech transcription team: an Australian English–speaking SLP (VS1), a bilingual Vietnamese- and English-speaking linguist (VS2), and two bilingual Vietnamese- and English-speaking SLPs (VS3, VS4). The transcription team undertook transcription training, reliability, and consensus checking informed by studies of Vietnamese (Masso et al., 2020; Phạm & McLeod, 2016, 2019) using the four-step VietSpeech Multilingual Transcription Protocol (Margetson, McLeod, Verdon, & Tran, 2023). The VSA score form provided target transcriptions for Standard, Northern, Central, and Southern Vietnamese dialects (Phạm, Le, & McLeod, 2016). Speech was transcribed using broad impressionistic transcription with the IPA and adding diacritics when a consonant was different from expected targets (e.g., the aspiration diacritic was added when an aspirated voiceless plosive was produced). After the assessment, transcriptions were independently checked by the transcribers using audio and video recordings. Typically, disagreements could be classified as "near functional equivalence" (Seifert et al., 2020, p. 84), acceptable dialectal variants (Phạm & McLeod, 2019), or common mismatches made by English-speaking SLPs transcribing Vietnamese speech (Masso et al., 2020; see Margetson, McLeod, Verdon, & Tran, 2023, for more details).

Reliability

Inter- and intrarater reliability for speech transcription of the children and adults' Vietnamese and English was calculated using a point-to-point agreement between 3–4 transcribers (see Table 5).

Vietnamese Transcription

Interrater reliability for Vietnamese transcription of the VSA was calculated for 10 (16.39%) child participants and 12 (15.58%) adult participants. For the 22 child and adult participants combined, the average interrater reliability for 12,056 Vietnamese consonants was 91.74% (range: 70.62–99.82), 1,732 Vietnamese vowels was 89.67% (range: 60.76–100.00), and 1,709 tones was 92.33% (range: 75.64–100.00; see Table 5). Intrarater reliability was undertaken by VS3 8 months after the original transcription for 10 child participants' productions of 1,370 Vietnamese consonants (M = 88.32%, range: 71.53–97.81), 790 Vietnamese vowels (M = 89.11%, range: 74.68–97.47), and 780 Vietnamese tones (M = 87.44%, range: 69.23–100.00) and 12 adult participants' productions of 1,644 Vietnamese consonants (M = 94.89%, range: 88.32–99.27), 948 Vietnamese vowels (M = 96.73%, range: 84.81–100.00), and 936 Vietnamese tones (M = 96.90%, range: 92.31–100.00). An acceptable to excellent level of inter- and intrarater reliability for speech transcription is demonstrated by these results (cf. Shriberg & Lof, 1991) and is similar to other studies of Vietnamese (Le et al., 2022; Phạm & McLeod, 2019).

English Transcription

Interrater reliability for English transcription of the DEAP was calculated for 10 (16.13%) child participants

Table 4. Cueing hierarchy used during children's Vietnamese assessment (n = 61) and English assessment (n = 62).

Vietnamese cue	Prompt	High Proficiency (n = 26)		Low Proficiency (n = 14)		English Proficient (n = 21)		Total	
		M	% (*SD*)	*M*	% (*SD*)	*M*	% (*SD*)	*M*	% (*SD*)
1. Spontaneous (Trả lời ngay)	Đây là con. . .?	28.42	36.91 (25.86)	17.29	22.45 (19.55)	13.05	16.95 (18.66)	20.57	26.72 (23.67)
2. Cue		0.73	0.95 (1.78)	0.50	6.49 (1.75)	0.48	6.18 (1.40)	0.59	0.77 (1.63)
3. Binary choice		0.00	0.00 (0.00)	0.36	0.46 (1.40)	0.00	0.00 (0.00)	0.08	0.11 (0.68)
4. Delayed imitation (Bắt chước gián đoạn)	*Voi*. Con nhắc lại!	47.15	61.24 (26.85)	56.50	73.38 (27.22)	63.19	82.07 (19.43)	54.82	71.19 (25.93)
0. No response		0.73	0.95 (4.08)	2.36	3.06 (10.38)	0.29	0.37 (1.17)	0.95	1.23 (5.64)
English cue	**Prompt**	**High Proficiency (n = 24)**		**Low Proficiency (n = 16)**		**English Proficient (n = 22)**		**Total**	
		M	% (*SD*)	*M*	% (*SD*)	*M*	% (*SD*)	*M*	% (*SD*)
1. Spontaneous	What's this?	65.88	80.34 (17.09)	54.13	66.01 (21.76)	68.45	83.48 (12.76)	63.76	77.75 (18.25)
2. Cue	What has a trunk?	4.33	5.28 (2.38)	2.69	3.28 (3.26)	3.64	4.43 (2.60)	3.66	4.46 (2.78)
3. Binary choice	It is an [*elephant*] or a [WORD]?	5.38	6.55 (8.02)	9.06	11.05 (7.32)	4.55	5.54 (5.24)	6.03	7.36 (7.20)
4. Delayed imitation	Say [*elephant*]	4.25	5.18 (10.55)	13.94	17.00 (18.28)	3.59	4.38 (7.80)	6.52	7.95 (13.20)
0. No response		0.08	0.10 (0.34)	0.69	0.84 (2.17)	0.32	0.39 (1.38)	0.32	0.39 (1.39)

Note. Prompt examples are for the word *voi* (elephant).

Table 5. Inter- and intrarater reliability for Vietnamese and English.

Language	Reliability	Focus	Group	Tokens	*M*	Range	Transcribers
Vietnamese	Interrater	Consonants and semivowels	10 children	5,480	86.57%	70.62–93.98	4 (VS1, VS2, VS3, VS4)[a]
			12 adults	6,576	96.05%	93.98–99.82	4 (VS1, VS2, VS3, VS4)[a]
			22 children + adults	12,056	91.74%	70.62–99.82	4 (VS1, VS2, VS3, VS4)[a]
		Vowels	10 children	785	84.33%	60.76–97.47	2 (VS2, VS3)
			12 adults	947	94.09%	86.08–100.00	2 (VS2, VS3)
			22 children + adults	1,732	89.67%	60.76–100.00	2 (VS2, VS3)
		Tones	10 children	774	87.73%	75.64–98.72	2 (VS2, VS3)
			12 adults	935	96.15%	90.91–100.00	2 (VS2, VS3)
			22 children + adults	1,709	92.33%	75.64–100.00	2 (VS2, VS3)
Vietnamese	Intrarater	Consonants and semivowels	10 children	1,370	88.32%	71.53–97.81	1 (VS3) 8 months apart
			12 adults	1,644	94.89%	88.32–99.27	1 (VS3) 8 months apart
		Vowels	10 children	790	89.11%	74.68–97.47	1 (VS3) 8 months apart
			12 adults	948	96.73%	84.81–100.00	1 (VS3) 8 months apart
		Tones	10 children	780	87.44%	69.23–100.00	1 (VS3) 8 months apart
			12 adults	936	96.90%	92.31–100.00	1 (VS3) 8 months apart
English	Interrater	Consonants	10 children	6,210	92.62%	73.75–99.03	3 (VS1, VS3, VS4)[b]
			12 adults	7,407	88.69%	81.48–93.72	3 (VS1, VS3, VS4)[b]
			22 children + adults	13,617	90.48%	73.75–99.03	3 (VS1, VS3, VS4)[b]
		Vowels	10 children	1,179	93.04%	81.36–98.31	2 (VS1, VS3)
			12 adults	1,408	83.24%	70.34–90.68	2 (VS1, VS3)
			22 children + adults	2,587	87.71%	70.34–98.31	2 (VS1, VS3)
English	Intrarater	Consonants	10 children	2,067	94.10%	82.61–100.00	1 (VS3) 8 months apart
			12 adults	2,469	88.70%	77.78–97.58	1 (VS3) 8 months apart
		Vowels	10 children	1,179	96.61%	92.37–99.15	1 (VS3) 8 months apart
			12 adults	1,408	84.45%	72.03–89.83	1 (VS3) 8 months apart

Note. VS1 is an Australian English–speaking SLP, VS2 is a bilingual Vietnamese-English–speaking linguist, and VS3 and VS4 are two bilingual Vietnamese-English–speaking SLPs. SLP = speech-language pathologist; VSA = Vietnamese Speech Assessment; DEAP = Diagnostic Evaluation of Articulation and Phonology.

[a]For seven of 10 child participants and 10 of 12 adult participants, VS1, VS2, VS3, and VS4 transcribed the VSA samples online (face-to-face) and checked transcriptions using audio and/or video recordings. For the remaining interrater reliability participants, two or three members of the transcription team transcribed online, and all four transcribed/checked the transcriptions using the recordings. [b]For seven of 10 child participants and 11 of 12 adult participants, three transcribers transcribed the DEAP samples online (face-to-face), and three transcribers checked transcriptions using audio and/or video recordings. For the remaining children, VS3 and VS4 transcribed the DEAP samples online, and three transcribers checked transcriptions using the recordings. For the remaining adult, VS1 transcribed online, and three transcribers checked the transcription from the recordings.

and 12 (15.58%) adult participants. For the 22 child and adult participants combined, the average interrater reliability for 13,617 English consonants was 90.48% (range: 73.75–99.03), and the average interrater reliability for 2,587 English vowels was 87.71% (range: 70.34–98.31; see Table 5). Intrarater reliability was undertaken by VS1 8 months after the original transcription for 10 child participants' productions of 2,067 English consonants

(M = 94.10%, range: 82.61–100.00) and 1,179 English vowels (M = 96.61%, range: 92.37–99.15) and 12 adult participants' productions of 2,469 English consonants (M = 88.70%, range: 77.78–97.58) and 1,408 English vowels (M = 84.45%, range: 72.03–89.83). An acceptable to excellent level of inter- and intrarater reliability for speech transcription is demonstrated by these results (cf. Shriberg & Lof, 1991) and is similar to other studies of English transcription.

Data Analysis

Cluster Analyses

To identify participants who shared similar characteristics in language proficiency, a cluster analysis was conducted similar to VietSpeech Study 1 where three cluster profiles were identified for 271 adults with Vietnamese heritage living in Australia: Vietnamese proficient (31%), similar proficiency (52%), and English proficient (17%; Wang et al., 2021). A cluster analysis is a person-centered analysis that identifies groups of participants who share similar response patterns to a set of predefined variables. Likert scale variables with five or more categories can be treated as continuous variables (Norman, 2010; Sullivan & Artino, 2013); therefore, Ward's hierarchical cluster method was appropriate for this analysis. Hierarchical cluster analysis using Ward's method was conducted in the SAS University Edition (SAS Institute) in order to group participants into relatively homogenous subgroups using standardized scores of Vietnamese and English language proficiencies (cf. Wang et al., 2021). A number of fit indices were consulted, including R^2, root-mean-square standard deviation, pseudo F, and pseudo t^2 statistics, to identify the optimal cluster solution (Milligan & Cooper, 1985).

Speech Analyses

The children's and adults' Vietnamese and Australian English consonant production were analyzed based on their productions of single words on the VSA and the DEAP. Within the VietSpeech Project, "correct" productions were described in different ways, acknowledging that judgments of correctness are an artifact of the frame of reference, which is influenced by social constructions of standardization, dialect spoken, and whether the speakers were monolingual or bilingual. The first definition of "correct" productions was written as PCC-S and related to Standard Vietnamese as provided on the score form of the VSA (Phạm, Le, & McLeod, 2016) and Standard Australian English as provided on the score form of the DEAP (Dodd et al., 2002). Standard Australian English consonants are similar to Standard British consonants and Standard U.S. consonants except that speakers in the United States use syllable-final "r" (e.g., *car* /ka/ in Australian and British English cf. /kaɹ/ in U.S. English; Crowe & McLeod, 2020). PCC-S does not take into account acceptable productions as a result of dialectal or cross-linguistic differences. Therefore, the second definition of "correct" was written as PCC-D and takes into account the dialect spoken by the participants (e.g., Southern, Central, and Northern Vietnamese). PCC-D was not reported for Australian English because there is limited dialectal difference in the monolingual production of Australian English consonants (Cox & Fletcher, 2017). However, an outcome of this research will be to define "correct" productions within the PCC-D framework for the Vietnamese–English dialect. Percentage of vowels correct–standard (PVC-S) and percentage of vowels correct–dialect (PVC-D) were reported for both Vietnamese and English because of dialectal variability for vowels, and Vietnamese tones were also considered using percentage of tones correct–standard (PTC-S) and percentage of tones correct–dialect (PTC-D).

By simultaneously assessing children and their family members, a greater understanding of what should be considered "correct" bilingual productions was determined (cf. McLeod, Margetson, et al., 2022). A contrastive analysis (McGregor et al., 1997) was used to compare children's speech productions with the speech of adults in their ambient environment to identify dialectal influences upon speech production. By integrating data regarding cross-linguistic, dialectal, maturational, language experience, and environmental (ambient phonology) factors, a list of late-acquired consonants were identified. This list helps identify which nonstandard productions have resulted from cross-linguistic transfer, with particular consideration of dialectal influences to differentiate acceptable productions from speech errors produced by children with SSD. A three-generation VietSpeech case study was used to pilot the protocols (McLeod, Margetson, et al., 2022).

Results

The results will be presented in corresponding order to the aims of the study.

Bilingual Language Experience: Participant Cluster Profiles

The multiple fit indices (see Table 6), together with considerations of theory and the meaningfulness of the cluster solutions, indicated that a three-cluster solution was the best solution for both child and adult participants. Table 7 provides the participants' scores across profiles for their Vietnamese and English language proficiency (to speak and understand) for children and adults, respectively, based on the total sample (69 children and 85

Table 6. Fit indices for cluster analysis of Vietnamese and English language proficiency with standardized scores for children (n = 69) and adults (n = 85) in the Vietnamese sample.

Cluster	Fit indices			
	R^2	RMSSTD	Pseudo F	Pseudo t^2
Child clusters				
2	.35	0.83	36.1	37.3
3	**.57**	**0.72**	**44.59**	**24.2**
4	.68	0.74	46.1	23.3
5	.75	0.59	48.7	27.3
Adult clusters				
2	.47	0.72	74	116
3	**.75**	**0.56**	**125**	**44.7**
4	.83	0.29	133	1038
5	.87	0.85	128	6.3

Note. R^2 measures variances between each cluster. RMSSTD is the aggregated value of the standard deviation for all variables included in each cluster. Pseudo F refers to the ratio of the between-clusters to within-cluster variance. Pseudo t^2 reports the difference between two merged clusters. A larger value of R^2 and pseudo F and a smaller value of RMSSTD and pseudo t^2 indicate a good cluster solution. Bolded text indicated the best cluster solution. RMSSTD = root-mean-square standard deviation.

adults), the Vietnamese sample (61 children and 77 adults), and the English sample (62 children and 77 adults). Three clusters emerged for the child participants (High Proficiency, Low Proficiency, and English Proficient), and three clusters emerged for the adult participants (Vietnamese Proficient, Similar Proficiency, and English Proficient). These cluster profiles were used to describe the results for the eligible child and adult participants for the Vietnamese data (N = 138) and English data (N = 139).

Cluster Profiles of the Entire Sample

The three statistically identified child participant cluster profiles (see Table 7) were as follows:

- *Child Cluster 1: High Proficiency* (n = 26, 37.68%). Children's reported ability to speak and understand both Vietnamese and English was high (*well–very well*).
- *Child Cluster 2: Low Proficiency* (n = 21, 30.43%). Children's reported ability to speak and understand both Vietnamese and English was low (*not well–average*).
- *Child Cluster 3: English Proficient* (n = 22, 31.88%). Children's reported ability to speak and understand English was higher (*well–very well*) than their ability to speak and understand Vietnamese (*not well–average*).

The children in the Low Proficiency cluster (M = 53.29 months, SD = 20.62) were younger than the children in the High Proficiency cluster (M = 71.19 months, SD = 19.18) and the English Proficient cluster (M = 69.68 months, SD = 20.33; see Table 7). Therefore, the Low Proficiency cluster was the youngest group, 76.19% of parents had concerns (*yes* or *a little*) about how they talked and made speech sounds, and many had not received services from an SLP. However, there was no significant interaction between age and parent concern within the Low Proficiency cluster, suggesting that parents rated proficiency with the age of their child in mind.

The three statistically identified adult participant cluster profiles (see Table 7) were as follows:

- *Adult Cluster 1: Vietnamese Proficient* (n = 40, 47.06%). Adults' reported ability to speak and understand Vietnamese (*well–very well*) was higher than their ability to speak and understand English (*not well–average*).
- *Adult Cluster 2: Similar Proficiency* (n = 36, 42.35%). Adults' reported ability to speak and understand Vietnamese and English was similarly high (*well–very well*).
- *Adult Cluster 3: English Proficient* (n = 9, 10.59%). Adults' reported ability to speak and understand English was higher (*very well*) than their ability to speak and understand Vietnamese (*not at all–not well*; see Table 7).

The three cluster profiles for the adult participants were similar to those identified by Wang et al. (2021) for 271 adults with Vietnamese heritage living in Australia.

Children's Intelligibility

The children's intelligibility in Vietnamese (see Table 8) and English (see Table 9) was described by their parents on the seven-item ICS (McLeod et al., 2012) where a score of 1 represented *never* intelligible and 5 represented *always* intelligible. The adults' intelligibility was not assessed. Overall, the children's intelligibility in English was rated more highly than the children's intelligibility in Vietnamese (ICS-VN). For Vietnamese, the total sample's average ICS score was 3.87 (SD = 0.65), indicating that, on average, all children were *sometimes* to *usually* intelligible across communication partners (see Table 8). The High Proficiency cluster had the highest average ICS-VN score (M = 4.31, SD = 0.65; *usually* to *always* intelligible), followed by the Low Proficiency cluster (M = 3.85, SD = 1.14) and the English Proficient cluster (M = 3.31, SD = 0.83; *sometimes* to *usually* intelligible). For English, the total sample's average ICS score was 4.24 (SD = 0.62), indicating that, on average, all children were *usually* to *always* intelligible across communication partners (see Table 9). The High Proficiency children had the highest average ICS score

Table 7. Vietnamese and English language proficiency profiles for the total sample of children (N = 69) and adults (N = 85).

Total sample		Children (N = 69)				Adults (N = 85)			
	M (*SD*)	High Proficiency (n = 26)	Low Proficiency (n = 21)	English Proficient (n = 22)	Total	Vietnamese Proficient (n = 40)	Similar Proficiency (n = 36)	English Proficient (n = 9)	Total
Vietnamese proficiency[a]	Speak	4.12 (0.71)	2.33 (1.20)	2.09 (0.68)	2.93 (1.28)	4.70 (0.46)	5.00 (0.00)	1.67 (0.87)	4.51 (1.08)
	Understand	4.19 (0.69)	2.76 (1.14)	2.82 (0.59)	3.32 (1.06)	4.70 (0.46)	4.97 (0.17)	1.67 (0.87)	4.49 (1.08)
	Read	2.08 (1.32)	1.70 (1.03)	1.50 (0.60)	1.78 (1.06)	4.70 (0.46)	4.81 (0.62)	1.44 (0.73)	4.40 (1.17)
	Write	1.88 (1.11)	1.35 (0.75)	1.27 (0.46)	1.53 (0.87)	4.70 (0.46)	4.75 (0.65)	1.22 (0.44)	4.35 (1.21)
English proficiency[a]	Speak	4.27 (0.53)	2.57 (0.75)	4.45 (0.51)	3.81 (1.02)	2.58 (0.64)	4.17 (0.45)	4.44 (0.88)	3.45 (1.02)
	Understand	4.38 (0.50)	2.90 (0.77)	4.50 (0.51)	3.97 (0.92)	2.70 (0.72)	4.22 (0.42)	4.44 (0.88)	3.53 (1.01)
	Read	3.23 (1.27)	2.05 (1.10)	3.14 (1.61)	2.85 (1.43)	2.68 (0.69)	4.17 (0.70)	4.44 (0.88)	3.49 (1.05)
	Write	3.11 (1.21)	1.65 (0.81)	3.00 (1.45)	2.65 (1.35)	2.55 (0.64)	3.94 (0.75)	4.33 (1.12)	3.33 (1.05)
Vietnamese sample		**Children (n = 61)**				**Adults (n = 77)**			
	M (*SD*)	High Proficiency (n = 26)	Low Proficiency (n = 14)	English Proficient (n = 21)	Total	Vietnamese Proficient (n = 39)	Similar Proficiency (n = 36)	English Proficient (n = 2)	Total
English proficiency[a]	Speak	4.27 (0.53)	2.79 (0.58)	4.43 (0.51)	3.98 (0.85)	2.56 (0.64)	4.17 (0.45)	5.00 (0.00)	3.38 (1.00)
	Understand	4.38 (0.50)	3.00 (0.39)	4.48 (0.51)	4.10 (0.77)	2.69 (0.73)	4.22 (0.42)	5.00 (0.00)	3.47 (0.99)
	Read	3.23 (1.27)	2.00 (0.91)	3.19 (1.63)	2.95 (1.42)	2.67 (0.70)	4.17 (0.70)	5.00 (0.00)	3.43 (1.04)
	Write	3.11 (1.21)	1.69 (0.85)	3.10 (1.41)	2.80 (1.33)	2.53 (0.64)	3.94 (0.75)	5.00 (0.00)	3.26 (1.02)
Vietnamese proficiency[a]	Speak	4.12 (0.71)	2.79 (1.12)	2.10 (0.70)	3.11 (1.21)	4.69 (0.47)	5.00 (0.00)	2.50 (0.71)	4.78 (0.53)
	Understand	4.19 (0.69)	3.14 (0.86)	2.81 (0.60)	3.48 (0.94)	4.69 (0.47)	4.97 (0.17)	2.50 (0.71)	4.77 (0.54)
	Read	2.08 (1.32)	1.77 (1.01)	1.52 (0.60)	1.82 (1.07)	4.69 (0.47)	4.81 (0.62)	1.50 (0.71)	4.66 (0.75)
	Write	1.88 (1.11)	1.46 (0.88)	1.29 (0.46)	1.58 (0.91)	4.69 (0.47)	4.75 (0.65)	1.50 (0.71)	4.64 (0.76)
English sample		**Children (n = 62)**				**Adults (n = 77)**			
	M (*SD*)	High Proficiency (n = 24)	Low Proficiency (n = 16)	English Proficient (n = 22)	Total	Vietnamese Proficient (n = 33)	Similar Proficiency (n = 36)	English Proficient (n = 8)	Total
English proficiency[a]	Speak	4.29 (0.55)	2.63 (0.72)	4.45 (0.51)	3.91 (0.96)	2.70 (0.47)	4.17 (0.45)	4.50 (0.93)	3.57 (0.92)
	Understand	4.42 (0.50)	2.94 (0.68)	4.50 (0.51)	4.06 (0.87)	2.85 (0.57)	4.22 (0.42)	4.50 (0.93)	3.66 (0.90)
	Read	3.38 (1.21)	2.13 (1.06)	3.14 (1.61)	2.98 (1.41)	2.82 (0.53)	4.17 (0.70)	4.50 (0.93)	3.62 (0.96)
	Write	3.25 (1.15)	1.73 (0.88)	3.00 (1.44)	2.79 (1.34)	2.67 (0.48)	3.94 (0.75)	4.38 (1.19)	3.44 (0.98)
Vietnamese proficiency[a]	Speak	4.13 (0.74)	2.44 (1.15)	2.09 (0.68)	2.97 (1.25)	4.70 (0.47)	5.00 (0.00)	1.50 (0.76)	4.51 (1.11)
	Understand	4.17 (0.70)	2.81 (1.05)	2.81 (0.59)	3.34 (1.01)	4.70 (0.47)	4.97 (0.17)	1.50 (0.76)	4.49 (1.11)
	Read	2.17 (1.34)	1.67 (0.98)	1.50 (0.60)	1.80 (1.06)	4.70 (0.47)	4.81 (0.62)	1.25 (0.46)	4.39 (1.20)
	Write	1.96 (1.12)	1.33 (0.82)	1.27 (0.46)	1.56 (0.90)	4.70 (0.47)	4.75 (0.65)	1.13 (0.35)	4.35 (1.23)

Note. Child Cluster 1: High Proficiency = children's reported ability to speak and understand both Vietnamese and English was high; Child Cluster 2: Low Proficiency = children's reported ability to speak and understand both Vietnamese and English was low; Child Cluster 3: English Proficient = children's reported ability to speak and understand English was higher (*well–very well*) than their ability to speak and understand Vietnamese; Adult Cluster 1: Vietnamese Proficient = adults' reported ability to speak and understand Vietnamese was higher than their ability to speak and understand English; Adult Cluster 2: Similar Proficiency = adults' reported ability to speak and understand Vietnamese and English was similarly high; Adult Cluster 3: English Proficient = adults' reported ability to speak and understand English was higher than their ability to speak and understand Vietnamese.

[a]1 = *not at all*; 2 = *not well*; 3 = *average*; 4 = *well*; 5 = *very well*.

Table 8. Vietnamese intelligibility, percentage of vowels correct, and percentage of tones correct for children (n = 61) and adults (n = 77) in the Vietnamese speech sample.

Speech element	Children (n = 61)								Adults (n = 77)							
	High Proficiency (n = 26)		Low Proficiency (n = 14)		English Proficient (n = 21)		Total		Vietnamese Proficient (n = 39)		Similar Proficiency (n = 36)		English Proficient (n = 2)		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Vietnamese intelligibility: ICS-VN (/5)[a]	4.31	0.65	3.85	1.14	3.31	0.83	3.87	0.93	—		—		—		—	
Standard Vietnamese vowels: PVC-S (/79)[b,c]	83.84%	6.83%	81.65%	8.41%	80.11%	7.63%	82.05%	7.54%	85.72%	6.58%	86.29%	6.17%	79.11%	0.90%	85.81%	6.36%
Northern vowels[d]	5.50%	4.44%	5.61%	4.57%	4.70%	4.20%	5.25%	4.33%	6.56%	4.73%	7.03%	4.61%	0.63%	0.90%	6.63%	4.69%
Central vowels[d]	0.00%	0.00%	0.00%	0.00%	0.42%	1.93%	0.15%	1.13%	0.45%	2.00%	0.32%	1.90%	1.90%	0.90%	0.43%	1.93%
Southern vowels[d]	8.42%	9.83%	7.59%	9.93%	9.34%	9.67%	8.55%	9.66%	7.14%	10.84%	6.22%	10.11%	17.72%	0.00%	6.99%	10.45%
English vowels[d]	0.93%	1.54%	2.44%	2.13%	2.29%	2.32%	1.74%	2.07%	0.00%	0.00%	0.04%	0.21%	0.00%	0.00%	0.02%	0.14%
Other vowels (i.e., incorrect)	1.31%	2.67%	2.71%	2.71%	3.13%	3.68%	2.26%	3.13%	0.13%	0.49%	0.11%	0.47%	0.63%	0.90%	0.13%	0.49%
Standard Vietnamese tones: PTC-S (/78)[b,c]	95.17%	3.98%	87.91%	13.08%	91.70%	5.28%	92.31%	7.82%	97.37%	3.34%	97.29%	4.19%	92.95%	2.72%	97.22%	3.78%
Northern tones[d]	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Central tones[d]	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.03%	0.21%	0.00%	0.00%	0.00%	0.00%	0.02%	0.15%
Southern tones[d]	2.86%	3.45%	2.29%	3.07%	2.38%	3.52%	2.56%	3.34%	2.07%	2.99%	1.75%	3.11%	5.77%	0.91%	2.01%	3.06%
Other tones (i.e., incorrect)	1.97%	2.72%	9.80%	12.43%	5.92%	4.77%	5.13%	7.33%	0.36%	0.83%	0.61%	1.49%	1.28%	1.81%	0.50%	1.20%

Note. Child Cluster 1: High Proficiency = children's reported ability to speak and understand both Vietnamese and English was high; Child Cluster 2: Low Proficiency = children's reported ability to speak and understand both Vietnamese and English was low; Child Cluster 3: English Proficient = children's reported ability to speak and understand English was higher (*well–very well*) than their ability to speak and understand Vietnamese; Adult Cluster 1: Vietnamese Proficient = adults' reported ability to speak and understand Vietnamese was higher than their ability to speak and understand English; Adult Cluster 2: Similar Proficiency = adults' reported ability to speak and understand Vietnamese and English was similarly high; Adult Cluster 3: English Proficient = adults' reported ability to speak and understand English was higher than their ability to speak and understand Vietnamese.

[a]Intelligibility in Context Scale (ICS) average scores were calculated based on the number of completed answers: 1 = *never*; 2 = *rarely*; 3 = *sometimes*; 4 = *usually*; 5 = *always*. [b]Vietnamese Speech Assessment (Phạm, Le, & McLeod, 2016). [c]Percentage of standard vowels correct (PVC-S) and percentage of standard tones correct (PTC-S) indicate that the consonants, vowels, and tones matched Standard Vietnamese productions. [d]Percentage of dialectal vowels correct (PVC-D) and percentage of dialectal tones correct (PTC-D) indicate that the consonants, vowels, and tones matched dialectal Vietnamese productions.

Table 9. English intelligibility and percentage of standard vowels correct (PVC-S) for children (n = 62) and adults (n = 77) in the English speech sample.

Speech element	Children (n = 62)								Adults (n = 77)							
	High Proficiency (n = 24)		Low Proficiency (n = 16)		English Proficient (n = 22)		Total		Vietnamese Proficient (n = 33)		Similar Proficiency (n = 36)		English Proficient (n = 8)		Total	
English (DEAP)	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
English intelligibility: ICS[a] (/5)	4.44	0.39	3.93	0.84	4.24	0.56	4.24	0.62	—	—	—	—	—	—	—	—
Australian English vowels: PVC-S[b,c] (/118)	96.40%	3.73%	94.54%	4.07%	96.88%	2.95%	96.09%	3.63%	83.98%	6.08%	87.57%	6.77%	94.81%	7.96%	86.78%	7.28%
American English vowels	1.20%	1.87%	0.95%	1.02%	0.96%	1.39%	1.05%	1.50%	1.59%	1.54%	2.10%	1.90%	1.59%	2.73%	1.83%	1.85%
Other vowels (i.e., incorrect)	2.40%	2.73%	4.45%	4.23%	2.16%	2.51%	2.84%	3.21%	14.43%	6.72%	10.10%	6.42%	3.39%	5.24%	11.27%	7.22%

Note. Child Cluster 1: High Proficiency = children's reported ability to speak and understand both Vietnamese and English was high; Child Cluster 2: Low Proficiency = children's reported ability to speak and understand both Vietnamese and English was low; Child Cluster 3: English Proficient = children's reported ability to speak and understand English was higher (*well–very well*) than their ability to speak and understand Vietnamese; Adult Cluster 1: Vietnamese Proficient = adults' reported ability to speak and understand Vietnamese was higher than their ability to speak and understand English; Adult Cluster 2: Similar Proficiency = adults' reported ability to speak and understand Vietnamese and English was similarly high; Adult Cluster 3: English Proficient = adults' reported ability to speak and understand English was higher than their ability to speak and understand Vietnamese.

[a]Intelligibility in Context Scale (ICS) average scores were calculated based on the number of completed answers: 1 = *never*; 2 = *rarely*; 3 = *sometimes*; 4 = *usually*; 5 = *always*. [b]Percentage of Standard Australian English vowels correct (PVC-S) indicate that the vowels matched Standard Australian English productions. [c]Percentage of dialectal vowels correct (PVC-D) indicate that the vowels matched dialectal English productions.

Table 10. Percentage of Vietnamese consonants correct (PCC-S, PCC-D)[a] across word positions and manner of articulation for children (n = 61) and adults (n = 77).

Speech element	Children (n = 61)								Adults (n = 77)							
	High Proficiency (n = 26)		Low Proficiency (n = 14)		English Proficient (n = 21)		Total		Vietnamese Proficient (n = 39)		Similar Proficiency (n = 36)		English Proficient (n = 2)		Total	
Vietnamese (VSA)[b]	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
Total PCC-S (/207)[a]	73.31%	8.49%	66.70%	8.87%	69.11%	8.23%	70.34%	8.78%	80.94%	5.08%	80.16%	5.22%	77.01%	0.52%	80.47%	5.09%
Total PCC-D (/137)[a]	92.20%	6.53%	82.55%	7.83%	85.74%	7.66%	87.76%	8.18%	98.41%	1.70%	97.65%	2.00%	97.81%	1.03%	98.04%	1.85%
Word initial (/84)	91.44%	7.62%	79.71%	7.70%	84.08%	8.23%	86.21%	9.11%	98.29%	1.81%	97.81%	2.07%	98.21%	0.84%	98.07%	1.91%
Within word (/1)	96.15%	19.61%	85.71%	36.31%	95.24%	21.82%	93.44%	24.96%	100.00%	0.00%	97.22%	16.67%	100.00%	0.00%	98.70%	11.40%
Word final (/52)	93.33%	6.70%	87.02%	9.75%	88.25%	8.35%	90.13%	8.40%	98.57%	2.98%	97.38%	4.09%	97.12%	1.36%	97.98%	3.54%
Plosives (/51)	92.96%	6.70%	86.24%	13.69%	87.53%	9.41%	89.55%	9.89%	98.59%	2.81%	97.44%	4.03%	98.04%	0.00%	98.04%	3.43%
Nasals (/37)	93.94%	5.92%	86.06%	8.82%	89.69%	7.59%	90.67%	7.78%	99.31%	1.48%	99.10%	1.83%	95.95%	1.91%	99.12%	1.72%
Fricatives (/30)	85.88%	13.46%	66.42%	14.08%	72.50%	14.63%	76.81%	16.04%	96.24%	4.79%	94.72%	4.94%	98.33%	2.36%	95.58%	4.85%
Lateral approximants (/3)	91.03%	27.58%	66.67%	47.14%	85.71%	32.61%	83.61%	35.29%	100.00%	0.00%	100.00%	0.00%	100.00%	0.00%	100.00%	0.00%
Semivowels (/16)	97.84%	4.31%	95.54%	5.16%	95.83%	6.65%	96.62%	5.42%	99.52%	2.21%	100.00%	0.00%	100.00%	0.00%	99.76%	1.58%
Voiceless (/57)	91.44%	8.51%	82.74%	10.49%	83.33%	10.24%	86.65%	10.31%	98.25%	3.01%	97.27%	3.78%	98.25%	0.00%	97.79%	3.37%
Voiced (/77)	94.11%	5.33%	84.72%	8.52%	90.08%	6.65%	90.56%	7.47%	98.77%	1.52%	98.70%	1.90%	97.40%	1.84%	98.70%	1.71%
Voiceless plosives (/39)	91.09%	8.52%	82.75%	16.02%	84.07%	11.72%	86.76%	12.09%	98.16%	3.67%	96.65%	5.27%	98.72%	1.81%	97.47%	4.49%
Voiced plosives (/12)	99.04%	2.72%	97.62%	6.88%	98.81%	5.46%	98.63%	4.85%	100.00%	0.00%	100.00%	0.00%	95.83%	5.89%	99.89%	0.95%

Note. Child Cluster 1: High Proficiency = children's reported ability to speak and understand both Vietnamese and English was high; Child Cluster 2: Low Proficiency = children's reported ability to speak and understand both Vietnamese and English was low; Child Cluster 3: English Proficient = children's reported ability to speak and understand English was higher (*well–very well*) than their ability to speak and understand Vietnamese; Adult Cluster 1: Vietnamese Proficient = adults' reported ability to speak and understand Vietnamese was higher than their ability to speak and understand English; Adult Cluster 2: Similar Proficiency = adults' reported ability to speak and understand Vietnamese and English was similarly high; Adult Cluster 3: English Proficient = adults' reported ability to speak and understand English was higher than their ability to speak and understand Vietnamese.

[a]Percentage of standard consonants correct (PCC-S) and percentage of dialect consonants correct (PCC-D) indicate that the consonants matched Standard, Northern, Central, or Southern Vietnamese productions. [b]Vietnamese Speech Assessment (VSA; Phạm, Le, & McLeod, 2016).

Table 11. Percentage of Standard Australian English consonants correct (PCC-S)[a] across word positions and manner of articulation for children (n = 62) and adults (n = 77).

Speech element	Children (n = 62)								Adults (n = 77)							
	High Proficiency (n = 24)		Low Proficiency (n = 16)		English Proficient (n = 22)		Total		Vietnamese Proficient (n = 33)		Similar Proficiency (n = 36)		English Proficient (n = 8)		Total	
English (DEAP[b])	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
Total PCC-S (/207)	88.84%	11.74%	69.56%	16.86%	85.04%	13.04%	82.51%	15.57%	77.90%	12.70%	90.68%	7.12%	95.71%	7.54%	85.73%	12.05%
Word initial (/94)	89.39%	11.15%	70.38%	16.57%	84.48%	14.11%	82.75%	15.54%	81.88%	11.83%	91.25%	7.07%	97.47%	4.69%	87.88%	10.71%
Within word (/48)	87.05%	13.77%	63.42%	18.13%	81.94%	16.78%	79.14%	18.48%	82.76%	12.49%	93.56%	5.34%	95.57%	7.99%	89.14%	10.77%
Word final (/65)	89.36%	12.48%	72.98%	19.17%	88.10%	10.34%	84.69%	15.32%	68.57%	17.94%	87.72%	10.25%	93.27%	11.89%	80.09%	17.33%
Consonant clusters (/31)	77.56%	22.67%	37.31%	25.51%	64.32%	29.35%	62.47%	30.12%	59.74%	24.04%	82.24%	13.13%	89.92%	14.16%	73.40%	22.09%
Plosives (/70)	94.78%	8.39%	81.52%	18.29%	93.27%	11.25%	90.82%	13.54%	77.13%	14.17%	89.39%	9.17%	96.96%	5.74%	84.92%	13.32%
Nasals (/27)	98.46%	1.87%	90.54%	11.49%	97.20%	4.54%	95.97%	7.18%	96.05%	4.73%	98.13%	3.38%	100.00%	0.00%	97.43%	4.05%
Fricatives (/60)	79.94%	18.37%	62.02%	19.67%	75.64%	17.75%	73.79%	19.58%	69.60%	17.01%	87.00%	11.33%	93.33%	10.50%	80.20%	16.73%
Affricates (/10)	89.58%	19.22%	61.81%	35.39%	88.64%	22.95%	82.08%	27.82%	57.78%	27.55%	83.96%	16.85%	93.75%	11.88%	73.76%	25.80%
Liquids (/31)	82.21%	23.28%	34.01%	30.02%	69.47%	31.45%	65.25%	33.79%	80.67%	18.13%	94.06%	7.33%	93.15%	13.40%	88.23%	14.94%
Glides (/9)	94.91%	12.25%	93.75%	10.71%	96.78%	8.59%	95.27%	10.57%	97.31%	6.23%	98.59%	5.95%	100.00%	0.00%	98.19%	5.78%
Voiceless (/90)	88.19%	12.50%	72.22%	16.70%	85.16%	13.12%	82.99%	15.15%	75.69%	14.91%	90.93%	7.66%	96.39%	5.75%	84.97%	13.84%
Voiced (/117)	89.34%	11.46%	67.53%	17.93%	84.95%	14.41%	82.15%	16.71%	79.61%	11.54%	90.48%	7.37%	95.19%	8.94%	86.31%	11.15%
Voiceless plosives (/41)	94.99%	7.17%	79.00%	20.33%	92.26%	12.98%	89.89%	14.91%	79.78%	14.38%	92.06%	7.87%	98.17%	3.39%	87.43%	12.82%
Voiced plosives (/29)	94.48%	10.50%	85.18%	16.72%	94.67%	10.19%	92.15%	12.78%	73.39%	16.56%	85.62%	13.16%	95.26%	9.20%	81.38%	16.11%

Note. Child Cluster 1: High Proficiency = children's reported ability to speak and understand both Vietnamese and English was high; Child Cluster 2: Low Proficiency = children's reported ability to speak and understand both Vietnamese and English was low; Child Cluster 3: English Proficient = children's reported ability to speak and understand English was higher (*well–very well*) than their ability to speak and understand Vietnamese; Adult Cluster 1: Vietnamese Proficient = adults' reported ability to speak and understand Vietnamese was higher than their ability to speak and understand English; Adult Cluster 2: Similar Proficiency = adults' reported ability to speak and understand Vietnamese and English was similarly high; Adult Cluster 3: English Proficient = adults' reported ability to speak and understand English was higher than their ability to speak and understand Vietnamese.

[a]PCC-S = percentage of Standard Australian English consonants correct. The definition of correct indicates that the consonants matched Standard Australian English. [b]Diagnostic Evaluation of Articulation and Phonology (DEAP; Dodd et al., 2002) samples comprised the Articulation and Phonology subtests plus two additional words (*chicken, go*) so that there were two + productions of each word-initial phoneme.

(M = 4.44, SD = 0.39), followed by the English Proficient cluster (M = 4.24, SD = 0.56) and then the Low Proficiency cluster (M = 3.93, SD = 0.84).

Children's Speech Production

The children's production of Standard Vietnamese consonants (including semivowels), vowels, and tones on the VSA (Phạm, Le, & McLeod, 2016; see Tables 8 and 10) and Standard Australian English consonants and vowels on the DEAP (Dodd et al., 2002; see Tables 9 and 11) was calculated and compared across clusters. Standard Vietnamese and dialectal variants of Vietnamese were considered as separate definitions of correct productions. Results are presented with respect to the children's cluster profiles.

Children's Vietnamese Consonant Accuracy

Children's Vietnamese consonant accuracy (for 137 consonants and semivowels) was calculated and compared across cluster profiles using Standard Vietnamese as the target (PCC-S: M = 70.34, SD = 8.78), and there was a significant correlation between age and PCC-S (r = .37, p = .003). Next children's Vietnamese consonant accuracy was calculated including dialectal variants as correct (PCC-D: M = 87.76, SD = 8.18; see Table 10). A paired-sample t test was conducted to compare the differences between PCC-S and PCC-D. Participants had significantly higher consonant accuracy for PCC-D compared with PCC-S, $p < 0.001$, Cohen's d = 3.55 (large effect). The highest average PCC-D was for the High Proficiency cluster (M = 92.20, SD = 6.53), followed by the English Proficient cluster (M = 85.74, SD = 7.66) and then the Low Proficiency cluster (M = 82.55, SD = 7.83; see Table 10).

Children's Vietnamese Vowel Accuracy

Standard Vietnamese vowel accuracy for 79 vowels on the VSA was used to document the child participants' percentage of vowels correct based on Standard Vietnamese (PVC-S; see Table 8). Overall, the children produced 82.05% (SD = 7.54) of vowels correctly using Standard Vietnamese as the target. The High Proficiency cluster had the highest PVC-S (M = 83.84, SD = 6.83), followed by the Low Proficiency cluster (M = 81.65, SD = 8.41) and then the English Proficient cluster (M = 80.11, SD = 7.63). Dialectal variants included Northern vowels (M = 5.25, SD = 4.33), Central vowels (M = 0.15, SD = 1.13), Southern vowels (M = 8.55, SD = 9.66), and English vowels (M = 1.74, SD = 2.07). Once all dialectal variants were included in the definition of correct (PVC-D), the children produced an average of 97.74% of vowels correctly.

Children's Vietnamese Tone Accuracy

The children had the opportunity to produce 78 tones within single words on the VSA. Overall, the children produced 92.31% (SD = 7.82) tones correctly using Standard Vietnamese as the target (PTC-S; see Table 8). PTC-D was mostly impacted by Southern Vietnamese dialectal pronunciation (typically relating to creaky (3) and dipping–rising (4) tones since Southern Vietnamese merges Tones 3 and 4). The High Proficiency cluster had the highest PTC-S (M = 95.17, SD = 3.98), followed by the English Proficient cluster (M = 91.70, SD = 5.28) and then the Low Proficiency cluster (M = 87.91, SD = 13.08). There was a greater amount of variation in the accuracy of tones by the Low Proficiency cluster (see Table 8).

Children's English Consonant Accuracy

The children's Standard Australian English consonant accuracy (PCC-S) was 82.51% (SD = 15.57) and was based on each child's productions of 207 English consonants on the DEAP (Dodd et al., 2002). The highest average PCC-S score was produced by the High Proficiency cluster (M = 88.84, SD = 14.00), followed by the English Proficient cluster (M = 85.04, SD = 13.03) and then the Low Proficiency cluster (M = 69.56, SD = 16.86; see Table 11).

Children's English Vowel Accuracy

The children's Standard Australian English vowel accuracy (PVC-S) was 96.09% (SD = 3.63%) based on the children's productions of vowels on the DEAP (see Table 9). Children achieved a similarly high PVC-S for all three proficiency clusters, ranging from 94.54% to 96.88% (see Table 9). A few vowels (1.05%) were produced using the Standard American English dialect (e.g., the first vowel in *zebra* was pronounced as /i/ instead of /e/), possibly demonstrating the influence of U.S. media or hearing Standard American English while living in Vietnam. There were 3.21% of the children's vowels that did not match either an Australian or Standard American English pronunciation, possibly demonstrating cross-linguistic influence from Vietnamese.

Children's Consonant Realizations and Phonological Patterns

Children's common, occasional, and rare realizations of consonants were determined for Vietnamese (see Tables 10 and 12) and English (see Tables 11 and 13) and then compared across cluster profiles for Vietnamese (see Table 14) and English (see Table 15). Similarly, children's occurrence of phonological patterns in Vietnamese (see Table 16) and English (see Table 17) were determined.

Table 12. Realizations of Vietnamese consonants by children (n = 61) and adults (n = 77).

Vietnamese manner	Consonant	No. of opportunities	Children (n = 61)				Adults (n = 77)			
			PCC-S[a]	Common (> 10%; including dialectal variants)	Occasional (5%–10%)	Rare[b] (1%–4.9%)	PCC-S[a]	Common (> 10%; including dialectal variants)	Occasional (5%–10%)	Rare[b] (1%–5%)
Plosives	/p/	5	74.50%	[p] (74.50%)	—	[pʰ] (4.97%)	82.34%	[p] (82.34%)	—	[pʰ] (3.38%)
						Ø (2.32%)		[b] (14.03%)		
						[t] (1.66%)				
	/b/	4	99.17%	[b] (99.17%)	—	—	99.68%	[b] (99.68%)	—	—
	/ʈʰ/	3	82.87%	[ʈʰ] (82.87%)	—	[tʰ] (4.42%)	100.00%	[ʈʰ] (100.00%)	—	—
						[t] (2.21%)				
						[d] (2.21%)				
						[k] (1.66%)				
						[h] (1.66%)				
						[f] (1.10%)				
						[θ] (1.10%)				
						[x] (1.10%)				
	/t/	10	75.99%	[t] (75.99%)	—	[tʰ] (4.11%)	85.97%	[t] (85.97%)	[k] (9.22%)	[tʰ] (2.86%)
				[k] (10.20%)		[kʰ] (3.13%)				[kᵖ] (1.69%)
						[p] (1.48%)				
						[kᵖ] (1.48%)				
						[d] (1.15%)				
						Ø (1.15%)				
	/d/	3	96.05%	[d] (96.05%)	—	[ɡ] (1.69%)	100.00%	[d] (100.00%)	—	—
	/ʈ/	3	6.67%	[c] (78.33%)	[t] (6.67%)	[ɡ] (2.78%)	30.74%	[c] (65.37%)	—	[ʦ] (2.60%)
						[ʤ] (1.67%)		[t] (30.74%)	(C) (1.30%)	
						[ʦ] (1.67%)				
						[p] (1.11%)				
						[t] (1.11%)				
						[k] (1.11%)				
						[ʃ] (1.11%)				
						[tʰ] (1.11%)				
						[tˀ] (1.11%)				
	/c/	3	86.26%	[c] (86.26%)	—	[ɡ] (3.85%)	97.84%	[c] (97.84%)	—	[ʦ] (2.16%)
						[ʤ] (2.20%)				
						[ʃ] (1.65%)				
						[t] (1.10%)				
						[f] (1.10%)				
	/k/	15	50.67%	[k] (50.67%)	[t] (6.32%)	[kʰ] (4.88%)	56.10%	[k] (56.10%)	[t] (5.63%)	Ø (3.12%)
				[kᵖ] (15.30%)	Ø (5.21%)	[p] (3.55%)		[kᵖ] (19.91%)		
				[c] (11.53%)				[c] (13.59%)		
	/ʔ/	5	57.89%	[ʔ] (57.89%)	—	—	81.77%	[ʔ] (81.77%)	—	—
				Ø (42.11%)				Ø (18.23%)		

(*table continues*)

Table 12. *(Continued).*

Vietnamese manner	Consonant	No. of opportunities	Children (*n* = 61)				Adults (*n* = 77)			
			PCC-S[a]	Common (> 10%; including dialectal variants)	Occasional (5%–10%)	Rare[b] (1%–4.9%)	PCC-S[a]	Common (> 10%; including dialectal variants)	Occasional (5%–10%)	Rare[b] (1%–5%)
Nasals	/m/	7	96.72%	[m] (96.72%)	—	[n] (2.11%)	99.63%	[m] (99.63%)	—	—
	/n/	11	80.00%	[n] (80.00%)	—	[m] (1.97%)	85.83%	[n] (85.83%)	—	[ŋ^{m}] (2.01%)
				[ŋ] (14.09%)		[ŋ^{m}] (1.82%)		[ŋ] (12.04%)		
						Ø (1.21%)				
	/ɲ/	4	84.23%	[ɲ] (84.23%)	—	[j] (2.90%)	99.35%	[ɲ] (99.35%)	—	—
				[n] (10.79%)		[m] (1.24%)				
	/ŋ/	15	52.83%	[ŋ] (52.83%)	—	[m] (2.55%)	59.83%	[ŋ] (59.83%)	—	—
				[ŋ^{m}] (16.65%)		Ø (1.22%)		[ŋ^{m}] (19.91%)		
				[n] (13.98%)				[ɲ] (13.68%)		
				[ɲ] (11.76%)				[n] (6.32%)		
Fricatives	/f/	3	96.11%	[f] (96.11%)	—	[b] (1.11%)	99.13%	[f] (99.13%)	—	—
	/v/	3	75.27%	[v] (75.27%)	—	[w] (4.40%)	99.13%	[v] (99.13%)	—	—
				[j] (12.64%)		[b] (2.20%)				
						[f] (2.20%)				
						[g] (1.10%)				
						[f/v] (1.10%)				
	/s/	3	86.44%	[s] (86.44%)	[θ] (5.65%)	[ʧ] (5.65%)	96.10%	[s] (96.10%)	—	[ʂ] (3.90%)
						[ʂ] (1.13%)				
						[j] (1.13%)				
	/z/	3	47.51%	[z] (47.51%)	—	[ð] (3.31%)	63.20%	[z] (63.20%)	—	[ʐ] (1.73%)
				[j] (37.02%)		[s] (2.76%)		[j] (33.33%)		
						[d] (1.66%)				
						[ɲ] (1.66%)				
						[l] (1.66%)				
						[θ] (1.10%)				
						[ʃ] (1.10%)				
	/ʂ/	3	12.92%	[s] (75.84%)	[θ] (7.30%)	[t] (1.12%)	21.65%	[s] (77.92%)	—	—
				[ʂ] (12.92%)				[ʂ] (21.65%)		
	/ʐ/	3	0.55%	[z] (28.57%)	[w] (7.69%)	[ð] (3.30%)	4.76%	[r] (39.39%	—	[ʐ] (4.76%)
				[ɹ] (20.33%)	[r] (6.04%)	(C) (3.30%)		[z] (36.36%)		[j] (1.73%)
				[ɹ^{w}, ɹ/w, w^{ɹ}] (11.54%)	[j] (6.04%)	[d] (2.20%)		[ɹ] (16.45%)		
						[s] (2.20%)				
						[θ] (1.65%)				
						[ʃ] (1.65%)				
						[ɣ] (1.10%)				

(*table continues*)

Table 12. *(Continued).*

Vietnamese manner	Consonant	No. of opportunities	Children (*n* = 61)				Adults (*n* = 77)			
			PCC-S[a]	Common (> 10%; including dialectal variants)	Occasional (5%–10%)	Rare[b] (1%–4.9%)	PCC-S[a]	Common (> 10%; including dialectal variants)	Occasional (5%–10%)	Rare[b] (1%–5%)
Fricatives (cont.)	/x/	3	52.46%	[x] (52.46%) [h] (32.79%)	[kʰ] (7.65%)	[k] (2.19%) [g] (1.64%) (C) (1.64%)	96.97%	[x] (96.97%)	—	[kʰ] (2.60%)
	/ɣ/	3	36.46%	[g] (48.07%) [ɣ] (36.46%)	[j] (7.18%)	[w] (2.21%) [ŋ] (1.10%) [ɹʷ, ɹ/w, wʲ] (1.10%) [g/ɣ] (1.10%) (C) (1.10%)	85.71%	[ɣ] (85.71%)	[g] (9.96%)	[g/ɣ] (4.33%)
	/h/	6	91.46%	[h] (91.46%)	Ø (6.89%)	—	98.05%	[h] (98.05%)	—	Ø (1.95%)
Lateral approximant	/l/	3	81.77%	[l] (81.77%)	[j] (9.94%)	[w] (2.76%) [d] (1.66%) [n] (1.66%) [v] (1.66%)	99.57%	[l] (99.57%)	—	—
Semivowels	/w/	10	96.02%	[w] (96.02%)	—	Ø (3.32%)	99.61%	[w] (99.61%)	—	—
	/j/	6	96.94%	[j] (96.94%)	—	Ø (2.22%)	100.00%	[j] (100.00%)	—	—

[a]The definition of correct indicates that the consonants matched Standard Vietnamese productions. Dialectal variants are included in the adjoining columns. [b]Mismatches that occurred < 1.0% were common but are not included in this table. (C) = unspecified consonant; Ø = omitted consonant. PCC-S = percentage of consonants correct–standard.

Table 13. Realizations of English consonants by children (n = 62) and adults (n = 77).

English manner[a]	Consonant		Children (n = 62)				Adults (n = 77)			
		No. of opportunities	PCC-S[b]	Common (> 10%)	Occasional (5%–10%)	Rare[c] (1%–4.9%)	PCC-S[b]	Common (> 10%)	Occasional (5%–10%)	Rare[c] (1%–4.9%)
Plosives	/p/	9	88.87%	[p] (88.87%)	—	[b] (3.41%), [k] (1.08%), Ø (4.49%)	89.02%	[p] (89.02%)	—	[pˉ] (4.05%) [b] (2.89%), Ø (1.45%), [f] (1.16%)
	/b/	15	93.10%	[b] (93.10%)	—	[p] (1.73%), Ø (1.51%)	89.84%	[b] (89.84%)	—	[bˉ] (3.30%), [p] (2.43%) [pˉ] (1.82%), Ø (1.22%)
	/t/	15	88.54%	[t] (88.54%)	—	Ø (4.11%), [k] (2.05%), [d] (1.62%)	83.94%	[t] (83.94%)	Ø (5.6%)	[tˉ] (3.13%), [t̪ʰ] (2.43%), [s] (1.22%), [c] (1.04%)
	/d/	5	91.29%	[d] (91.29%)	—	[t] (1.94%), [d̪] (1.61%)	88.51%	[d] (88.51%)	—	[dˉ] (4.70%), [t] (2.87%), [tˉ] (1.31%)
	/k/	17	91.70%	[k] (91.70%)	—	[t] (2.48%), [g] (1.05%), Ø (1.81%)	89.66%	[k] (89.66%)	—	[kˉ] (4.59%), [g] (2.37%), Ø (1.84%)
	/g/	10	89.84%	[g] (89.84%)	—	[k] (3.55%), [d] (2.58%), [ɣ] (1.45%)	63.93%	[g] (63.93%), [ɣ] (14.06%)	[k] (7.81%), [gˉ] (6.51%)	[kˉ] (2.99%), Ø (1.43%)
Nasals	/m/	7	96.64%	[m] (96.54%)	—	Ø (1.85%)	100.00%	[m] (100.00%)	—	—
	/n/	15	96.86%	[n] (96.86%)	—	Ø (1.62%), [m] (1.08%)	99.05%	[n] (99.05%)	—	—
	/ŋ/	5	93.20%	[ŋ] (93.20%)	—	[n] (3.24%), Ø (1.62%)	89.03%	[ŋ] (89.03%)	[n] (9.66%)	—
Fricatives	/f/	11	95.58%	[f] (95.58%)	—	—	95.60%	[f] (95.60%)	—	Ø (2.02%)
	/v/	6	80.38%	[v] (80.38%)	[b] (7.08%) Ø (5.99%)	[f] (1.63%) [w] (1.63%)	79.74%	[v] (79.74%), Ø (11.98%)	—	[b] (3.27%), [f] (2.40%), [w] (1.09%)
	/θ/	6	19.46%	[f] (54.05%), [θ] (19.46%)	[t] (8.92%), Ø (6.49%)	[s] (2.16%), [p] (1.35%), [d] (1.35%), [ʔ] (1.08%), [t̪ʰ] (1.35%)	45.65%	[θ] (45.65%), [t] (20.87%), [t̪ʰ] (18.48%),	—	[s] (3.70%), Ø (3.26%), [f] (2.39%), [ð] (1.30%), [tˉ] (1.09%), (C) (1.09%)
	/ð/	3	41.53%	[d] (45.90%), [ð] (41.53%)	[v] (6.01%)	[b] (2.19%), [g] (1.64%), [d̪] (1.09%)	61.30%	[ð] (61.30%), [d] (25.21%)	[z] (5.65%)	[t] (3.04%), [t̪ʰ] (2.61%), [θ] (1.30%)
	/s/	18	76.76%	[s] (76.76%), [θ] (10.09%)	Ø (6.67%)	[ʃ] (2.88%)	87.93%	[s] (87.93%)	—	[s] (4.77%), Ø (3.47%), [ʃ] (1.37%)

(*table continues*)

Table 13. *(Continued).*

English manner[a]	Consonant	No. of opportunities	Children (*n* = 62) PCC-S[b]	Common (> 10%)	Occasional (5%–10%)	Rare[c] (1%–4.9%)	Adults (*n* = 77) PCC-S[b]	Common (> 10%)	Occasional (5%–10%)	Rare[c] (1%–4.9%)
Fricatives (cont.)	/z/	6	65.23%	[z] (65.23%), Ø (10.24%)	[ð] (8.63%), [s] (6.47%)	[θ] (3.23%), [j] (1.08%), [ʃ] (1.08%)	70.09%	[z] (70.09%), [s] (10.92%)	Ø (9.83)	[ɹ] (1.75%), [dʒ] (1.75%)
	/ʃ/	6	86.29%	[ʃ] (86.29%)	[s] (5.65%)	[θ] (3.49%), (C) (1.34%), [ʧ] (1.08%)	77.06%	[ʃ] (77.06%), [s] (10.61%),	—	[ʂ] (3.90%), [ʧ] (3.03%), [t] (1.08%), [tˢ] (1.08%)
	/ʒ/	1	73.33%	[ʒ] (73.33%)	[dʒ] (8.33%), [z] (6.67%)	[c] (1.67%), [θ] (1.67%), [ð] (1.67%), [ʃ] (1.67%), [ʧ] (1.67%), Ø (1.67%), (C) (1.67%)	66.23%	[ʒ] (66.23%), [z] (25.97%)	—	[ʃ] (3.90%), [dʒ] (2.60%), [ʂ] (1.30%)
	/h/	3	99.46%	[h] (99.46%)	—	—	99.13%	[h] (99.13%)	—	—
Affricates	/ʧ/	5	84.79%	[ʧ] (84.79%)	—	[ʃ] (4.85%), [t] (3.24%), [c] (1.29%), (C) (1.62%), [s] (1.29%)	79.43%	[ʧ] (79.43%), [c] (11.72%)	—	[t] (1.82%), [ʃ] (1.56%), [ʂ] (1.04%), [dʒ] (1.04%)
	/dʒ/	5	79.35%	[dʒ] (79.35%)	—	Ø (3.87%), [d] (2.90%), [ʧ] (2.26%), [ʒ] (2.26%), [z] (1.94%), [t] (1.61%), [θ] (1.29%)	68.15%	[dʒ] (68.10%)	[z] (8.62%), [ʒ] (6.27%), Ø (5.22%)	[s] (3.13%), [g] (1.57%), [j] (1.57%), [ʃ] (1.04%) [j] (1.57%), [ʃ] (1.04%)
Liquids	/ɹ/	18	54.23%	[ɹ] (54.23%), [w] (26.62%), [ɹʷ, ɹ/w, wɹ] (12.41%)	Ø (5.22%)	—	86.15%	[ɹ] (86.15%)	[r] (8.70%)	Ø (3.55%)
	/l/	13	81.25%	[l] (81.25%)	[w] (9.13%),	Ø (4.00%), [j] (3.75%)	91.17%	[l] (91.17%)	—	Ø (4.71%), [n] (1.91%)
Glides	/j/	3	89.73%	[j] (89.73%)	—	[l] (4.86%), Ø (2.16%), [d] (1.08%), [w] (1.08%)	96.09%	[j] (96.09%)	—	[z] (1.74%), [dʒ] (1.30%)
	/w/	6	98.11%	[w] (98.11%)	—	Ø (1.08%)	99.35%	[w] (99.35%)	—	—

[a]Diagnostic Evaluation of Articulation and Phonology (DEAP) samples comprised the Articulation and Phonology subtests plus two additional words (*chicken, go*) so that there were two + productions of each word-initial consonant. [b]The definition of correct indicates that the vowels matched Standard Australian English. [c]Mismatches that occurred < 1.0% were common but are not included in this table. (C), unspecified consonant; Ø, omitted consonant. PCC-S = percentage of consonants correct–standard.

Table 14. Children's Standard Vietnamese consonant accuracy[a] across proficiency clusters (n = 61).

Standard Vietnamese	Consonant	No. of opportunities	Total tokens correct[a] (%)				Valid data
			High Proficiency (n = 26)	Low Proficiency (n = 14)	English Proficient (n = 21)	Total	
Plosives	/p/	5	75.38%	76.12%	72.38%	74.50%	61
	/b/	4	100.00%	100.00%	97.62%	99.17%	61
	/ʈʰ/	3	92.31%	63.41%	83.87%	82.87%	61
	/t/	10	78.46%	76.81%	72.38%	75.99%	61
	/d/	3	97.37%	90.00%	98.36%	96.05%	61
	/ʈ/	3	6.58%	4.88%	7.94%	6.67%	61
	/c/	3	93.59%	73.17%	85.71%	86.26%	61
	/k/	15	52.47%	48.53%	49.84%	50.67%	61
	/ʔ/	5	61.54%	56.52%	54.29%	57.89%	61
Nasals	/m/	7	100.00%	90.82%	96.60%	96.72%	61
	/n/	11	80.43%	79.05%	80.09%	80.00%	61
	/ɲ/	4	91.18%	72.73%	83.33%	84.23%	61
	/ŋ/	15	55.44%	50.25%	51.27%	52.83%	61
Fricatives	/f/	3	98.68%	90.24%	96.83%	96.11%	61
	/v/	3	78.21%	51.22%	87.30%	75.27%	61
	/s/	3	92.11%	84.21%	80.95%	86.44%	61
	/z/	3	50.00%	32.50%	53.97%	47.51%	61
	/ʂ/	3	6.67%	27.50%	11.11%	12.92%	61
	/ʐ/	3	0.00%	0.00%	1.59%	0.55%	61
	/x/	3	74.36%	42.86%	31.75%	52.46%	61
	/ɣ/	3	57.69%	21.95%	19.35%	36.46%	61
	/h/	6	89.74%	91.36%	93.65%	91.46%	61
Lateral approximants	/l/	3	91.03%	57.50%	85.71%	81.77%	61
Semivowels	/w/	10	97.68%	95.56%	94.26%	96.02%	61
	/j/	6	98.05%	93.83%	97.60%	96.94%	61

[a]The definition of correct indicates that the consonants matched Standard Vietnamese productions.

Children's Vietnamese Consonant Realizations

Vietnamese consonant realizations were compared across word positions and manner of articulation. Word-initial consonants were less likely to be correct than word-final and within-word consonants across all three clusters (see Table 10). For example, the Low Proficiency cluster produced an average of 79.71 (SD = 7.70) word-initial consonants correct compared with 87.02 (SD = 9.75) word-final consonants correct. Next, manner of articulation was compared across clusters. The High Proficiency cluster produced most Vietnamese plosives, nasals, lateral approximants, and semivowels correctly, with voiced plosives most likely to be correct (M = 99.04, SD = 2.72) and fricatives were least likely to be correct (M = 85.88, SD = 13.46; see Table 10). The Low Proficiency cluster produced more Vietnamese plosives, nasals, and semivowels correctly than lateral approximants (M = 66.67, SD = 47.14) and fricatives (M = 66.42, SD = 14.08). Vietnamese voiced plosives (M = 97.62, SD = 6.88) were more likely to be correct than Vietnamese voiceless plosives (M = 82.75, SD = 16.02). The English Proficient cluster demonstrated a similar pattern to the previous cluster, producing more plosives, nasals, and semivowels correctly than lateral approximants (M = 85.71, SD = 32.61) and fricatives (M = 72.50, SD = 14.63; see Table 10).

The PCC-S and common, occasional, and rare realizations for each Vietnamese consonant are found in Table 12. Overall, the children's most accurate Vietnamese consonants (80%–100%) with respect to Standard Vietnamese were as follows:

- plosives: /b/ (99.17%), /ʈʰ/ (82.87%), /d/ (96.05%), /c/ (86.26%);
- nasals: /m/ (96.72%), /n/ (80.00%), /ɲ/ (84.23%);
- fricatives: /f/ (96.11%), /s/ (86.44%), /h/ (91.46%);
- lateral approximant: /l/ (81.77%); and
- semivowels: /w/ (96.02%), /j/ (96.94%).

Table 15. Children's Standard Australian English consonant accuracy (PCC-S)[a] across proficiency clusters (n = 62).

Standard Australian English	Consonant	No. of opportunities	Total tokens correct (%)				
			High Proficiency (n = 24)	Low Proficiency (n = 16)	English Proficient (n = 22)	Total	Valid data
Plosives	/p/	9	95.37%	74.83%	91.92%	88.87%	62
	/b/	15	96.10%	85.71%	95.15%	93.10%	62
	/t/	15	93.04%	75.63%	92.99%	88.54%	62
	/d/	5	92.50%	81.25%	97.23%	91.29%	62
	/k/	17	96.57%	83.96%	91.94%	91.70%	62
	/ɡ/	10	93.75%	80.00%	92.73%	89.84%	62
Nasals	/m/	7	97.62%	92.86%	98.04%	96.54%	62
	/n/	15	99.44%	91.56%	97.86%	96.86%	62
	/ŋ/	5	96.67%	85.00%	95.41%	93.20%	62
Fricatives	/f/	11	95.08%	91.95%	98.76%	95.58%	62
	/v/	6	88.19%	60.22%	86.15%	80.38%	62
	/θ/	6	27.78%	6.25%	20.00%	19.46%	62
	/ð/	3	50.00%	30.43%	40.00%	41.53%	62
	/s/	18	85.85%	63.03%	76.71%	76.76%	62
	/z/	6	77.08%	45.26%	66.67%	65.23%	62
	/ʃ/	6	87.50%	78.13%	90.91%	86.29%	62
	/ʒ/	1	79.17%	60.00%	76.19%	73.33%	62
	/h/	3	98.61%	100.00%	100.00%	99.46%	62
Affricates	/ʧ/	5	93.33%	64.56%	90.00%	84.79%	62
	/ʤ/	5	85.83%	58.75%	87.27%	79.35%	62
Liquids	/ɹ/	18	74.19%	19.23%	57.83%	54.23%	62
	/l/	13	93.59%	55.88%	85.92%	81.25%	62
Glides	/j/	3	90.28%	85.42%	92.31%	89.73%	62
	/w/	6	97.22%	97.92%	99.24%	98.11%	62

Note. Diagnostic Evaluation of Articulation and Phonology samples comprised the Articulation and Phonology subtests plus two additional words (*chicken, go*) so that there were two + productions of each word-initial phoneme.

[a]The definition of correct indicates that the consonants matched Standard Australian English productions.

The following voiceless plosives were less likely (< 80%) to be correct, /p/ (74.50%), /t/ (75.99%), and /k/ (50.67%), and the retroflex plosive /ʈ/ (6.67%) was typically produced as the dialectal variant [c] (78.33%) or one of a large number of other substitutions. The nasal /ŋ/ typically was produced with Standard Vietnamese pronunciation [ŋ] (52.83%) and was substituted with other nasals including dialectal variants: [ŋ^{m}] (16.65%), [n] (13.98%), and [ɲ] (11.76%). The majority of the fricatives were produced correctly using the Standard form < 80% of the time: /v/ (75.27%), /z/ (47.51%), /ʂ/ (12.92%), /ʐ/ (0.55%), /x/ (52.46%), and /ɣ/ (36.46%). The fricative /ɣ/ was more likely to be produced as the English consonant [g] (48.07%) than [ɣ] (36.46%). Twenty different consonants were used as productions of /ʐ/ possibly due to cross-linguistic, dialectal, maturational, language experience, and environmental (ambient phonology) factors (see Table 12).

The order of consonant accuracy for the three proficiency clusters is shown in Table 14. The summaries for each proficiency cluster present each consonant divided into early (90%–100%), middle (70%–89%), and late (< 70%) acquisition in order of accuracy from highest to lowest.

- High Proficiency: Early: /b, m, f, j, w, d, c, ʈʰ, s, ɲ, l/; Middle: /h, n, t, v, p, x/; Late: /ʔ, ɣ, ŋ, k, z, ʂ, ʈ, ʐ/
- Low Proficiency: Early: /b, w, j, h, m, f, d/; Middle: /s, n, t, p, c, ɲ/; Late: /ʈʰ, l, ʔ, v, ŋ, k, x, z, ʂ, ɣ, ʈ, ʐ/
- English Proficient: Early: /d, b, j, f, m, w, h/; Middle: /v, c, l, ʈʰ, ɲ, s, n, t, p/; Late: /ʔ, z, ŋ, k, x, ɣ, ʂ, ʈ, ʐ/

Children's Vietnamese Phonological Patterns

Children's Vietnamese phonological patterns are provided in Table 16. A number of substitution (systemic) processes were common, but the actual occurrence was low: Fronting (of velars and palatals) was produced by 51

Table 16. Vietnamese phonological patterns[a] for children in proficiency clusters (n = 61).

Phonological pattern	Word position	No. of words phonological process is possible	Target consonants	Example	High Proficiency (n = 26)				Low Proficiency (n = 14)				English Proficient (n = 21)				Total			
					No. of children		Occurrence[b]		No. of children		Occurrence[b]		No. of children		Occurrence[b]		No. of children		Occurrence[b]	
					n	%	*M*	*SD*	*n*	%	*M*	*SD*	*n*	%	*M*	*SD*	*n*	%	*M*	*SD*
Substitution (systemic) patterns[a]																				
Velar and palatal fronting	WI, WF	48	/c, k, ʔ, ɲ, ŋ, x, ɣ/	/c/ → [t] /ɲ/ → [n]	21	80.77	7.69%	7.31%	11	78.57	11.31%	9.66%	19	90.48	11.81%	9.39%	51	83.61	9.94%	8.70%
Stopping of fricatives	WI	30	/f, v, s, z, ʂ, ʐ, x/	/v/ → [b] /s/ → [t]	15	57.69	4.87%	5.52%	12	85.71	9.76%	11.36%	18	85.71	7.14%	4.63%	45	73.77	6.78%	7.17%
Gliding of fricatives and nasals	WI, WF	67	/f, v, s, z, ʂ, ʐ, x, m, n, ɲ, ŋ/	/ʐ/ → [j] /h/ → [w] /ɲ/ → [j]	14	53.85	3.90%	4.61%	10	71.43	4.69%	5.82%	14	66.67	3.20%	4.10%	38	62.30	3.84%	4.70%
Denasalization (does not include dialectal *n* → l)	WI, WF	37	/m, n, ɲ, ŋ/	/m/ → [b] /n/ → [d] /ɲ/ → [c]	5	19.23	0.62%	1.39%	4	28.57	0.97%	1.71%	6	28.57	1.42%	2.65%	15	24.59	0.97%	1.98%
Nasalization (does not include dialectal l → *n*)	WI, WF	100	All except /m, n, ɲ, ŋ/	/d/ → [n] /ʔ/ → [ŋ] /l/ → [m]	26	100.00	2.73%	1.00%	14	100.00	2.64%	1.01%	21	100.00	2.67%	1.06%	61	100.00	2.69%	1.01%
Aspiration	WI, WW, WF	134	All except /tʰ/	/t/ → [tʰ] /n/ → [tʰ] /s/ → [tʰ]	12	46.15	1.03%	1.35%	7	50.00	1.71%	3.13%	16	76.19%	2.67%	3.33%	35	57.38%	1.75%	2.66%
Deaspiration	WI	3	/tʰ/	/tʰ/ → [t] /tʰ/ → [m] /tʰ/ → [s]	3	11.54	5.13%	15.47%	7	50.00	30.95%	38.04%	3	14.29	7.94%	23.34%	13	21.31	12.02%	26.55%
Glottal replacement	WI, WW, WF	126	All except /ʔ, h/	/t/ → [ʔ] /x/ → [h]	8	30.77	0.55%	0.95%	8	57.14	0.96%	1.04%	12	57.14	1.17%	1.19%	28	45.90	0.86%	1.08%
Syllable structure processes[a]																				
Initial consonant deletion	WI	84	/n/ → [Ø]		23	88.46	3.94%	3.26%	12	85.71	4.68%	3.57%	19	90.48	4.02%	2.59%	54	88.52	4.14%	3.09%
Final consonant deletion	WF	52	/–n/ → [Ø]		9	34.62	0.89%	1.46%	10	71.43	3.16%	2.98%	9	42.86	1.37%	1.94%	28	45.90	1.58%	2.21%
Semivowel deletion	WI, WW, WF	16	/–j–/ → [Ø]		4	15.38	1.44%	4.07%	7	50.00	4.91%	5.58%	7	33.33	3.27%	5.46%	18	29.51	2.87%	5.05%

Note. WI = word initial; WF = word final; WW = within word.

[a]Phonological patterns are defined in Phạm and McLeod (2019). [b]Dialectal variants were excluded from the calculation of the occurrence of phonological processes.

Table 17. English phonological patterns for children according to proficiency clusters (n = 62).

Phonological pattern	No. of words phonological process is possible	Examples consonant	Example word	High Proficiency (n = 24)				Low Proficiency (n = 16)				English Proficient (n = 22)				Total			
				No. of children		Occurrence[a]		No. of children		Occurrence[a]		No. of children		Occurrence[a]		No. of children		Occurrence[a]	
				n	%	*M*	*SD*	*n*	%	*M*	*SD*	*n*	%	*M*	*SD*	*n*	%	*M*	*SD*
Substitution (systemic) patterns																			
Velar fronting	30	/k/ → [t]	duck – [dʌt]	5	20.83	0.97%	2.30%	8	50.00	6.25%	13.38%	4	18.18	3.64%	14.87%	17	27.42	3.28%	11.25%
Palatal fronting	16	/ʃ/ → [s]	sheep – [sɪp]	6	25.00	3.39%	7.59%	6	37.50	6.25%	10.70%	5	22.73	2.84%	8.12%	17	27.42	3.93%	8.64%
Stopping of fricatives	57	/ʃ/ → [t]	sheep – [ʦip]	20	83.33	3.65%	3.14%	13	81.25	8.00%	6.72%	16	72.73	4.86%	4.40%	49	79.03	5.21%	4.95%
Stopping of affricates	9	/ʤ/ → [d]	sausage – [sɒsɪd]	5	20.83	6.02%	14.65%	6	37.50	9.03%	14.18%	2	9.09	1.01%	3.27%	13	20.97	5.02%	12.02%
Deaffrication	9	/ʤ/ → [s]	orange – [ɒɹɛns]	3	12.50	1.39%	3.75%	8	50.00	15.28%	18.98%	5	22.73	3.54%	7.18%	16	25.81	5.73%	12.03%
Gliding of liquids	31	/ɹ/ → [w]	frog – [fwɒg]	9	37.50	7.66%	14.45%	14	87.50	44.35%	27.89%	11	50.00	18.18%	24.53%	34	54.84	20.97%	26.32%
Prevocalic voicing	46	/p/ → [b]	pram – [bɹæm]	3	12.50	0.36%	1.05%	9	56.25	2.85%	3.43%	8	36.36	1.98%	5.11%	20	32.26	1.58%	3.65%
Postvocalic devoicing	23	/d/ → [t]	bird – [bɜt]	8	33.33	5.62%	10.99%	11	68.75	11.96%	14.16%	9	40.91	5.73%	11.42%	28	45.16	7.29%	12.14%
Fricative simplification	9	/θ/ → [f]	thumb – [fʌm]	19	79.17	36.57%	26.11%	14	87.50	31.25%	22.30%	19	86.36	43.94%	24.60%	52	83.87	37.81%	24.76%
Liquid simplification	18	/ɹ/ → [l]	ring – [lɪŋ]	4	16.67	0.93%	2.12%	2	12.50	0.69%	1.90%	0	0.00	0.00%	0.00%	6	9.68	0.54%	1.66%
Backing	38	/p/ → [k]	pram – [kɹæm]	5	20.83	0.88%	2.86%	9	56.25	6.41%	10.26%	8	36.36	2.27%	3.66%	22	35.48	2.80%	6.12%
Assimilation	2	/j/ → [l]	yellow – [lɛloʊ]	3	12.50	12.50%	33.78%	1	6.25	6.25%	25.00%	1	4.55	2.27%	10.66%	5	8.06	7.26%	25.36%
Syllable structure processes																			
Consonant cluster simplification	31	/kɹ/ → [kl]	crab – [klæb]	19	79.17	10.22%	10.28%	16	100.00	33.27%	17.30%	19	86.36	27.42%	24.86%	54	87.10	22.27%	20.61%
Cluster reduction	31	/kɹ/ → [k]	crab – [kæb]	12	50.00	8.60%	17.79%	15	93.75	22.58%	22.23%	13	59.09	4.99%	7.76%	40	64.52	10.93%	17.66%
Epenthesis	30	/kɹ/ → [kəɹ]	crab – [kəræb]	3	12.50	0.42%	1.13%	6	37.50	2.71%	4.08%	1	4.55	0.15%	0.71%	10	16.13	0.91%	2.43%
Initial consonant deletion	72	/h/ → [Ø]	this – [ɪs]	7	29.17	1.10%	2.56%	7	43.75	3.56%	6.04%	5	22.73	0.82%	2.00%	19	30.65	1.64%	3.76%
Final consonant deletion	58	/–n/ → [Ø]	train – [tɹeɪ]	14	58.33	2.87%	3.99%	13	81.25	9.05%	8.27%	10	45.45	1.80%	3.08%	37	59.68	4.09%	5.92%

[a]Dialectal variants were excluded from the calculation of the occurrence of phonological process.

children (83.61%). However, the occurrence per child was low (M = 9.94%, SD = 8.70%): stopping of fricatives by 45 children (73.77%, M = 6.78%, SD = 7.17%), gliding by 38 children (62.30%, M = 3.84%, SD = 4.70%), nasalization (excluding dialectal /l/ → [n]) by 61 children (100.00%, M = 2.69%, SD = 1.01%), aspiration by 35 children (57.38%, M = 1.75%, SD = 2.66%), and glottal replacement by 28 children (45.90%, M = 0.86%, SD = 1.08%). Few children produced denasalization and deaspiration. The Low Proficiency cluster often (but not always as in the example of fronting) was more likely to demonstrate these substitution (systemic) patterns.

Syllable structure patterns also were produced by many of the children, but again the occurrence per child was low (see Table 16). The most common was initial consonant deletion, which was produced by 54 children (88.52%, M = 4.14%, SD = 3.09%), followed by final consonant deletion by 28 children (45.90%, M = 1.58%, SD = 2.21%) and semivowel deletion by 18 children (88.52%, M = 2.87%, SD = 5.05%).

Children's English Consonant Realizations

Children's English consonant realizations were compared across word position and phoneme categories (see Table 11). The children's consonant accuracy (PCC-S) was not markedly influenced by word position across the three proficiency clusters; word-initial and word-final English consonant accuracy was similar, with slightly lower accuracy for within-word English consonants. Consonant clusters were less accurate than singleton consonants for all three proficiency clusters, and this was particularly apparent for the Low Proficiency cluster who produced 37.31% consonant clusters correct compared with the High Proficiency cluster who had 77.56% consonant clusters correct. Plosives, nasals, and glides were more likely to be accurate compared with fricatives, affricates, and liquids. The Low Proficiency cluster was more likely to produce the voiced plosives /b, d, g/ correctly compared with voiceless plosives /p, t, k/ (see Table 11), and this finding was similar for Vietnamese (see Table 10). In contrast, voicing (voiced vs. voiceless consonants) did not markedly influence English consonant accuracy across the High Proficiency and English Proficient clusters.

The PCC-S and mismatches (common, occasional, and rare) for each English consonant are found in Table 13. Overall, the children's most accurate English consonants (80%–100%) were as follows:

- plosives: /b/ (93.10%), /k/ (91.70%), /d/ (91.29%), /g/ (89.84%), /p/ (88.87%), /t/ (88.54%);
- nasals: /n/ (96.86%), /m/ (96.54%), /ŋ/ (93.20%);
- fricatives: /h/ (99.46%), /f/ (95.58%), /ʃ/ (86.29%), /v/ (80.38%);
- affricates: /ʧ/ (84.79%); and
- glides and liquids: /w/ (98.11%), /j/ (89.73%), /l/ (81.25%).

The next most accurate English consonants (produced correctly between 70% and 80% on average) were the fricatives /s/ (76.76%) and /ʒ/ (73.33%) and the affricate /ʤ/ (79.35%). The children's least accurate English consonants (< 70%) were the fricative /θ/ (19.46%) that was typically substituted with [f] (95.58%); the fricative /ð/ (41.53%) that was typically substituted with [d] (45.90%); the fricative /z/ (65.23%) that was either deleted, interdentalized /ð/, or substituted with /s/; and the liquid /ɹ/ (54.23%) that was typically substituted with [w] (26.62%; see Table 13).

The order of Standard Australian English consonant accuracy for each of the children's three clusters was similar with some minor differences as shown in Table 15. The following cluster profile summaries present each consonant in order of accuracy from highest to lowest, divided into early (90%–100%), middle (70%–89%), and late (< 70%) acquisition.

- High Proficiency: Early: /n, h, m, w, ŋ, k, b, p, f, g, l, ʧ, t, d, j/; Middle: /v, ʃ, s, ʤ, ʒ, z, ɹ /; Late: /ð, θ/
- Low Proficiency: Early: /h, w, m, f, n/; Middle: /b, j, ŋ, k, d, g, ʃ, t, p/; Late: /ʧ, s, v, ʒ, ʤ, l, z, ð, ɹ, θ/
- English Proficient: Early: /h, w, f, m, n, d, ŋ, b, t, g, j, k, p, ʃ, ʧ/; Middle: /ʤ, v, l, s, ʒ/; Late: /z, ɹ, ð, θ/

Children's English Phonological Patterns

English phonological patterns were analyzed (see Table 17). The most common substitution (systemic) phonological patterns involved fricatives. Overall, 52 (83.87%) children demonstrated at least one occurrence of fricative simplification (M occurrence = 37.81%, SD = 24.76%), and 49 (79.03%) children demonstrated at least one occurrence of stopping of fricatives; however, the actual occurrence was low (M occurrence = 5.21%, SD = 4.95%). Gliding was produced at least once by 34 (54.84%) children (M occurrence = 20.97%, SD = 26.32%). Postvocalic devoicing, postvocalic voicing, and backing were produced by just over a third of the children, but the occurrence per child was low (see Table 17). The Low Proficiency cluster was more likely to demonstrate these substitution patterns.

Syllable structure patterns were produced by many of the children, and the most common patterns were consonant cluster simplification (n = 54, 87.10%)

and cluster reduction (n = 40, 64.52%; see Table 17). Other common syllable structure patterns were final consonant deletion (n = 37, 59.68%) and initial consonant deletion (n = 19, 30.65%), but again, the occurrence per child was low (see Table 17). It is important to note that consonant clusters do not occur in Vietnamese and final consonants are rare; therefore, cross-linguistic influences may have impacted the children's productions of these syllable shapes. Again, the Low Proficiency cluster was more likely to demonstrate these syllable structure patterns.

Adults' Speech Production

The adult participants' accuracy of Vietnamese consonants, semivowels, vowels, and tones on the VSA (Phạm, Le, & McLeod, 2016; see Table 8) and English consonants and vowels on the DEAP (Dodd et al., 2002; see Table 9) was calculated and compared across cluster profiles. Standard and dialectal variants were compared across cluster profiles.

Adults' Vietnamese Consonant Accuracy

Adults' Vietnamese consonant and semivowel accuracy on the VSA was calculated first using Standard Vietnamese as the target (PCC-S: M = 80.47, SD = 5.09) and then by adding dialectal variants as acceptable targets (PCC-D: M = 98.04, SD = 1.85; see Table 10). A paired-sample t test was conducted to compare the differences between PCC-S and PCC-D. Participants had significantly higher consonant accuracy for PCC-D compared with PCC-S, $p < .001$, Cohen's d = 3.99 (large effect). The adults produced most of the 137 Vietnamese consonants correctly, especially when dialectal variants were the reference point for correct productions (PCC-D; see Table 8). The Vietnamese Proficient cluster had the highest average PCC-D (M = 98.41, SD = 1.70), followed by the Similar Proficiency cluster (M = 97.65, SD = 2.00) and the English Proficient cluster (M = 97.81, SD = 1.03).

Adults' Vietnamese Vowel Accuracy

Overall, the adults produced 85.81% (SD = 6.36) of vowels correctly using Standard Vietnamese as the target (PCC-S; see Table 8). The adults in the Similar Proficiency cluster (M = 86.29, SD = 6.17) had the highest, followed by the Vietnamese Proficient cluster (M = 85.72, SD = 6.58) and then the English Proficient cluster (M = 79.11, SD = 0.90; see Table 8). Once all dialectal variants were included in the definition of correct (PVC-D), the adults produced an average of 99.87% of vowels correctly. This definition included production of Northern vowels (M = 6.63, SD = 4.69), Central vowels (M = 0.43, SD = 1.93), Southern vowels (M = 6.99, SD = 10.45), and English vowels (M = 0.02, SD = 0.14).

Adults' Vietnamese Tone Accuracy

The adult participants produced most tones (M = 97.22, SD = 3.78) similarly to the Standard Vietnamese target (PTC-S; see Table 8). The highest PTC-S was produced by the Vietnamese Proficient cluster (M = 97.37, SD = 3.34), the Similar Proficiency cluster (M = 97.29, SD = 4.19), and then the English Proficient cluster (M = 92.95, SD = 2.72; see Table 8). Tones produced according to the Southern dialect accounted for most nonstandard tone productions. Once all dialectal variants were included in the definition of correct (PTC-D), the adults produced an average of 99.50% of tones correctly.

Adults' English Consonant Accuracy

The adults' accuracy of Australian English consonants was calculated based on their productions of words on the DEAP (Dodd et al., 2002; see Table 11). The English Proficient cluster (M = 95.91, SD = 7.54) had the highest average PCC-S based on each participants' productions of 207 consonants, followed by the Similar Proficiency cluster (M = 90.67, SD = 7.12) and then the Vietnamese Proficient cluster (M = 77.90, SD = 12.70; see Table 11). The adults' English consonants were generally more likely to match Australian English consonant targets than their children's consonants likely due to their children's ongoing maturation of speech production skills.

Adults' English Vowel Accuracy

The adults in the English Proficient cluster (M = 94.81, SD = 7.96) had the highest average PVC-S based on each adult participants' productions of 118 Standard Australian English vowels, followed by the Similar Proficiency cluster (M = 87.57, SD = 6.77) and then the Vietnamese Proficient cluster (M = 83.98, SD = 6.08; see Table 9). Similar to the children, some vowels (1.83%) were produced using Standard American English (e.g., *zebra*); however, 11.27% of the vowels did not match either Standard Australian or Standard American English pronunciation, mostly demonstrating cross-linguistic influences from Vietnamese. For example, the word *moon* /mun/ was frequently produced as [mʊn], and unstressed schwa vowels in *elephant* and *television* often were stressed.

Adults' Vietnamese Consonant Realizations

Adults' common, occasional, and rare realizations of consonants were determined for Vietnamese (see Table 12) and English (see Table 13) and then compared across cluster profiles for Vietnamese (see Table 13) and English (see Table 15). Vietnamese consonant realizations were compared across word positions and manner of articulation (see Table 10). Due to the adults' high overall accuracy of the production of Vietnamese consonants, the proficiency clusters had similar levels of accuracy regarding word position, plosives, nasals, fricatives, lateral approximants, and

Table 18. Adults' Vietnamese consonant accuracy[a] across proficiency clusters (n = 77).

			Total tokens correct[a] (%)				
Standard Vietnamese	Consonant	No. of opportunities	Vietnamese Proficient (n = 39)	Similar Proficiency (n = 36)	English Proficient (n = 2)	Total	Valid data
Plosives	/p/	5	82.05%	100.00%	100.00%	82.34%	77
	/b/	4	100.00%	100.00%	87.50%	99.68%	77
	/ʈʰ/	3	100.00%	100.00%	100.00%	100.00%	77
	/t/	10	87.95%	85.28%	60.00%	85.97%	77
	/d/	3	100.00%	100.00%	100.00%	100.00%	77
	/ʈ/	3	32.48%	25.00%	100.00%	30.74%	77
	/c/	3	98.29%	98.15%	83.33%	97.84%	77
	/k/	15	56.23%	55.74%	60.00%	56.10%	77
	/ʔ/	5	81.54%	82.68%	70.00%	81.77%	77
Nasals	/m/	7	99.27%	100.00%	100.00%	99.63%	77
	/n/	11	86.95%	86.11%	59.09%	85.83%	77
	/ɲ/	4	100.00%	99.31%	87.50%	99.35%	77
	/ŋ/	15	59.83%	59.81%	60.00%	59.83%	77
Fricatives	/f/	3	98.28%	100.00%	100.00%	99.13%	77
	/v/	3	99.15%	99.07%	100.00%	99.13%	77
	/s/	3	95.73%	97.22%	83.33%	96.10%	77
	/z/	3	63.25%	66.67%	0.00%	63.20%	77
	/ʂ/	3	23.08%	18.52%	50.00%	21.65%	77
	/ʐ/	3	6.84%	2.78%	0.00%	4.76%	77
	/x/	3	99.15%	99.44%	100.00%	96.97%	77
	/ɣ/	3	92.31%	77.78%	100.00%	85.71%	77
	/h/	6	97.44%	98.61%	100.00%	98.05%	77
Lateral approximants	/l/	3	99.15%	100.00%	100.00%	99.57%	77
Semivowels	/w/	10	99.23%	100.00%	100.00%	99.61%	77
	/j/	6	100.00%	100.00%	100.00%	100.00%	77

[a]The definition of correct indicates that the consonants matched Standard Vietnamese productions.

semivowels (see Table 10). Standard Vietnamese consonant accuracy (see Table 18) is presented from most to least accurate below for the three proficiency clusters:

- Vietnamese Proficient: 90%–100% /b, ʈʰ, d, ɲ, j, m, w, v, x, l, c, f, h, s, ɣ/; 70–89% /t, n, p, ʔ/; < 70% /z, ŋ, k, ʈ, ʂ, ʐ/
- Similar Proficiency: 90%–100% /b, ʈʰ, d, j, m, w, l, f, p, x, ɲ, v, h, c, s, n/; 70%–89% /t, ʔ, ɣ/; < 70% /z, ŋ, k, ʈ, ʂ, ʐ/
- English Proficient: 90%–100% /ʈʰ, d, j, m, w, l, f, p, x, v, h, ɣ, ʈ/; 70%–89% /b, ɲ, c, s, ʔ/; < 70% /t, ŋ, k, n, ʂ, z, ʐ/. There were only two participants in this cluster.

The consonants that were not produced similarly to Standard Vietnamese predominantly were dialectal variants or due to cross-linguistic influences (see Table 12). For example, /z/ was produced as [j] 33.33% of the time, which reflects the typical pronunciation in Southern Vietnam.

Adults' English Consonant Realizations

Adults' English consonant realizations were compared across cluster profiles (see Table 12). There were similar results to the children for the English Proficient and Similar Proficiency clusters regarding word position and voicing. Each of these clusters had the highest accuracy for plosives, nasals, fricatives, liquids, and glides. The Vietnamese Proficient and Similar Proficiency clusters demonstrated less accuracy than the English Proficient cluster for consonant clusters, affricates, and voiceless plosive consonants possibly reflecting cross-linguistic influences since consonant clusters and affricates do not exist in Vietnamese and there are few voiceless plosive consonants (see Table 12).

The order of consonant accuracy (cf. Standard Australian English) for the adults' cluster profiles was different from each other and was different from the children's (presented from most to least accurate, see Table 19):

- Vietnamese Proficient: 90%–100% /m, n, w, h, j, f, l, b, d, k, ŋ, p/; 70%–89% /s, ɹ, t/; < 70% /v, ʧ, ʃ, ʒ, z, ʤ, ð, g, θ/

Table 19. Adults' English consonant accuracy[a] across proficiency clusters (n = 77).

			Total tokens correct (%)				
Standard Australian English	Consonant	No. of opportunities	Vietnamese Proficient (n = 33)	Similar Proficiency (n = 36)	English Proficient (n = 8)	Total	Valid data
Plosives	/p/	9	82.09%	92.90%	100.00%	89.02%	77
	/b/	15	85.43%	92.19%	97.50%	89.84%	77
	/t/	15	74.55%	90.13%	95.00%	83.94%	77
	/d/	5	83.64%	91.57%	95.00%	88.51%	77
	/k/	17	83.18%	93.29%	100.00%	89.66%	77
	/g/	10	47.87%	72.22%	92.50%	63.93%	77
Nasals	/m/	7	100.00%	100.00%	100.00%	100.00%	77
	/n/	15	98.79%	99.07%	100.00%	99.05%	77
	/ŋ/	5	82.32%	92.74%	100.00%	89.03%	77
Fricatives	/f/	11	91.11%	98.73%	100.00%	95.60%	77
	/v/	6	69.39%	87.44%	87.50%	79.74%	77
	/θ/	6	23.35%	55.81%	91.67%	45.65%	77
	/ð/	3	48.98%	67.59%	83.33%	61.30%	77
	/s/	18	79.26%	93.97%	96.53%	87.93%	77
	/z/	6	56.19%	78.70%	87.50%	70.09%	77
	/ʃ/	6	60.61%	90.28%	85.42%	77.06%	77
	/ʒ/	1	57.58%	66.67%	100.00%	66.23%	77
	/h/	3	97.98%	100.00%	100.00%	99.13%	77
Affricates	/ʧ/	5	64.24%	91.06%	90.00%	79.43%	77
	/ʤ/	5	51.22%	77.09%	97.50%	68.15%	77
Liquids	/ɹ/	18	75.63%	94.41%	92.36%	86.15%	77
	/l/	13	87.82%	93.56%	94.23%	91.17%	77
Glides	/j/	3	94.90%	96.30%	100.00%	96.09%	77
	/w/	6	98.48%	100.00%	100.00%	99.35%	77

[a]The definition of correct indicates that the consonants matched Standard Australian English productions. Diagnostic Evaluation of Articulation and Phonology (DEAP) samples comprised the Articulation and Phonology subtests plus two additional words (*chicken, go*) so that there were two + productions of each word-initial phoneme.

- Similar Proficiency: 90%–100% /m, w, h, n, f, j, ɹ, s, l, k, p, ŋ, b, d, ʧ, ʃ, t/; 70%–89% /v, z, ʤ, g/; < 70% /ʒ, ð, θ/
- English Proficient: 90%–100% /m, w, h, n, f, j, k, p, ŋ, ʒ, b, ʤ, s, d, t, l, g, ɹ, θ, ʧ/; < 90% / v, z, ʃ, ð/.

Cross-linguistic transfer was evident for English consonants that were not in the Vietnamese consonant inventory. For example, the consonant /g/ was one of the most difficult consonants for the Vietnamese Proficient (47.87%) and Similar Proficiency (72.22%) adult clusters compared to the English proficient cluster (92.50%). These findings could inform selection of consonant targets for English-language teaching for Vietnamese speakers.

Consonant Matches Between Children and Adult Family Members

Comparisons were made between children's and family members' production of each consonant within each Vietnamese and English word.

Vietnamese Consonant Matches Between Children and Family Members

Children's and family members' productions of each Vietnamese consonant within the words on the VSA were compared. Matching was undertaken for the 54 children who produced all 77 words on the VSA. The children were most likely to produce Vietnamese consonants in a similar way to their mother (M = 82.23%, SD = 9.82%, valid n = 49) and other adults in their lives (M = 78.77%, SD = 11.13%, valid n = 28) and were least likely to match the consonant pronunciation of their siblings (M = 77.30%, SD = 7.42%, valid n = 10).

English Consonant Matches Between Children and Family Members

Children's and family members' productions of each English consonant within the words on the DEAP were compared. Matching was undertaken for the 58 children who produced all 82 words on the DEAP. The children were likely to produce the English consonants in a similar

way to their mother (M = 75.06%, SD = 14.33%, valid n = 53) and other adults in their lives (M = 75.57%, SD = 14.48%, valid n = 28). They were least likely to match the consonant pronunciation of their younger siblings (M = 71.37%, SD = 15.83%, valid n = 11), possibly due to the impact of maturation (McLeod, Margetson, et al., 2022).

Discussion

This is a landmark study since it presents the largest bilingual data set of children's speech acquisition to date. This large-scale study of 154 Vietnamese-English–speaking children and adult family members used the VietSpeech Multilingual Assessment Protocol (see the Appendix) to demonstrate that children's acquisition of Vietnamese and English was influenced by cross-linguistic, dialectal, maturational, language experience, and environmental (ambient phonology) factors. This study identified the impact of bilingual language experience on speech production (Aim 1); compared children's intelligibility (Aim 2) and production of Vietnamese and Australian English including dialectal variants in the definition of accuracy (Aim 3); identified consonant realizations and phonological patterns (Aim 4); considered ambient phonology produced by adult family members and compared adult family members' production of Vietnamese and Australian English (Aim 5); and finally considered matches between children's and adult family members' production of consonants, vowels, and tones (Aim 6).

Impact of Multilingual Language Experience

Children's intelligibility and acquisition of Vietnamese and Australian English was impacted by speaking two languages (see Tables 8–11). On average, they were rated by their parents to be *usually* intelligible in English (M = 4.24) and *sometimes–usually* intelligible in Vietnamese (M = 3.87; see Tables 8 and 9) on the 5-point ICS (McLeod et al., 2012). The children achieved higher consonant accuracy scores for English compared with Vietnamese. Specifically, they could produce more Standard English consonants correctly (PCC-S: 82.51) than Standard Vietnamese (PCC-S: 70.34; see Tables 10 and 11); however, the children had a greater number of Vietnamese consonants correct if dialect was considered (PCC-D: 87.76). A similar pattern was found for vowels (English PVC-S: 96.09; English PVC-D: 99.98; Vietnamese PVC-S: 82.05; Vietnamese PVC-D: 97.74; see Tables 8 and 9).

Vietnamese voiced plosives, nasals, vowels, and tones were more often correct than voiceless plosives and fricatives (see Table 10). English plosives, nasals, glides, and vowels were more often correct than fricatives and affricates (see Table 11). The latest acquired Standard Vietnamese consonants by children were /t, v, p, k, ŋ, x, z, ɣ, ʈ, ʂ, ʐ/ (< 80%; see Table 14), and the latest acquired English consonants were /s, ʒ, ʤ, θ, ð, z, ɹ/ (< 80%; see Table 15) and none of the consonants in these lists overlapped. Vietnamese word-initial consonants had lower accuracy than word-final consonants (see Table 10), whereas word position rarely influenced English consonant accuracy (see Table 11). The children's nonadult realizations of specific consonants also differed cross-linguistically. For example, the voiceless plosives /p/ and /t/ were more likely to be produced correctly in English words (88.89%; see Table 13) than in Vietnamese words (74.50%; see Table 12). For /t/ in English words, children used [k, d] or omitted it, whereas there were seven different nonadult productions for /t/ in Vietnamese words [k, t^h, k^h, p, k^p, d] and omitted. Additionally, Vietnamese /ʈʰ/ is a separate phoneme, and there were eight different nonadult productions for /ʈʰ/ including the fricatives [h, f, θ, x].

Bidirectional cross-linguistic transfer was found across the participants, particularly for unshared phonemes. For example, some children produced the English plosive /g/ as the Vietnamese fricative [ɣ] (see Table 13); conversely, some children produced the Vietnamese fricative /ɣ/ as the English plosive [g] (see Table 12). Some children produced the English fricative /θ/ as the Vietnamese plosive [ʈʰ], and other children produced the Vietnamese plosive /ʈʰ/ as the English fricative [θ]. Previous researchers also have identified cross-linguistic transfer in multilingual speech acquisition. For example, Fabiano-Smith and Barlow (2010) found 25% of the Spanish–English bilingual children demonstrated cross-linguistic transfer, but bidirectional transfer was rare. More data are required to fully document which nonadult realizations are typical; however, understanding cross-linguistic differences between the languages is important for accurate differential diagnosis between typical multilingual acquisition and SSD.

Impact of Dialect

Researchers have highlighted the importance of considering dialectal variations in the interpretation of speech data to reduce misdiagnosis of multilingual and multidialectal speakers (Ball & Bernhardt, 2008; MacLeod & Demers, 2023; McAlister et al., 2023; Toohill et al., 2012). In this study, distinctions were made between accuracy of standard and acceptable dialectal productions based on data from linguistic reviews (e.g., Phạm & McLeod, 2016) and typically developing monolingual children (Phạm & McLeod, 2019). The acronyms PCC-S and PCC-D were coined in the VietSpeech study (McLeod, Margetson, et al., 2022) and are recommended

for use by speech researchers and clinicians. The impact of dialect was considered by comparing the participants' production of Standard Vietnamese consonants (PCC-S) with dialectal Vietnamese consonants (PCC-D) and Standard English and Vietnamese vowels (PVC-S) and dialectal vowels (PVC-D). Dialectal differences between consonants are not a feature of Australian English (Cox & Fletcher, 2017); however, in the future, dialectal variants of bilingual Vietnamese–Australian English could be classified.

The children and adults were found to produce more Vietnamese consonants, semivowels, vowels, and tones correctly when dialectal variants were considered to be correct. For example, both child and adult participants had a significantly higher Vietnamese consonant accuracy for PCC-D (children: 87.76, adults: 99.87) compared with PCC-S (children: 70.34, adults: 80.47) with a large effect size (Cohen's d = 3.55 children, d = 3.99 adults; see Table 10). Vietnamese vowels were produced correctly by over 80% of participants using Standard Vietnamese (PVC-S children: 82.05, adults: 85.81; see Table 8). Most of the nonstandard productions of Vietnamese vowels were explained by their pronunciations matching with acceptable dialectal variants from Northern and Southern Vietnam and occasionally with using English vowels (see Table 8). Therefore, when dialectal variants were considered, most participants could produce most vowels correctly (PVC-D children: 97.74, adults: 99.87). Vietnamese tones were produced correctly by most participants using Standard Vietnamese as the target (PTC-S children: 92.31, adults: 97.22; see Table 8). Southern Vietnamese speakers typically accounted for the group who produced nonstandard productions of tones because Southern Vietnamese speakers pronounce five tones (Le et al., 2022), whereas Standard and Northern Vietnamese speakers use six tones (Phạm & McLeod, 2019). Thus, when dialectal variants were considered, most participants could produce most tones correctly (PTC-D children: 97.74, adults: 99.50).

The importance of dialectal differences when considering children's Vietnamese consonant realizations was evident in Table 12. There was a large variety of nonstandard productions that monolingual English-speaking SLPs may label as "atypical errors" or evidence of a motor speech problem if dialectal variants were unknown (cf. Margetson et al., 2023b). For example, only 50.67% of children's productions of /k/ matched the Standard Vietnamese production /k/. One of the most common productions (15.30%) was when children produced as [k^{p}] often with puffed cheeks. This is a typical dialectal production of /k/; however, if it is not understood as a dialectal variant on an assessment score form, it may be misinterpreted as an atypical motor speech error by English-speaking SLPs.

Impact of Maturation, Proficiency, and Language Experience

Across the entire sample of children (N = 69), post hoc correlations revealed that age was significantly correlated with PCC-S for English (r = .65, p < .001) and PCC-S for Vietnamese (r = .37, p = .003) but not with their mean scores on the ICS or iTalk (Speech). Post hoc correlations also revealed that parent concern was significantly correlated with PCC-S for English (r = −.42, p < .001) and PCC-S for Vietnamese (r = −.37, p = .003) and was significantly negatively correlated with ICS for English (r = −.48, p < .001) and Vietnamese (r = −.42, p = .001) but not with their mean scores for iTalk (Speech) for English or Vietnamese. Therefore, age and parental concern were not the only explanatory factors for speech acquisition. Parental report of children's proficiency in spoken languages added another dimension to understanding children's speech acquisition.

Children's language proficiency and use was reported by their parents in the preassessment questionnaire (see Table 7) for Vietnamese and English. These data were compared to their performance on speech tasks to identify whether experience impacts speech. Participants were statistically clustered based on their reported proficiency in speaking and understanding both spoken languages (cf. Wang et al., 2021). The three identified child clusters had approximately equal numbers: High Proficiency, 37.68%; Low Proficiency, 30.43%; English Proficient, 31.88% (see Table 7). Children's English and Vietnamese consonant accuracy and intelligibility ranged from highest to lowest for the following clusters: High Proficiency > English Proficient > Low Proficiency. The Low Proficiency cluster was the youngest group, and 76.19% of their parents had concerns (*yes* or *a little*) about how they talked and made speech sounds. However, there was no significant interaction between age and concern for the Low Proficiency group, suggesting that parents rated proficiency with the age of their child in mind.

There were three clusters of adults (Vietnamese Proficient, Similar Proficiency, and English Proficient). As would be expected, the Vietnamese Proficient cluster had a higher level of accuracy of Vietnamese consonants, vowels, and tones than the English Proficient cluster. The English Proficient clusters were higher for English consonants and vowels (see Table 11). Results were reported using these three clusters so that language proficiency informed outcomes (see Supplemental Material S4), as speakers would not be expected to achieve speech accuracy in languages in which they are not proficient. Consideration of language proficiency reduces the risk of misdiagnosis in multilingual speakers (Paradis et al., 2010).

Impact of the Ambient Language Within the Child's Environment

Adult Family Members' Vietnamese and English Speech Production

The adults' productions of consonants, vowels, and tones were generally more likely to match Vietnamese consonant targets than their children's productions (see Tables 8 and 10), possibly due to the children's maturing speech (i.e., speech still contained developmentally appropriate errors), as well as cross-linguistic influences due to children's greater exposure to English during their speech acquisition (since most children were born in Australia and most adults were born in Vietnam). For English consonants, English-proficient adults had more matches than English-proficient children and similarly proficient adults; however, for English vowels, English-proficient children were more likely to match Standard English vowels.

While it was not surprising that the majority (82.23%) of children's consonants were produced the same way as their mother, there were also high matches between the children and other adults in their lives (78.77%), as well as with their siblings (77.30%). Some of the children's mismatches were accounted for by maturation as illustrated in the three-generational VietSpeech case study (McLeod, Margetson, et al., 2022). The case study showed mismatches occurred when children pronounced a consonant in one word the same way as their mother and in a different word the same way as their grandparent or sibling. Inconsistency or variability in consonant productions can impact SLPs' clinical decision making since monolingual children's inconsistency has been interpreted as signaling difficulties with phonological or motor planning requiring intervention (e.g., McIntosh & Dodd, 2008). Without undertaking contrastive analyses with family members, SLPs may erroneously conclude that children's inconsistency/variability indicates SSD and requires intervention.

Theoretical Implications

The emergence approach (Davis & Bedore, 2013), the theory of linguistic multicompetence (Cook, 2016), and the Culturally Responsive Teamwork Framework (Hopf et al., 2021) informed the research design, data collection, and analysis extending previous research to understand children's development. In keeping with the emergence approach (Davis & Bedore, 2013), these bilingual Vietnamese–Australian children's speech acquisition was influenced by both by intrinsic factors (e.g., maturation) and extrinsic factors (interactions with adults from their ambient phonological environment).

As outlined in the introduction, the theory of linguistic multicompetence (Cook, 2016) considers relationships between languages, that monolingual norms are not appropriate for multilingual speakers, and that "the multicompetent speaker perceives and encodes information differently from monolinguals." Examples of cross-linguistic transfer within the data support the concept of the multiple languages in the one mind influencing the other (Paradis & Genesee, 1996). The level of exposure to languages has been found to impact multilingual consonant accuracy (Ruiz-Felter et al., 2016). For example, in this study, children who were in the English Proficient cluster had fewer nonstandard productions of English consonants than children in the Low Proficiency cluster (see Table 15). Similarly, adult family members who were in the Vietnamese Proficient cluster had more nonstandard productions of English consonants than those who were in the English Proficient cluster or the Similar Proficiency cluster (see Table 19). Gildersleeve-Neumann et al. (2008) examined English acquisition in Spanish–English bilingual children and found that balanced bilinguals produced more English phoneme errors than monolingual English speakers and bilingual children with high proficiency in English.

Clinical Implications

This culturally responsive VietSpeech research demonstrates how children's acquisition of Vietnamese and English was influenced by cross-linguistic, ambient phonology, dialectal, maturation, and language experience factors. SLPs can use these results as a resource with multilingual Vietnamese–English population and use the lessons from this research within clinical practice with multilingual children. Specifically, these data highlight the usefulness of the VietSpeech Multilingual Assessment Protocol of (a) assessing all spoken languages, (b) assessing the speech of important adults in the child's life to understand the impact of the ambient environment upon speech production and undertake contrastive analysis, (c) defining what is considered "correct" by considering the dialect, and (d) collecting a comprehensive language profile during case history discussions to map language experience and proficiency. The VietSpeech Multilingual Assessment Protocol can support SLPs to think critically about whether children genuinely present with SSD across all languages spoken, or whether their speech production does not conform to the standardized norm because is reflective of their rich and diverse linguistic influences (see McLeod, Verdon, et al., 2017).

This study has additional applications for SLPs working in clinical practice. First, the study identified common nonstandard productions in English for Vietnamese–Australian

adults and children. These data provide a useful clinical resource for SLPs attempting to differentially diagnose between typical multilingual speech development and the presence of a true SSD. Productions of Vietnamese or English that did not match the standard or dialectal targets could be classified five ways identified in the VietSpeech three-generation case study:

- cross-linguistic transfer: where a phoneme that was in one language was used in the other language (e.g., *gà* (chicken): /ɣɑ/ → [gɑ] using English /g/ instead of Vietnamese /ɣ/);
- cross-dialectal transfer: where a phoneme in one dialect is used in another dialect;
- developmental/maturation production: where the production matches the form typically used by a younger child (e.g., *rùa* (turtle): /ruo/ → [ɹʷuo]; this example also demonstrates cross-linguistic transfer);
- assimilation: where the phoneme has a similar voice/place/manner to the target phoneme (e.g., *nhíp* (tweezers): /ɲip/ → [ɲim]); and
- error: where the pronunciation cannot be explained by the reasons above (e.g., *lược* (comb): /lɯɤk/ → [vɯɤk]; McLeod, Margetson, et al., 2022, p. 18).

One explanation for cross-linguistic transfer may be due to phonetic similarity (Flege, 1981), whereby multilingual children may collapse more subtle differences between phonemes across their languages, at times producing them with the same form in both languages, while at other times reverting to the standard production in each language, as was observed in the Vietnamese production of /t/ realized as [t̪ʰ] on some occasions and as the target [t] on other occasions. In English, these are allophones, but in Vietnamese, they are discrete phonemes.

This culturally responsive research demonstrated the impact of dialect on children's speech. Data were considered in relation to the Vietnamese (and English) dialects that children were hearing in their ambient environment. This is particularly useful when working with multilingual Vietnamese children living outside Vietnam, as their primary input in Vietnamese will be from those in their immediate environment and based on the dialect spoken by those communication partners and not through formal channels such as Vietnamese schooling, which adheres to the Standard Vietnamese form.

The current data also highlight the importance of SLPs' understanding of language features and differences between the languages spoken by the child when undertaking a multilingual assessment. For example, Vietnamese is a monosyllabic language with no consonant clusters and few final consonants. Armed with this knowledge, lower accuracy of English consonant clusters, word-final consonants, and polysyllabic words can be interpreted as potential cross-linguistic transfer, rather than features of SSD. The comparison between child and adult data further emphasized this point since adult family members were less likely to correctly produce consonant clusters and consonants in word-final position than their children, who may have been more proficient in English and less affected by cross-linguistic transfer (McLeod, Margetson, et al., 2022).

Data from adult family members may be clinically useful for SLPs engaged in Intelligibility Enhancement to provide multilingual speakers with strategies for effective communication in social, education, and employment settings (Blake et al., 2020, 2021). This study identified that Vietnamese speakers found the following English features to be difficult: consonants not found in Vietnamese (e.g., /θ, ɡ/), voiceless plosives /p, t, k/, and consonants in unfamiliar contexts such as word-final consonants and consonant clusters.

Limitations

While this study contributes to current understanding of speech production by Vietnamese–Australian children and adults, there are a number of limitations and considerations for future research. First, it is important to recognize that "typical" multilingual data are difficult to collect due to the diversity in the language experiences of multilingual speakers, even when they share the same languages. These data may not be applicable to all Vietnamese–English speakers since speech production varies among multilingual speakers, depending on personal, social, and environmental factors. Australian Vietnamese migrants differ in socioeconomic and sociocultural ways—as well as in the Vietnamese dialects spoken. The first wave of Vietnamese migrants came to Australia from Southern Vietnam as refugees during the war in the 1970s. Participants in this study were primarily from the second wave of migration who relocated from across Vietnam to Australia for educational or employment reasons. Further research is needed not only in relation to Vietnamese migration but also to understand multilingual speech and language acquisition among other populations and languages.

Second, the intended participant sample was redefined due to extensive Australian government restrictions on visiting homes and travel due to COVID-19. Specifically, the age range of the sample was expanded in order to access more participants. Even so, the size of the sample was still smaller than planned.

Finally, single-word elicitations were collected (not conversational samples), and analysis was largely conducted at the phoneme level, rather than the word or

phrase level. Some word-level (phonotactic) features were considered; however, the focus on phoneme analysis limited identification of some phonological and prosodic patterns. Future research may examine further phonological aspects and consider the impact of cross-linguistic transfer upon connected speech samples for Vietnamese–English speakers.

Conclusions

Children's acquisition of Vietnamese and English consonants, vowels, and tones was influenced by cross-linguistic, dialectal, maturational, language experience, and environmental (ambient phonology) factors. This study supports a paradigm shift in multilingual research highlighting the usefulness of the VietSpeech Multilingual Assessment Protocol that includes (a) assessing all languages spoken by multilingual children, (b) considering the ambient phonology in the environment by assessing adult family members, (c) considering dialect in the definition of accuracy, and (d) acknowledging language experience to inform differential diagnosis of SSD in multilingual populations. Despite more than half of the world's population speaking more than one language, in the field of speech-language pathology, this study presents the most comprehensive analysis of the largest bilingual data set of the speech of children and adult family members. The data presented in this study provide SLPs with a road map for considering the acquisition and production of speech sounds across two languages in both children and adults. This study demonstrates how family contrastive analysis can be used to consider the impact of ambient phonology upon a multilingual children's speech acquisition, which can assist in differentiation between productions influenced by ambient phonology, cross-linguistic transfer, and the presence of SSD. Enhanced accuracy in the identification of SSD among multilingual children will support equitable access to early intervention for those children who need it and reduce overdiagnosis of children whose speech differs as a result of linguistic diversity.

This study describes the VietSpeech Multilingual Assessment Protocol for culturally responsive speech assessment and analysis of multilingual children. Based on the combined theoretical contributions of the emergence approach (Davis & Bedore, 2013), the theory of linguistic multicompetence (Cook, 2016), and the Culturally Responsive Teamwork Framework (Hopf et al., 2021), the VietSpeech Multilingual Assessment Protocol provides a decolonizing pathway to deconstruct narrow interpretations of speech production in relation to standardized norms, allowing the consideration of multiple factors that influence typical speech acquisition. While the VietSpeech study supports the paradigm shift for Vietnamese-English–speaking children in Australia, it can be used to develop greater understanding of diverse linguistic populations around the world. With increased understanding of the influences on children's multilingual and multidialectal speech acquisition, accuracy of differential diagnosis within speech-language pathology can be enhanced.

Data Availability Statement

Data are available from the authors upon reasonable request.

Acknowledgments

This research was supported by an Australian Research Council Discovery Grant (DP180102848) awarded to Sharynne McLeod and Sarah Verdon and an Australian Postgraduate Awards Scholarship awarded to Van H. Tran. The authors thank Katherine White, Holly McAlister, Craig Margetson, Olivia Richards, Jake Sheader, and Evelyn Sowter for assistance with data entry and analysis.

References

Australian Bureau of Statistics. (2011). *2011 Household form*. http://www.abs.gov.au/ausstats/abs@.nsf/lookup/2903.0Main%20Features162011

Australian Bureau of Statistics. (2016). *Cultural diversity in Australia, 2016*. https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/2071.0~2016~Main%20Features~Cultural%20Diversity%20Article~60#

Australian Bureau of Statistics. (2018). *IRSAD*. https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/2033.0.55.001~2016~Main%20Features~IRSAD~20

Australian Bureau of Statistics. (2021a). *Cultural diversity: Census*. https://www.abs.gov.au/statistics/people/people-and-communities/cultural-diversity-census/2021

Australian Bureau of Statistics. (2021b). *Snapshot of Australia*. https://www.abs.gov.au/statistics/people/people-and-communities/snapshot-australia/2021

Ball, J., & Bernhardt, B. M. (2008). First Nations English dialects in Canada: Implications for speech-language pathology. *Clinical Linguistics & Phonetics, 22*(8), 570–588. https://doi.org/10.1080/02699200802221620

Bedore, L. M., & Peña, E. D. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism, 11*(1), 1–29. https://doi.org/10.2167/beb392.0

Bialystok, E. (2011). Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 65*(4), 229–235. https://doi.org/10.1037/a0025406

Blake, H. L., McLeod, S., & Verdon, S. (2020). Intelligibility Enhancement assessment and intervention: A single-case experimental design with two multilingual university students.

Clinical Linguistics & Phonetics, 34(1–2), 1–20. https://doi.org/10.1080/02699206.2019.1608470

Blake, H. L., McLeod, S., Verdon, S., & Fuller, G. (2018). The relationship between spoken English proficiency and participation in higher education, employment and income from two Australian censuses. *International Journal of Speech-Language Pathology, 20*(2), 202–215. https://doi.org/10.1080/17549507.2016.1229031

Blake, H. L., Verdon, S., & McLeod, S. (2021). Exploring multilingual speakers' perspectives on their intelligibility in English. *Speech, Language and Hearing, 24*(3), 133–144. https://doi.org/10.1080/2050571X.2019.1585681

Camarota, S. A., & Zeigler, K. (2014). *One in five U.S. residents speaks foreign language at home, record 61.8 million: Spanish, Chinese, and Arabic speakers grew most since 2010*. Center for Immigration Studies. https://cis.org/sites/cis.org/files/camarota-language.pdf

Cao, X. H. (2006). *Tiếng Việt mãy vấn đề ngữ âm–ngữ pháp–ngữ nghĩa* [Vietnamese some issues in phonology-syntax-semantics]. Khoa học xã hội.

Clyne, M. (2008). The monolingual mindset as an impediment to the development of plurilingual potential in Australia. *Sociolinguistic Studies, 2*(3), 347–366. https://doi.org/10.1558/sols.v2i3.347

Cook, V. (2016). Premises of multi-competence. In V. Cook & Li Wei (Eds.), *The Cambridge handbook of linguistic multi-competence* (pp. 1–25). Cambridge University Press. https://doi.org/10.1017/CBO9781107425965.001

Cox, F., & Fletcher, J. (2017). *Australian English pronunciation and transcription* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/9781316995631

Crowe, K., & McLeod, S. (2020). Children's English consonant acquisition in the United States: A review. *American Journal of Speech-Language Pathology, 29*(4), 2155–2169. https://doi.org/10.1044/2020_AJSLP-19-00168

Crowther, D., Kim, S., Lee, J., Lim, J., & Loewen, S. (2021). Methodological synthesis of cluster analysis in second language research. *Language Learning, 71*(1), 99–130. https://doi.org/10.1111/lang.12428

Czech Statistics Office. (2011). *Cizinci v České republice podle dat sčítání lidu - 2011*. https://www.czso.cz/csu/czso/cizinci-v-ceske-republice-podle-dat-scitani-lidu-2011-dxfg9l7fpy

Davis, B. L., & Bedore, L. M. (2013). *An emergence approach to speech acquisition: Doing and knowing*. Psychology Press. https://doi.org/10.4324/9780203375303

Đinh, L. T., & Nguyễn, V. H. (1998). *Cơ cấu ngữ âm tiếng Việt* [The structure of Vietnamese phonology]. Giáo dục.

Đỗ, X. T., & Lê, H. T. (2005). *Tiếng Việt 2* [Vietnamese 2]. Đại học Sư phạm Hà Nội.

Đoàn, T. T. (2003). *Ngữ âm tiếng Việt* [Vietnamese phonetics]. Đại học Quốc gia Hà Nội.

Dodd, B., Hua, Z., Crosbie, S., Holm, A., & Ozanne, A. (2002). *Diagnostic Evaluation of Articulation and Phonology (DEAP)*. The Psychological Corporation.

Easton, C., & Verdon, S. (2021). The influence of linguistic bias upon speech-language pathologists' attitudes toward clinical scenarios involving nonstandard dialects of English. *American Journal of Speech-Language Pathology, 30*(5), 1973–1989. https://doi.org/10.1044/2021_AJSLP-20-00382

Eberhard, D. M., Gary, F. S., & Charles, D. F. (2023). *Ethnologue: Languages of the world (26th ed.)*. http://www.ethnologue.com

Ellis, C., Jacobs, M., & Kendall, D. (2021). The impact of racism, power, privilege, and positionality on communication sciences and disorders research: Time to reconceptualize and seek a pathway to equity. *American Journal of Speech-Language Pathology, 30*(5), 2032–2039. https://doi.org/10.1044/2021_AJSLP-20-00346

Fabiano-Smith, L., & Barlow, J. A. (2010). Interaction in bilingual phonological acquisition: Evidence from phonetic inventories. *International Journal of Bilingual Education and Bilingualism, 13*(1), 81–97. https://doi.org/10.1080/13670050902783528

Fabiano-Smith, L., & Goldstein, B. A. (2010). Phonological acquisition in bilingual Spanish-English–speaking children. *Journal of Speech, Language, and Hearing Research, 53*(1), 160–178. https://doi.org/10.1044/1092-4388(2009/07-0064)

Farrugia-Bernard, A. M. (2018). Speech-language pathologists as determiners of the human right to diversity in communication for school children in the U.S. *International Journal of Speech-Language Pathology, 20*(1), 170–173. https://doi.org/10.1080/17549507.2018.1406002

Flege, J. E. (1981). The phonological basis of foreign accent: A hypothesis. *TESOL Quarterly, 15*(4), 443–455. https://doi.org/10.2307/3586485

General Statistics Office. (2020). *Kết quả toàn bộ Tổng điều tra dân số và nhà ở năm 2019* [Completed results of the 2019 Viet Nam population and housing census]. https://www.gso.gov.vn/en/data-and-statistics/2020/11/completed-results-of-the-2019-viet-nam-population-and-housing-census/

Gildersleeve-Neumann, C. E., Kester, E. S., Davis, B. L., & Pena, E. D. (2008). English speech sound development in preschool-aged children from bilingual English–Spanish environments. *Language, Speech, and Hearing Services in Schools, 39*(3), 314–328. https://doi.org/10.1044/0161-1461(2008/030)

Gildersleeve-Neumann, C. E., & Wright, K. L. (2010). English speech acquisition in 3- to 5-year-old children learning Russian and English. *Language, Speech, and Hearing Services in Schools, 41*(4), 429–444. https://doi.org/10.1044/0161-1461(2009/09-0059)

Glascoe, F. P. (2000). *Parents' Evaluation of Developmental Status: Authorized Australian version*. Centre for Community Child Health.

Goldfeld, S., O'Connor, M., Mithen, J., Sayers, M., & Brinkman, S. (2014). Early development of emerging and English-proficient bilingual children at school entry in an Australian population cohort. *International Journal of Behavioral Development, 38*(1), 42–51. https://doi.org/10.1177/0165025413505945

Goldstein, B. A., & Iglesias, A. (2001). The effect of dialect on phonological analysis: Evidence from Spanish-speaking children. *American Journal of Speech-Language Pathology, 10*(4), 394–406. https://doi.org/10.1044/1058-0360(2001/034)

Goldstein, B. A., & Washington, P. S. (2001). An initial investigation of phonological patterns in typically developing 4-year-old Spanish–English bilingual children. *Language, Speech, and Hearing Services in Schools, 32*(3), 153–164. https://doi.org/10.1044/0161-1461(2001/014)

Grosjean, F. (2021). *Life as a bilingual*. Cambridge University Press. https://doi.org/10.1017/9781108975490.002

Hambly, H., Wren, Y., McLeod, S., & Roulstone, S. (2013). The influence of bilingualism on speech production: A systematic review. *International Journal of Language & Communication Disorders, 48*(1), 1–24. https://doi.org/10.1111/j.1460-6984.2012.00178.x

Hoàng, T. C. (2004). *Phương ngữ học Tiếng Việt* [Vietnamese dialectology]. Đại học Quốc gia Hà Nội.

Hopf, S. C., Crowe, K., Verdon, S., Blake, H. L., & McLeod, S. (2021). Advancing workplace diversity through the Culturally Responsive Teamwork Framework. *American Journal of Speech-Language Pathology, 30*(5), 1949–1961. https://doi.org/10.1044/2021_AJSLP-20-00380

Huỳnh, C. T. (2014). *Tiếng Sài Gòn* [The Saigon dialect]. Chính trị quốc gia Sự thật.

Kirby, J. P. (2011). Vietnamese (Hanoi Vietnamese). *Journal of the International Phonetic Association, 41*(3), 381–392. https://doi.org/10.1017/S0025100311000181

Klein, R. M., Christie, J., & Parkvall, M. (2016). Does multilingualism affect the incidence of Alzheimer's disease? A worldwide analysis by country. *SSM—Population Health, 2,* 463–467. https://doi.org/10.1016/j.ssmph.2016.06.002

Kritikos, E. P. (2003). Speech-language pathologists' beliefs about language assessment of bilingual/bicultural individuals. *American Journal of Speech-Language Pathology, 12*(1), 73–91. https://doi.org/10.1044/1058-0360(2003/054)

Kroll, J. F., & Dussias, P. E. (2017). The benefits of multilingualism to the personal and professional development of residents of the U.S. *Foreign Language Annals, 50*(2), 248–259. https://doi.org/10.1111/flan.12271

Lã, T. B. L., Phan, T. H. X., & Nguyễn, T. T. N. (2011). *Tiếng Việt và tiếng Việt thực hành* [Vietnamese and Vietnamese in use]. Đại học Sư phạm Hà Nội.

Le, X. T. T., McLeod, S., & Phạm, B. (2022). Consonant accuracy and intelligibility of Southern Vietnamese children. *Speech, Language and Hearing, 25*(3), 315–324. https://doi.org/10.1080/2050571X.2021.1888195

Li, X. X., & To, C. K. S. (2017). A review of phonological development of Mandarin-speaking children. *American Journal of Speech-Language Pathology, 26*(4), 1262–1278. https://doi.org/10.1044/2017_AJSLP-16-0061

MacLeod, A. A. N., & Demers, C. (2023). Transmitting White monolingual Anglo-American norms: A concept analysis of "quality of language" in parent–child interactions. *Applied Psycholinguistics*. Advanced online publication. https://doi.org/10.1017/S014271642300005X

Margetson, K., McLeod, S., & Verdon, S. (2023a). Cross-linguistic transfer and ambient phonology: Impact on diagnosis of speech sound disorders in a longitudinal bilingual case study. *Journal of Monolingual and Bilingual Speech, 4*(3), 311–339. https://doi.org/10.1558/jmbs.23672

Margetson, K., McLeod, S., & Verdon, S. (2023b). *Diagnosing speech sound disorder in bilingual Vietnamese-English-speaking children: Are English-only assessments sufficient?* [Manuscript submitted for publication]. School of Education Charles Sturt University.

Margetson, K., McLeod, S., Verdon, S., & Tran, V. H. (2023). Transcribing multilingual children's and adults' speech. *Clinical Linguistics and Phonetics, 37*(4–6), 415–435. https://doi.org/10.1080/02699206.2022.2051073

Márquez, M. C., & Porras, A. M. (2020). Science communication in multiple languages is critical to its effectiveness in communication. *Frontiers in Communication*, *5*(31). https://doi.org/10.3389/fcomm.2020.00031

Masso, S., McLeod, S., Cronin, A., & Phạm, B. (2020). Transcription of Vietnamese adults' and children's consonants by English-speaking speech-language pathologists. *Folia Phoniatrica et Logopaedica, 72*(2), 92–107. https://doi.org/10.1159/000500245

McAlister, H., Hopf, S. C., & McLeod, S. (2023). Effect of dialect on identification and severity of speech sound disorder in Fijian children. *Speech, Language and Hearing, 26*(1), 48–60. https://doi.org/10.1080/2050571X.2022.2052506

McGregor, K. K., Williams, D., Hearst, S., & Johnson, A. C. (1997). The use of contrastive analysis in distinguishing difference from disorder: A tutorial. *American Journal of Speech-Language Pathology, 6*(3), 45–56. https://doi.org/10.1044/1058-0360.0603.45

McIntosh, B., & Dodd, B. (2008). Evaluation of core vocabulary intervention for treatment of inconsistent phonological disorder: Three treatment case studies. *Child Language Teaching and Therapy, 24*(3), 307–327. https://doi.org/10.1177/0265659007096295

McLeod, S. (2020). Intelligibility in Context Scale: Cross-linguistic use, validity, and reliability. *Speech, Language and Hearing, 23*(1), 9–16. https://doi.org/10.1080/2050571X.2020.1718837

McLeod, S., Baker, E., McCormack, J., Wren, Y., Roulstone, S., Crowe, K., Masso, S., White, P., & Howland, C. (2017). Cluster-randomized controlled trial evaluating the effectiveness of computer-assisted intervention delivered by educators for children with speech sound disorders. *Journal of Speech, Language, and Hearing Research, 60*(7), 1891–1910. https://doi.org/10.1044/2017_JSLHR-S-16-0385

McLeod, S., & Crowe, K. (2018). Children's consonant acquisition in 27 languages: A cross-linguistic review. *American Journal of Speech-Language Pathology, 27*(4), 1546–1571. https://doi.org/10.1044/2018_AJSLP-17-0100

McLeod, S., Crowe, K., & Shahaeian, A. (2015). Intelligibility in Context Scale: Normative and validation data for English-speaking preschoolers. *Language, Speech, and Hearing Services in Schools, 46*(3), 266–276. https://doi.org/10.1044/2015_LSHSS-14-0120

McLeod, S., & Harrison, L. J. (2009). Epidemiology of speech and language impairment in a nationally representative sample of 4- to 5-year-old children. *Journal of Speech, Language, and Hearing Research, 52*(5), 1213–1229. https://doi.org/10.1044/1092-4388(2009/08-0085)

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, 55*(2), 648–656. https://doi.org/10.1044/1092-4388(2011/10-0130)

McLeod, S., Harrison, L. J., Whiteford, C., & Walker, S. (2016). Multilingualism and speech-language competence in early childhood: Impact on academic and social-emotional outcomes at school. *Early Childhood Research Quarterly, 34,* 53–66. https://doi.org/10.1016/j.ecresq.2015.08.005

McLeod, S., Margetson, K., Wang, C., Tran, V. H., Verdon, S., White, K., & Phạm, B. (2022). Speech acquisition within a 3-generation Vietnamese–English family: The influence of maturation and ambient phonology. *Clinical Linguistics & Phonetics, 36*(4–5), 470–493. https://doi.org/10.1080/02699206.2020.1862915

McLeod, S., Verdon, S., Baker, E., Ball, M. J., Ballard, E., David, A. B., Bernhardt, B. M., Bérubé, D., Blumenthal, M., Bowen, C., Brosseau-Lapré, F., Bunta, F., Crowe, K., Cruz-Ferreira, M., Davis, B., Fox-Boyer, A., Gildersleeve-Neumann, C., Grech, H., Goldstein, B., . . . The International Expert Panel on Multilingual Children's Speech. (2017). Tutorial: Speech assessment for multilingual children who do not speak the same language(s) as the speech-language pathologist. *American Journal of Speech-Language Pathology, 26*(3), 691–708. https://doi.org/10.1044/2017_AJSLP-15-0161

McLeod, S., Verdon, S., Tran, V. H., Margetson, K., & Wang, C. (2022). SuperSpeech: Multilingual speech and language maintenance intervention for Vietnamese–Australian children and families via telepractice. *Language, Speech, and Hearing Services in Schools, 53*(3), 675–697. https://doi.org/10.1044/2021_LSHSS-21-00146

McLeod, S., Verdon, S., Wang, C., & Tran, V. H. (2019). Language proficiency, use, and maintenance among people with Vietnamese heritage living in Australia. *Journal of*

Monolingual and Bilingual Speech, 1(1), 55–79. https://doi.org/10.1558/jmbs.10973

Meneghini, R., & Packer, A. L. (2007). Is there science beyond English? Initiatives to increase the quality and visibility of non-English publications might help to break down language barriers in scientific communication. *EMBO Reports, 8*(2), 112–116. https://doi.org/10.1038/sj.embor.7400906

Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika, 50,* 159–179. https://doi.org/10.1007/BF02294245

Nancarrow, S., McGill, N., Baldac, S., Lewis, T., Moran, A., Harris, N., Johnson, T., & Mulcair, G. (2023). Diversity in the Australian speech-language pathology workforce: Addressing Sustainable Development Goals 3, 4, 8, and 10. *International Journal of Speech-Language Pathology, 25*(1), 119–124. https://doi.org/10.1080/17549507.2023.2165149

Nguyen, A., Shin, F., & Krashen, S. (2001). Development of the first language is not a barrier to second-language acquisition: Evidence from Vietnamese immigrants to the United States. *International Journal of Bilingual Education and Bilingualism, 4*(3), 159–164. https://doi.org/10.1080/13670050108667725

Nguyễn, T. C. (1997). *Giáo trình lịch sử ngữ âm tiếng Việt* [Textbook of Vietnamese historical phonology]. Giáo dục.

Nguyễn, V. L., & Edmondson, J. A. (1998). Tones and voice quality in modern northern Vietnamese: Instrumental case studies. *Mon–Khmer Studies, 28,* 1–18.

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education, 15*(5), 625–632. https://doi.org/10.1007/s10459-010-9222-y

Oh, J. S., & Fuligni, A. J. (2010). The role of heritage language development in the ethnic identity and family relationships of adolescents from immigrant backgrounds. *Social Development, 19*(1), 202–220. https://doi.org/10.1111/j.1467-9507.2008.00530.x

Paradis, J., Emmerzael, K., & Duncan, T. S. (2010). Assessment of English language learners: Using parent report on first language development. *Journal of Communication Disorders, 43*(6), 474–497. https://doi.org/10.1016/j.jcomdis.2010.01.002

Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition, 18*(1), 1–25. https://doi.org/10.1017/S0272263100014662

Peña, E. D., Gutiérrez-Clellen, V. F., Iglesias, A., Goldstein, B. A., & Bedore, L. M. (2018). *Inventory to Assess Language Knowledge (ITALK)*. Brookes.

Peña, E. D., Spaulding, T. J., & Plante, E. (2006). The composition of normative groups and diagnostic decision making: Shooting ourselves in the foot. *American Journal of Speech-Language Pathology, 15*(3), 247–254. https://doi.org/10.1044/1058-0360(2006/023)

Pham, A. H. (2003). *Vietnamese tones: A new analysis*. Routledge.

Phạm, B., Le, X. T. T., & McLeod, S. (2016). *Vietnamese Speech Assessment: Research version*. Authors.

Phạm, B., & McLeod, S. (2016). Consonants, vowels, and tones across Vietnamese dialects. *International Journal of Speech-Language Pathology, 18*(2), 122–134. https://doi.org/10.3109/17549507.2015.1101162

Phạm, B., & McLeod, S. (2019). Vietnamese-speaking children's acquisition of consonants, semivowels, vowels, and tones in Northern Viet Nam. *Journal of Speech, Language, and Hearing Research, 62*(8), 2645–2670. https://doi.org/10.1044/2019_JSLHR-S-17-0405

Phạm, B., McLeod, S., & Harrison, L. J. (2017). Validation and norming of the Intelligibility in Context Scale in Northern Viet Nam. *Clinical Linguistics & Phonetics, 31*(7–9), 665–681. https://doi.org/10.1080/02699206.2017.1306110

Phạm, B., McLeod, S., & Le, X. T. T. (2016). Development of the Vietnamese speech assessment. *Journal of Clinical Practice in Speech-Language Pathology, 18*(3), 126–130.

Poole, I. (1934). The genetic development of consonant sounds in the speech of children. *The Elementary English Review, 11,* 159–161.

Ruiz-Felter, R., Cooperson, S. J., Bedore, L. M., & Peña, E. D. (2016). Influence of current input–output and age of first exposure on phonological acquisition in early bilingual Spanish–English-speaking kindergarteners. *International Journal of Language & Communication Disorders, 51*(4), 368–383. https://doi.org/10.1111/1460-6984.12214

Ryan, C. (2013). *Language use in the United States: 2011*. https://www.census.gov/library/publications/2013/acs/acs-22.html

Sabri, M., & Fabiano-Smith, L. (2018). Phonological development in a bilingual Arabic-English–speaking child with bilateral cochlear implants: A longitudinal case study. *American Journal of Speech-Language Pathology, 27*(4), 1506–1522. https://doi.org/10.1044/2018_AJSLP-17-0162

Seifert, M., Morgan, L., Gibbin, S., & Wren, Y. (2020). An alternative approach to measuring reliability of transcription in children's speech samples: Extending the concept of near functional equivalence. *Folia Phoniatrica et Logopaedica, 72*(2), 84–91. https://doi.org/10.1159/000502324

Shriberg, L. D., & Lof, G. L. (1991). Reliability studies in broad and narrow phonetic transcription. *Clinical Linguistics & Phonetics, 5*(3), 225–279. https://doi.org/10.3109/02699209108986113

Statistics Canada. (2012). *Linguistic characteristics of Canadians: Language, 2011 census of population (98-314-X2011001)*. Minister of Industry, Statistics Canada. http://www12.statcan.gc.ca/census-recensement/2011/as-sa/98-314-x/98-314-x2011001-eng.cfm

Stenius, K., Kerr-Corrêa, F., Obot, I., Furtado, E. F., Lima, M. C. P., & Babor, T. F. (2017). Beyond the Anglo-American world: Advice for researchers from developing and non–English-speaking countries. In T. F. Babor, K. Stenius, R. Pates, M. Miovský, J. O'Reilly, & P. Candon (Eds.), *Publishing addiction science: A guide for the perplexed* (pp. 71–88). Ubiquity Press. https://doi.org/10.5334/bbd.d

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2022). Feature generalization in Dutch–German bilingual and monolingual children's speech production. *First Language, 42*(1), 101–123. https://doi.org/10.1177/01427237211058937

Stow, C., & Dodd, B. (2005). A survey of bilingual children referred for investigation of communication disorders: A comparison with monolingual children referred in one area in England. *Journal of Multilingual Communication Disorders, 3*(1), 1–23. https://doi.org/10.1080/14769670400009959

Sullivan, G. M., & Artino, A. R., Jr. (2013). Analyzing and interpreting data from Likert-type scales. *Journal of Graduate Medical Education, 5*(4), 541–542. https://doi.org/10.4300/JGME-5-4-18

Tang, G., & Barlow, J. (2006). Characteristics of the sound systems of monolingual Vietnamese-speaking children with phonological impairment. *Clinical Linguistics & Phonetics, 20*(6), 423–445. https://doi.org/10.1080/02699200500100910

To, C. K. S., McLeod, S., & Cheung, P. S. P. (2015). Phonetic variations and sound changes in Hong Kong Cantonese: Diachronic review, synchronic study and implications for speech sound assessment. *Clinical Linguistics & Phonetics, 29*(5), 333–353. https://doi.org/10.3109/02699206.2014.1003329

Toohill, B. J., McLeod, S., & McCormack, J. (2012). Effect of dialect on identification and severity of speech impairment in Indigenous Australian children. *Clinical Linguistics & Phonetics, 26*(2), 101–119. https://doi.org/10.3109/02699206.2011.595523

Tran, V. H., McLeod, S., Verdon, S., & Wang, C. (2021). Vietnamese–Australian parents: Factors associated with language use and attitudes towards home language maintenance. *Journal of Multilingual and Multicultural Development*. Advance online publication. https://doi.org/10.1080/01434632.2021.1904963

Tran, V. H., Verdon, S., & McLeod, S. (2022). Consistent and persistent: Successful home language maintenance among Vietnamese–Australian families. *Journal of Home Language Research, 5*(1), 5. https://doi.org/10.16993/jhlr.43

Tran, V. H., Verdon, S., McLeod, S., & Wang, C. (2022). Family language policies of Vietnamese–Australian families. *Journal of Child Science, 12*(1), e67–e78. https://doi.org/10.1055/s-0042-1743490

Tran, V. H., Wang, C., McLeod, S., & Verdon, S. (2021). Vietnamese–Australian children's language proficiency and use. *International Journal of Bilingualism, 25*(6), 1735–1763. https://doi.org/10.1177/13670069211034587

Verdon, S., McLeod, S., & McDonald, S. (2014). A geographical analysis of speech-language pathology services to support multilingual children. *International Journal of Speech-Language Pathology, 16*(3), 304–316. https://doi.org/10.3109/17549507.2013.868036

Verdon, S., McLeod, S., Tran, V., & Margetson, K. (2021). *An ecological systems model of home language maintenance in children* [Manuscript in preparation]. School of Allied Health, Exercise and Sports Science, Charles Sturt University.

Verdon, S., McLeod, S., & Winsler, A. (2014). Language maintenance and loss in a population study of young Australian children. *Early Childhood Research Quarterly, 29*(2), 168–181. https://doi.org/10.1016/j.ecresq.2013.12.003

Verdon, S., Tran, V., McLeod, S., Margetson, K., & Wang, C. (2021). *Parents' perceptions of a bilingual speech and language maintenance support program for Vietnamese–Australian children and families via telepractice* [Manuscript in preparation]. School of Allied Health, Exercise and Sports Science, Charles Sturt University.

Wang, C., Verdon, S., McLeod, S., & Tran, V. H. (2021). Profiles of linguistic multicompetence in Vietnamese–English speakers. *American Journal of Speech-Language Pathology, 30*(4), 1711–1727. https://doi.org/10.1044/2021_AJSLP-20-00296

Williams, C. J., & McLeod, S. (2012). Speech-language pathologists' assessment and intervention practices with multilingual children. *International Journal of Speech-Language Pathology, 14*(3), 292–305. https://doi.org/10.3109/17549507.2011.636071

Appendix

VietSpeech Multilingual Assessment Protocol

The VietSpeech Multilingual Assessment Protocol involves

(a) assessing and analyzing both/all languages spoken by the child within the child's cultural community,

(b) considering ambient phonology by assessing children and adult family members in both/all languages,

(c) defining accuracy using both the standard form of the language (e.g., percentage of consonants correct–standard [PCC-S]) and by accepting dialectal variants of the language (e.g., percentage of consonants correct–dialect [PCC-D]), and

(d) describing and analyzing the child's speech based on proficiency in both/all languages.