Bilingual phonological development across generations: Segmental accuracy and error patterns in second- and third-generation British Bengali children

Abstract

Introduction: While developmental norms for speech sound development have been widely reported for monolingual children, and increasingly for bilingual children, little is known about speech sound development across different generations of children growing up in heritage language settings. The purpose of the present study was to gain a better understanding of inter-generational differences in the phonological development of British Bengali children.

Methods: Typically-developing second-generation and third-generation Bengali heritage children living in Wales (n=19), aged between 4 and 5 years, participated in a picture-naming task in Sylheti and English. The single-word speech samples were transcribed phonetically and analyzed in terms of consonant and vowel accuracy measures, and error patterns. Subsequently, logistic mixed-effects regression models were fitted to identify the factors that predict accurate speech patterns in the children's productions.

Results:

The results revealed high levels of accuracy in consonant and vowel production by both sets of children, particularly in English. On Sylheti consonants, second-generation children significantly outperformed third-generation children, however only on language-specific sounds. In contrast, generation was not a significant predictor for accuracy on English consonants, but all children performed better on shared sounds than on English-specific categories, and on stops than affricates. The third-generation children exhibited a greater number of error types in Sylheti than the second-generation children, and more common replacement of Sylheti dental stops with alveolars.

Conclusion: The results suggest that third-generation children have less developed pronunciation patterns in the heritage language, but not the majority language, than their age-matched second-generation peers, however only on language-specific sounds. These findings indicate that differentiating between the phonological norms of monolingual and bilingual children may not be clinically sufficiently sensitive, at least in the minority language, and that more fine-grained language use variables, such as the generation to which a bilingual child belongs, need to be considered.

Keywords: bilingual phonological acquisition; heritage language; Sylheti-English bilingualism; intergenerational transmission; phonological process analysis

1. Introduction

While knowledge of bilingual phonological development in heritage language settings has increased significantly in recent years, at least for common immigrant groups, evidence on children speaking less common languages is lacking. For instance, the last ten years have witnessed a burgeoning of research on speech sound development in mid- to large-size samples of Spanish/English bilingual children in the US (see Author 4, Author 1 & other, 2018; Cooperson, Bedore & Peña, 2013; Fabiano-Smith & Goldstein, 2010; Goldstein & Bunta, 2012; Keffala, Scarpino, Scheffner Hammer, Rodriguez, Lopez & Goldstein, 2020; Gildersleeve-Neumann, Kester, Davis, & Peña, 2008; Gildersleeve-Neumann, Peña, Davis & Kester, 2009; Ruiz-Felter, Cooperson, Bedore & Peña, 2016; Scarpino, Scheffner Hammer, Goldstein, Rodriguez & Lopez, 2019, among others). These studies, which have included cross-sectional and longitudinal studies as well as comparisons with age-matched monolinguals, have much aided speech-language pathologists differentiate typical from atypical development in bilingual children. However, research on bilingual phonological development in children speaking less common languages has lagged behind, especially in the context of smaller immigrant communities. Even more lacking have been studies that have examined differences in speech development across different generations of children raised with a heritage language. The present study aims to address this gap in the literature by presenting consonant and vowel accuracy measures and error patterns in the heritage language and the host language of typically developing second- and third-generation British Bengali children raised in a small bilingual community in Wales.

1.1.Phonological development in heritage language settings

Recent years have seen an increased interest in the study of heritage languages and their acquisition (e.g. Kupisch & Rothman, 2018; Montrul, 2016; Polinsky, 2018).

Heritage language speakers are bilinguals who acquire a minority language in a majority language environment (Lohndal, Rothman, Kupisch & Westergaard, 2019). While their dominance patterns may be variable in early developmental stages, by the time they reach adulthood, they are virtually always dominant in the majority language (Lohndal et al., 2019). The status of their heritage language, in turn, has been widely debated. Some have maintained that the grammars of heritage languages are inherently incomplete (Benmamoun, Montrul & Polinsky, 2013; Montrul, 2016; Polinsky, 2018), while others have argued that they constitute coherent grammars in their own right with different developmental paths and outcomes as those of monolinguals (Bayram, Kupisch, Pascual y Cabo & Rothman, 2019; Kupisch & Rothman, 2018).

In terms of their pronunciation patterns, heritage language speakers have been shown to outperform L2 learners (Amengual, 2019; Chang, Yao, Haynes & Rhodes, 2011; Kupisch et al., 2014; Oh, Jun, Knightly & Au, 2003). Thus, Chang et al. (2011) demonstrated that heritage Mandarin speakers in the United States were more accurate in their production of cross-linguistically similar stops, fricatives and vowels than L1 American English learners of Mandarin. Similarly, Kupisch et al. (2014) showed that heritage language speakers from Germany, France and Italy were perceived to be less foreign-accented in their minority language than L2 learners.

These findings have been explained on the basis of differences in early linguistic experience, with heritage language speakers typically having exposure from birth, or shortly thereafter. During this period, monolingual and bilingual children start out as language-general perceivers, but subsequently become highly attuned to the regularities of their native language(s) (Bosch & Sebastián-Gallés, 2003; Fennell, Sin-Mei Tsui, & Hudon, 2016; Werker & Tees, 1984). There is a consensus that linguistic experience within the first year of life is particularly influential for subsequent development. For

example, international adoptees without subsequent exposure to their birth language were able to retain some residual knowledge of L1 speech patterns (Hyltenstam, Bylund, Abrahamsson, & Park, 2009; Pierce, Chen, Delcenserie, Genesee, & Klein, 2015). Moreover, they were able to access it during re-exposure and outperform individuals without that experience (Choi, Broersma, & Cutler, 2017).

Nevertheless, early experience does not automatically lead to native-like pronunciation patterns. Indeed, a substantial number of studies has shown differences in the minority language of heritage language speakers and monolinguals (Kupisch et al., 2014; Author 1 & Author 2, 2018; McCarthy, Evans & Mahon, 2013; Oh et al., 2003). For example, McCarthy et al. (2013) showed that Sylheti heritage speakers in London produced Sylheti vowels much like monolingual controls, but differed from them on Sylheti stops. Author 1 and Author 2 (2018) similarly found that while second-generation female heritage speakers of Sylheti produced voiced stops that were similar to those of first-generation female immigrants who had grown up in Bangladesh, they realized Sylheti coronals as alveolars and unaspirated voiced velars, showing transfer of English properties to the Sylheti sound system. Differences between the productions of heritage language and monolingual speakers may also be perceptually salient. Indeed, Kupisch et al. (2014) showed that, despite exposure to both languages from birth, the heritage language speakers in their study were perceived to have a foreign accent in their heritage languages.

On the contrary, heritage language speakers tend to display monolingual-like or native/near-native productions in the majority language (Evans, Mistry & Moreiras, 2007; Author 1 & Author 2, 2018; McCarthy et al., 2013). For instance, McCarthy et al. (2013) found that second-generation Sylheti heritage speakers in London produced English vowels and stops that were indistinguishable from those of monolingual

English speakers. Similarly, the second-generation female heritage speakers of Sylheti in Author 1 and Author 2 (2018) produced all English stops at the correct place of articulation and were hence more accurate than the first-generation female immigrants who had grown up in Bangladesh. Kirkham and McCarthy (2020) also reported successful acquisition of the English allophonic lateral contrast by Sylheti/English bilingual children, in line with previous findings that show monolingual-like acquisition of allophony by heritage language speakers in other contexts (Barlow, Branson, & Nip, 2013; Burrows, Jarmulowicz, & Oller, 2019).

The amount of cumulative early experience in the heritage language might explain why some heritage speakers develop native-like pronunciation in the heritage language but others do not. For instance, Amengual (2019) found that only sequential Spanish/English bilinguals – those who heard exclusively Spanish at home and learned English later in school – produced native-like spirantization patterns in Spanish as adults. On the other hand, matched heritage Spanish speakers who were simultaneous Spanish/English bilinguals failed to produce native-like spirants, even though they had been exposed to Spanish from birth. These findings suggest that there may be a *threshold* of early linguistic experience that allows for native or near native heritage language pronunciation. It is also possible that contact with – and hence transfer from – the societal language early on overrides a speaker's ability to develop a native accent in a language that will be heard and used less over time.

Indeed, it might be regular and sustained use of the heritage language across at least childhood that may be necessary for the development of a native accent, especially after schooling in the societal language begins. For instance, Oh et al. (2003) found that heritage speakers who had used Korean in early childhood but stopped using it after school entry at age 5 were significantly less accurate in their production of Korean stops

and were perceived to sound less native than native Korean speakers. Likewise, Mora and Nadeu (2012) showed that Catalan-Spanish bilinguals from predominantly Catalan backgrounds who started to be exposed to Spanish at school age failed to display native Catalan features in their speech if they used Spanish frequently. However, this was not the case for speakers whose daily use of Spanish was low. Taken together, these results suggest that early and extensive experience in the heritage language, possibly coupled with limited early contact with the societal language and sustained and regular use of the heritage language throughout childhood, might be prerequisites for the successful acquisition of a native accent in the minority language. These experiences might then also interact with learner-internal characteristics such as sound learning (i.e. pronunciation-specific) aptitude (Saito, 2017; Saito, Sun & Tierney, 2018) and learner-external sociolinguistic factors, such as variety status (Author 1 et al., 2019), to account for different degrees of native features in the speech of heritage language speakers.

1.2.Inter-generational differences in pronunciation patterns

Studies in heritage language settings have primarily compared the language skills of first-generation immigrants to those of second-generation heritage language speakers, but rarely examined language maintenance and shift in the third or subsequent generations. This is because seminal work showed that language shift typically occurs within three generations, with the first generation often being dominant if not monolingual in the native (immigrant) language, the second generation becoming bilingual by virtue of learning the heritage language in the home and the societal language outside of the home, and the third generation virtually becoming monolingual in the societal language (Fishman's (1964) "three-generation model"). However, recent work in contexts where language may be highly indexical of ethnic identity has shown

that the third generation may still learn (and transmit) the heritage language to various extents (Anderson-Mejías, 2005; Sevinç, 2016; Author 1 & Author 2, 2018), making third-generation heritage language speakers particularly interesting in the study of intergenerational language maintenance.

Indeed, in the context of Turkish immigrants in the Netherlands, Sevinç (2016) found that the third-generation heritage language speakers did learn and speak Turkish. However, despite their commitment to retaining Turkish language and identity, their language history, their self-rated language proficiency, and current language practices differed significantly from those of first- and second-generation bilinguals, suggesting a possible outset of intergenerational language shift. In particular, the third-generation bilinguals in this study felt Turkish and valued the heritage language; yet, they preferred Dutch when responding to their parents and made no specific efforts to practice and maintain Turkish. Thus, ethnic affiliation did not ensure heritage language maintenance. Similarly, Al-Azami (2014), in a case study of language shift in a British-Bengali child, found that, despite feeling Bengali, the child literally replaced Bengali with English once she came in contact with this language, suggesting that, although important, ethnic affiliation does not guarantee heritage language maintenance, especially when parental attitudes do not discourage societal language use in the home.

Few studies exist on intergenerational differences in pronunciation patterns. Nagy and Kochetov (2013) examined the degree of sociolinguistic variation induced by language contact by exploring the VOT patterns of voiceless stops across three generations of heritage language speakers: immigrants who were born and raised in either Russia, Ukraine or Italy but had lived in Toronto for at least 20 years (first generation); speakers who were the children of immigrants but were born in Toronto or had arrived before age 6 (second generation); and speakers who had at least one parent

who was a second generation heritage language speaker (third generation). The authors found a gradual lengthening of VOT patterns in Russian and Ukrainian across each successive generation, suggesting a drift towards English longer-lag VOT values. Indeed, the third generation had VOT values in the heritage language that were equivalent to English VOT values for English monolinguals. Nagy and Kochetov (2013) also found that more limited heritage language use and lower ethnic orientation were related to longer (i.e. English-like) VOT values, suggesting that the language practices and ethnic affiliation of third-generation heritage language speakers are different from those of the second generation and show gradual drift towards the societal language. Interestingly, however, this pattern was not found for Italian heritage language speakers possibly due to the long-established presence of the Italian community in Toronto as well as the educational opportunities offered by the city to maintain the language.

Author 1 and Author 2 (2018) also found inter-generational differences in the pronunciation patterns of Bangladeshi heritage families. Specifically, the authors found that the Sylheti stops produced by Sylheti-English bilinguals exhibited incremental changes towards English-like stops across successive generations, with the third-generation children's productions showing the greatest influence from English. The authors interpreted this finding as dependent upon linguistic experience: the third-generation children came from primarily English-speaking homes and lived in ethnically heterogeneous areas, factors that significantly decreased the amount of Sylheti they heard in their immediate environment. Moreover, the third-generation children were also likely to hear Sylheti input that contained non-native features, which, in the context of reduced input, could ultimately be conducive to a non-native accent. In Author 1 and Author 2's (2018, p. 278) words: "since the GEN 3 CHILDREN have

restricted exposure to Sylheti, their mothers' non-native productions may be partly responsible for their own non-native realizations in the heritage language." Taken together, these results suggest that third-generation heritage language speakers may be faced with both psychological (i.e. identity-related) and environmental (i.e. input-related) challenges that limit their likelihood to attain a native accent in the heritage language.

1.3.The present study

The purpose of the present study was to investigate the phonological development of second- and third-generation British Bengali children growing up in Cardiff, South Wales. In the latest census, 451,529 individuals in the United Kingdom considered their ethnicity to be Bengali (Office for National Statistics [ONS], 2011), with a heavy concentration in large urban areas, such as Birmingham or London, notably Camden and Tower Hamlets (cf. Blackledge et al., 2008; McCarthy et al., 2013; McCarthy, Mahon, Rosen & Evans, 2014, for details). In contrast, the Bengali community in Wales is comparatively small, with only some 0.3% of the population (i.e. 10,687 individuals) considering themselves British Bangladeshis (Office for National Statistics [ONS], 2011). Of these, 5,207 individuals indicated Sylheti Bengali as their main home language with approximately half of them residing in Cardiff, mostly in the inner-city wards of Riverside and Grangetown. These communities have a close-knit social structure with shops, restaurants and community centers.

The overwhelming majority of the Bengali population in the UK originates from the rural province of Sylhet in north-eastern Bangladesh (Chalmers & Miah, 1996), where Sylheti is spoken. Sylheti is an Indo-Aryan language that is closely related to Standard Bengali (SB), but the two languages are not mutually intelligible (Chalmers & Miah, 1996; Chatterji, 1926; Hamid, 2007). Unlike Sylheti speakers in Bangladesh,

British-born heritage speakers are not usually competent in SB (Chalmers & Miah, 1996) but they tend to ascribe greater prestige to it since it constitutes the language of formal education in Bangladesh (Blackledge et al., 2008).

Overall, the phonological system of Sylheti is less complex than that of SB with fewer consonant and vowel categories (Chalmers & Miah, 1996; Gope, 2018; Gope & Mahanta, 2015; Khan, 2010). Thus, compared with SB, Sylheti has a reduced obstruent system characterized by de-aspiration, spirantization and de-affrication (Gope, 2018; Gope & Mahanta, 2015). Nevertheless, as Table 1 shows, the language distinguishes 9 stops, including the breathy voiced categories /bʰ/ and /gʰ/, 2 affricates, 3 nasals, 6 fricatives as well as /r/, /j/ and /l/. Its vowel system, in turn, encompasses the 7 monophthongs /i e ɛ ɑ o ɔ u/, as well as a range of diphthongs. The language also allows a number of heterosyllabic vowel sequences (Chalmers & Miah, 1996; Khan, 2010; McCarthy et al., 2013).

Table 1. Consonant inventory: Sylheti and English

	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Retroflex	Palatal	Velar	Glottal
Stop	**b** *p* bʰ		t̪ d̪	*t* *d*		ʈ ɖ		**k** **g** gʰ	
Fricative		**f** *v*	*θ* *ð*	**s** **z**	**ʃ** *ʒ*			x	**h**
Affricate					*tʃ* *dʒ*		cʃ ɟʒ		
Nasal	**m**			**n**				**ŋ**	
Trill				r					
Approximant	*w**				ɹ		**j**		
Lateral Approximant				**l**					

Key: *labio-velar approximant; Sylheti only = standard font; English only = *italics*; shared = **bold**.

Note 1: Historically, Sylheti also contains the stops /p/, /pʰ/ and /kʰ/. However, Sylheti speakers have been shown to realize them consistently as fricatives, i.e. [f] (or [ɸ]) and [x] (Author 1 & Author 2, 2018; Gope, 2018; Gope & Mahanta, 2015; McCarthy et al., 2013).

Note 2: While the English coronal stops /t/ and /d/ are typically realized at the alveolar place of articulation, there is a fair amount of individual variation, with some speakers realizing them as dentals (Dart, 1998). We are grateful to an anonymous reviewer for drawing our attention to this issue.

A small number of studies have examined the speech of Sylheti speakers in the United Kingdom. In McCarthy et al.'s (2013) study of three groups of adult Sylheti

speakers from the London Bengali community, i.e. first-generation Bangladeshis who arrived in the UK in adulthood (late arrivals), first-generation Bangladeshis who arrived in early childhood (early arrivals), and second-generation heritage speakers, Sylheti monophthongs were produced natively by speakers in all three groups, but only the late arrivals also produced Sylheti stops natively. Author 1 and Author 2 (2018), in turn, found that only the first-generation of Sylheti-speaking mothers living in Cardiff – but not the second-generation – produced Sylheti stops in line with native Sylheti patterns. Moreover, both groups exhibited an influence from English with Sylheti voiceless categories realized with long-lag, rather than short-lag VOT patterns. The authors ascribed the different findings in the two communities to differences in social structure. In particular, in contrast to the Bengali community in Cardiff, the London communities are characterized by a steady influx of new arrivals from Bangladesh, which, in turn, supports the maintenance of homeland norms.

No studies are available from Sylheti-speaking children in Asia and only three studies have investigated the speech of Bangladeshi heritage children in the UK. The first, McCarthy et al. (2014), examined the perception and production of English stops in children from the Bengali community in London. This study revealed changes during the first school year in the direction of their monolingual peers' patterns. However, Sylheti stops were not examined. The only studies to date that have examined children's productions in Sylheti and English are Kirkham and McCarthy (2020), and Author 1 and Author 2 (2018). Kirkham and McCarthy (2020) investigated the productions of laterals by Sylheti-English bilingual children and found that the children produced monolingual-like Sylheti laterals (i.e. clear laterals) in all positions, although English laterals showed transfer of phonetic detail from Sylheti. Author 1 and Author 2 (2018), in turn, found that third-generation heritage speakers, aged around 4 years, were

significantly less accurate in their Sylheti stop productions than age-matched second-generation children, although both sets of children were native-like in their production of English stops. While this study provides important insights into inter-generational differences in speech production, it is limited to a small number of consonantal categories. In order to be of clinical value, it is essential to have a broader understanding of the speech patterns of heritage language children across different generations. The current study aims to address this need on the basis of consonant and vowel accuracy measures and error patterns from two generations of British Bengali children. Specifically, it sought to answer the following research questions:

(RQ1) Do age-matched second- and third-generation British Bengali children differ in the accuracy with which they produce consonants and vowels in the heritage language, i.e. Sylheti, and the host language, i.e. English?

(RQ2) To what extent is the accuracy with which second- and third-generation British Bengali children produce consonants in Sylheti and English predicted by (a) overlap in the sound inventories of the two languages, (b) word position, and (c) manner class?

(RQ3) Do age-matched second- and third-generation British Bengali children differ in the number and type of error patterns they produce?

With respect to (RQ1), we predicted that second-generation children would outperform third-generation children on consonant production in the heritage language, but not the host language, based on evidence from the few existing studies on the speech patterns of Sylheti-English bilingual children in the UK (Author 1 & Author 2, 2018; Kirkham & McCarthy, 2020; McCarthy et al., 2014). Previous work on vowel productions in bilingual preschoolers suggests high degrees of accuracy in the majority and minority language (Gildersleeve-Neumann et al., 2008; Goldstein & Bunta, 2012;

Goldstein et al., 2010). However, no prior data are available on vowel productions from British Bengali children.

With respect to (RQ2), we hypothesized that cross-linguistically shared sounds would be produced more accurately than language-specific categories based on prior work suggesting enhanced cue strength (Author 4, Author 1 & other, 2018; Fabiano-Smith & Goldstein, 2010), in line with MacWhinney's *Unified Competition Model* (MacWhinney, 2005). The predictions for word position and manner were less equivocal. Thus, while a number of studies have shown that children struggle more with word-final consonants (e.g., Smit, 1993), some more recent work has shown greater accuracy on word-final than word-initial sounds (e.g. McLeod & Masso, 2019). Similarly, while some types of sounds are generally earlier acquired, e.g. stops and nasals, and others later, e.g. rhotics, the evidence from previous studies on heritage language children did not allow any clear predictions to be made (Author 4, Author 1 & other, 2018; Ruiz-Felter et al., 2016).

Finally, with respect to (RQ3), we predicted broadly similar types of error patterns in both sets of children as they were of the same age and had been exposed to the same two languages. However, since the third-generation children had received less input in the heritage language, we hypothesized that early error patterns may be more likely to persist than in the second-generation children.

2. Methods

2.1. Participants

The present study included a total of 19 typically developing Sylheti-English bilingual children: nine second-generation Bengali heritage children (GENERATION 2), aged 3;8-5;0, and ten third-generation Bengali heritage children (GENERATION 3),

aged 3;8-4;11 (see Table 2 for details). All were born in Cardiff, United Kingdom, and have been living in the same location throughout life. Generational status was assigned to children based on whether both their parents had spent their formative years in Bangladesh before moving to the United Kingdom (GENERATION 2 children) or whether both their parents had been raised entirely in Sylheti-speaking homes in the United Kingdom with English as their predominant language following mainstream education (GENERATION 3 children). Demographic details were established on the basis of a detailed language background questionnaire completed by the children's mothers.

Table 2. Participant background information

	GENERATION 2	**GENERATION 3**	**Difference**
Number of participants	9	10	-
Gender	5 m; 4 f	7 m; 3 f	-
Age in months	54 (SD: 6.5)	51 (SD: 6.3)	t (17)= 1.24, p= .232
***Mean % Sylheti at home**	49 (SD: 11.8)	5 (SD: 7.6)	t (17)= 9.945, p< .001
***Mean % Sylheti in community**	25 (SD: 5)	9 (SD: 6.7)	t (17)= 6.032, p< .001
***Mean % Sylheti media**	0 (SD: 0)	0 (SD: 0)	-

*Estimated language use patterns; mothers reported for their children.

The GENERATION 2 children are almost exclusively addressed in Sylheti by both their Bangladesh-born parents, who arrived in the United Kingdom when they were in their late teens or in early adulthood, with only occasional use of English in the presence of non-Sylheti-speaking visitors. All children also have older siblings who they predominantly address in Sylheti when their parents are present, but use Sylheti and English as well as code-switching when they are absent. Since most of the GENERATION 2 children live in multigenerational homes, they also receive consistent Sylheti input from other native Sylheti-speaking adults, notably their grandparents. As they live in inner-city areas with a high concentration of Bangladeshi heritage speakers, outside of the home they are additionally exposed to Sylheti in a range of different settings as part of their everyday life, such as the local grocer, the mosque or neighbors.

On the other hand, they use both Sylheti and English with friends in the local community. None of these children attend day care settings as their mothers are all stay-at-home parents. Based on maternal reports, the GENERATION 2 children use Sylheti on average 49% of the time in the home and 25% of the time in the local community, but never consume Sylheti language media. They spend an average of one week per year in Bangladesh to visit relatives.

The GENERATION 3 children, in contrast, live in ethnically heterogeneous areas of the city, and thus have fewer opportunities to use Sylheti in the community. They receive regular Sylheti input via interactions with their grandparents and other Sylheti speakers in the community, for example whilst attending the local mosque. However, input in the home is largely in English. Thus, while their second-generation parents (both mothers and fathers) are competent Sylheti speakers and regularly use the language to communicate with their parents and community members, the children's dominant language is English since they were born, raised and educated in the United Kingdom. Moreover, the GENERATION 3 children all attend an English-medium pre-school, and interactions between siblings are virtually always in English. Thus, based on maternal reports, these children use Sylheti 5% of the time in the home on average and 9% of the time in the community, and hence they hear this language significantly less than their age-matched second-generation counterparts. Moreover, they never travel to Bangladesh for family visits. Nevertheless, both sets of children have functional productive skills in Sylheti and were able to sustain a conversation in the language with the second author, a fluent Sylheti-English bilingual, who collected the data. No formal test of their proficiency in Sylheti was carried out, however. Note also that none of the children had any speech, language or communication difficulties, and all had normal hearing. The mothers of all participants gave their informed consent for inclusion of

their children before the study commenced; the children themselves also indicated their assent. The research reported in this manuscript was reviewed and approved by the XXX ethics committee (ethics reference number: XXX).

2.2. Materials

The children's phonological development in Sylheti and English was assessed in a picture-naming task, designed specifically for the purposes of this study, which encompassed a total of 122 Sylheti words and 125 English ones (see Appendix for full details). All were monosyllabic or bisyllabic, with the exception of one trisyllabic word in each language. The target items were words that were expected to be familiar to young children in this community and that could be elicited via pictorial representation. A few loanwords from English were included in the Sylheti set since they are fully integrated into the language. Both sets of words include multiple instances of all singleton consonants and vowels that occur in the two languages. Moreover, where phonotactically admissible, word-initial and word-final consonants were systematically distinguished.

2.3. Procedure

Data collection took place in individual sessions in a quiet room in the participants' homes in the presence of an adult family member. Each participant was seen twice, once in a Sylheti session, and once in an English one, with the two sessions scheduled on different days so as to set them into a monolingual language mode (Grosjean, 2001). Recordings were made using a Zoom H2 Handy Recorder with integrated condenser microphone, positioned around 30 centimeters from the participant's mouth. Each session commenced with a brief interaction between the participant and the second author, a fluent Sylheti-English bilingual. Following

familiarization with the task, the children were asked to name the pictures depicted in random order in a bound book. Where items could not be elicited spontaneously, the children were given phonemic or semantic prompts, and if these were unsuccessful, the words were modelled by the experimenter, as in previous studies (Author 1 et al., 2014, 2015; Munro, Ball, Müller, Duckworth & Lyddy, 2005). This procedure follows the methodology of many recent studies (e.g. Author 4, Author 1 & Other, 2018; Keffala et al. 2020; Ruiz-Felter et al. 2016, among others) and is based on evidence showing that differences in accuracy in children's speech productions across imitated and spontaneously produced items either do not occur or are negligible (e.g., Goldstein, Fabiano & Iglesias, 2004; Siegel, Winitz & Conkey, 1963; Smith & Ainsworth, 1967). This yielded a total of 122 (words) × 19 (children) = 2318 Sylheti words, and 125 (words) × 19 (children) = 2375 English words. Of these, 24 Sylheti words (i.e. 1.04%) and 45 English words (i.e. 1.89%) were discarded due to poor recording quality/overlapping background noise, yielding a total of 4624 words for analysis. On average, the recording sessions lasted 25 minutes.

2.4. Analysis

A whole-word approach to transcription was adopted. All data were transcribed by a phonetically trained transcriber who is an early Sylheti-English bilingual, using the symbols of the International Phonetic Alphabet (IPA) (International Phonetic Association, 2005). Specifically, in addition to base symbols, diacritics were used, notably those critical for the distinction of language-specific categories, such as [ː], [ʰ], [◌̥] and [◌̤]. However, not all subphonemic details were transcribed. For example, we did not differentiate degrees of openness or lip rounding in vowels. Moreover, no symbols from the extensions to the IPA (ExtIPA) chart (Ball, Howard & Miller, 2018) were used. This approach is in line with many previous studies on children's speech

development (e.g. Author 4, Author 1 & Other, 2018; Fabiano-Smith & Goldstein, 2010; Kehoe & Girardier, 2020; Munro et al., 2005).

Upon initial transcription, 20% of the items were randomly selected from the pool of productions and independently reanalyzed by a second phonetically trained transcriber. Mean inter-transcriber reliability reached 95.3%, ranging from 94.09% for the Sylheti tokens to 96.33% for the English tokens. Any differences between the two transcribers were resolved by consensus.

As in many previous studies of bilingual children's phonological development (e.g., Author 4, Author 1 & Other, 2018; Goldstein & Bunta, 2012; Ruiz-Felter et al., 2016; Scarpino, 2011) we calculated a range of accuracy measures for descriptive purposes, before running inferential statistical analyses. This included overall consonant accuracy in terms of Percent of Consonants Correct-Revised (PCC-R) (Shriberg, Austin, Lewis, McSweeney, & Wilson, 1997). PCC-R specifies the percentage of correctly produced consonants out of the total number of consonants targeted. However, the measure disregards developmental speech sound distortions and productions that differ from adults' in fine phonetic detail, such as the degree of aspiration in initial voiceless plosives. In addition, we calculated PCC-R separately for each manner class (e.g., plosives, fricatives, affricates, etc.) and for word-initial and word-final consonants. Moreover, since past studies of bilingual speech sound development have found differential accuracy for sounds shared or "phonetically similar" (Flege, 1981) between languages and language-specific (i.e. unique) sounds (Author 4, Author 1 & Other, 2018; Fabiano-Smith & Goldstein, 2010), we calculated PCC-R for sounds that occur in both Sylheti and English (/b k ɡ m n ŋ f s z ʃ h j l/) as well as for unshared consonants, i.e. those only occurring in English (/p t d tʃ dʒ v θ ð ʒ w ɹ/) or Sylheti (/bʰ t̪

d̪ ʈ ɖ gʰ cʃ ɟʒ x r/). Finally, the percentage accuracy of vowel productions (PVC) was calculated for each language.

In order to examine the factors that predict accurate pronunciation patterns, a series of logistic mixed-effects regression models were conducted using the *lme4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2017) packages for R (R Core Team, 2020) in RStudio (R Studio Team, 2016). Mixed-effects modelling was chosen primarily due to the size of the dataset and the need to account for inter-speaker and inter-item variation when predicting which factors influence accuracy (Johnson, 2009: 365).

Each model contained *accuracy* as the dependent variable with 'accurate' fitted as the baseline. Separate models were run on the consonant and the vowel data. With respect to the former, independent variables included in the modelling were *language*, *generation*, *manner*, *overlap*, and *position* along with a number of interactions. For the vowel analysis, only *language* and *generation* were considered as independent variables. *Word* and *speaker* were included as random effects in all models. Table 3 shows the dependent variables included in the study along with the baseline factor level and remaining factor levels.

A general-to-specific approach to modelling was taken, whereby non-significant variables were removed from the model in order to improve the model fit. The robustness of one model over another was ascertained based on the results of a log-likelihood test. The best-fitting models are presented in the results sections below.

Table 3. Independent variables and factor levels included in the mixed-effects modelling; the first factor level listed was set as the baseline.

Independent Variable	**Factor levels**
Language	English Sylheti

Generation	2 3
Position	final initial
Manner	affricate fricative liquid/glide nasal stop
Overlap	shared unshared

Finally, in addition to accuracy, we examined the types of phonological processes produced by the two sets of children and their percentage of occurrence. Errors were coded in terms of the following categories: *affrication, de-affrication, de-voicing, final consonant deletion, fricative simplification, fronting, gliding, spirantization, stopping,* and *trill deviation*. We included these categories since they have been attested in previous studies on phonological development in monolingual and bilingual children (e.g. Goldstein & Bunta, 2012; Prezas, Hodson & Schommer-Aikins, 2014; Shriberg & Kwiatkowski, 1980). In addition, we included three processes to account for cross-linguistic differences between Sylheti and English consonants. The first, *de-aspiration*, refers to aspirated plosives, notably Sylheti /bʰ/ and /gʰ/, being realized without aspiration. The second, *de-retroflexion*, in turn, refers to retroflex plosives being realized at a more anterior place of articulation, i.e. as alveolars. The final process, *de-dentalization*, refers to Sylheti dental plosives being realized as alveolars. Note that we avoided referring to the latter two processes as instances of *fronting* and *backing*, respectively, despite changes in place of articulation since these terms have specific significance in clinical contexts. We only considered *fronting* to have taken place if a palatal or velar consonant was realized in a more anterior position, e.g. /ŋ/>[n] or /ʃ/ > [s].

3. Results

3.1. Accuracy: consonant productions

The 4624 elicited words contained a total of 7597 consonant tokens for analysis. Of this total, 56.47% of tokens appeared in initial position (n = 4290) and 43.53% were final tokens (n = 3307). Of the total number of tokens, 60.96% were shared phonemes across English and Sylheti (n = 4631) compared to 39.04% of unshared phonemes (n = 2966). Table 4 shows the number of consonant tokens by language and generation, table 5 the number and percentage of tokens by manner in English and Sylheti.

Table 4. Number of consonant tokens by language and generation.

	GENERATION 2	**GENERATION 3**	***Total***
English	1863	2065	*3928*
Sylheti	1743	1926	*3669*
Total	*3606*	*3991*	*7597*

Table 5. Number and percentage of consonant tokens by manner in English and Sylheti

	English **% (n)**	**Sylheti** **% (n)**	***Total tokens***
Affricate	9.09% (n=357)	2.40% (n=88)	*445*
Fricative	29.07% (n=1142)	31.81% (n=1167)	*2309*
Liquid/Glide	12.60% (n=495)	13.41% (n=492)	*987*
Nasal	14.36% (n=564)	16.03% (n=588)	*1152*
Stop	34.88% (n=1370)	36.35% (n=1334)	*2704*
Total	100% (n=3928)	100% (n=3669)	7597

Of the English tokens, 96.21% were produced accurately (n = 3779) compared to 86.81% of Sylheti tokens (n = 3185). Tables 6 and 7 show the percentage of accurately produced tokens by generation in English and Sylheti, respectively. This includes measures of overall consonant accuracy (PCC-R) as well as accuracy by position (PCC-R word-initial; PCC-R word-final), cross-linguistic overlap (PCC-R shared; PCC-R unshared) and manner class (PCC-R stops; PCC-R fricatives; PCC-R affricates; PCC-R nasals; PCC-R liquids/glides).

Table 6. Phonological accuracy measures: English; standard deviations in parenthesis

	GENERATION 2	GENERATION 3
	% (SD)	% (SD)
PCC-R	97.32 (2.88)	95.22 (6.33)
PCC-R Word-initial	96.54 (3.87)	94.33 (6.97)
PCC-R Word-final	98.07 (1.86)	96.02 (6.06)
PCC-R Shared	99.25 (0.78)	97.4 (4.56)
PCC-R Unshared	94.58 (6.53)	92.32 (9.80)
PCC-R Stops	99.84 (0.48)	99.58 (0.95)
PCC-R Fricatives	96.28 (3.70)	91.77 (10.24)
PCC-R Affricates	95.81 (8.39)	87.37 (30.51)
PCC-R Nasals	98.86 (1.71)	99.63 (1.17)
PCC-R Liquids/ Glides	91.28 (12.38)	91.6 (11.38)

Table 7. Phonological accuracy measures: Sylheti; standard deviations in parenthesis

	GENERATION 2	GENERATION 3
	% (SD)	% (SD)
PCC-R	90.07 (1.86)	83.81 (5.47)
PCC-R Word-initial	88.25 (1.4)	82.04 (5.61)
PCC-R Word-final	93.02 (3.27)	86.55 (6.84)
PCC-R Shared	99.01 (0.54)	98.71 (1.26)
PCC-R Unshared	74.28 (5.09)	56.75 (14.86)
PCC-R Stops	78.79 (5.6)	66.11 (16.76)
PCC-R Fricatives	97.82 (1.63)	98.0 (2.43)
PCC-R Affricates	37.22 (16.79)	51.0 (24.47)
PCC-R Nasals	99.64 (1.08)	100.0 (0.0)
PCC-R Liquids	100 (0)	92.28 (11.47)

The results for both languages revealed high levels of accuracy on consonant production by both sets of children, in particular in English. To investigate the factors

that predict accurate consonant production patterns, logistic mixed effects modelling was initially undertaken on the entire dataset (*n*=7597). *Accuracy* was included as the dependent variable and *language*, *generation*, *manner*, *position*, and *overlap* were included as independent variables. *Word* and *participant* were set as random effects and a number of interactions between the independent variables were tested. These models failed to converge, however, and it was therefore decided to undertake separate statistical analyses of the English and Sylheti data with an adjusted alpha level (α = 0.025). The remainder of the analysis treats the English and Sylheti datasets separately.

3.1.1. English data

Table 8 shows the results of the best-fitting model conducted on the English data. The table shows that *overlap* and *manner* were significant predictors of *accuracy*, but, importantly, not *generation* nor any interaction involving *generation*.

Table 8. Regression coefficients with *z*- and *p*-values for the final model predicting production accuracy of consonant production in English (*n* = 3928).

	β	***SE***	***z***	***p***
(Intercept)	**5.11516**	**0.6432**	**7.953**	**<0.001**
Overlap (Unshared)	**-1.66171**	**0.34325**	**-4.841**	**<0.001**
Manner (Fricative)	-0.21734	0.43688	-0.497	0.619
Manner (Approximant)	-0.01634	0.48064	-0.034	0.973
Manner (Nasal)	1.56481	0.74159	2.11	0.035
Manner (Stop)	**3.07758**	**0.57086**	**5.391**	**<0.001**

As Table 8 shows, sounds which were not part of the phonologies of both English and Sylheti were less likely to be produced accurately (β = -1.661, z = -4.841, p < 0.001). Figure 1 shows the percentage of accurate and inaccurate productions

according to whether the phonemes are shared or unshared between English and Sylheti. Of the shared tokens, 98.37% were produced accurately (n = 2300) compared to 93.02% of unshared tokens (n = 1479). It should be noted that, due to the relatively low number of tokens which were inaccurate, these results should be treated with caution.

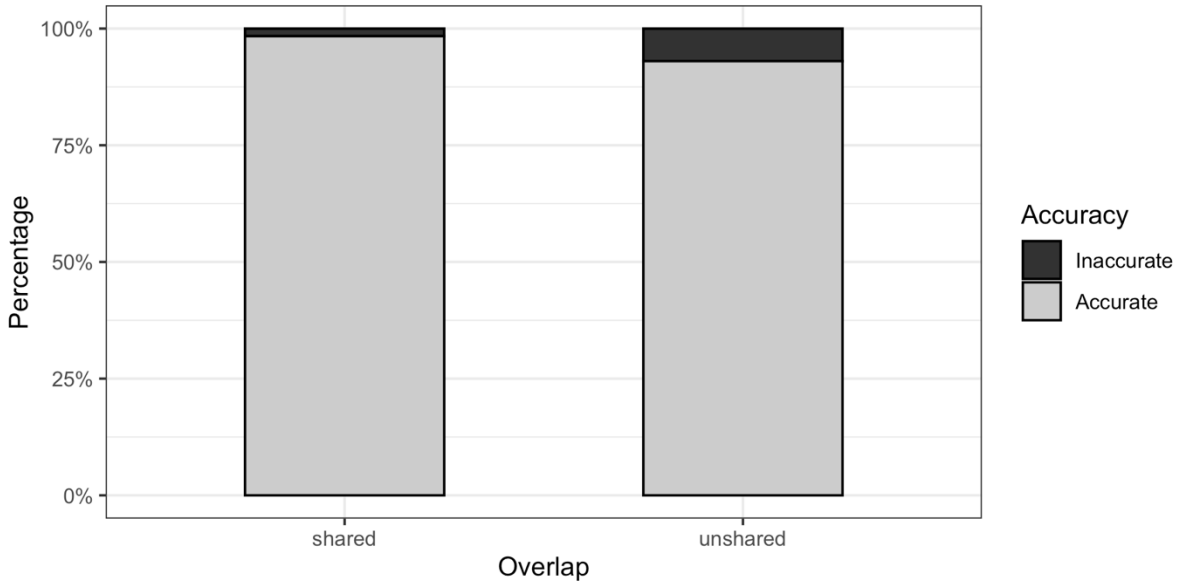

Figure 1. Percentage of accurate and inaccurate productions in English according to whether the phonemes are shared or unshared between English and Sylheti.

Table 8 also shows that *manner* was a significant predictor of *accuracy* in the English data. Specifically, stop consonants were more likely to be produced accurately compared to the baseline factor level, affricates (β = 3.078, z = 5.391, p < 0.001). Fricatives, approximants, and nasals, in contrast, did not differ significantly in accuracy from baseline based on an adjusted alpha level of 0.025. Figure 2 shows the percentage of accurate and inaccurate productions in English according to manner of articulation.

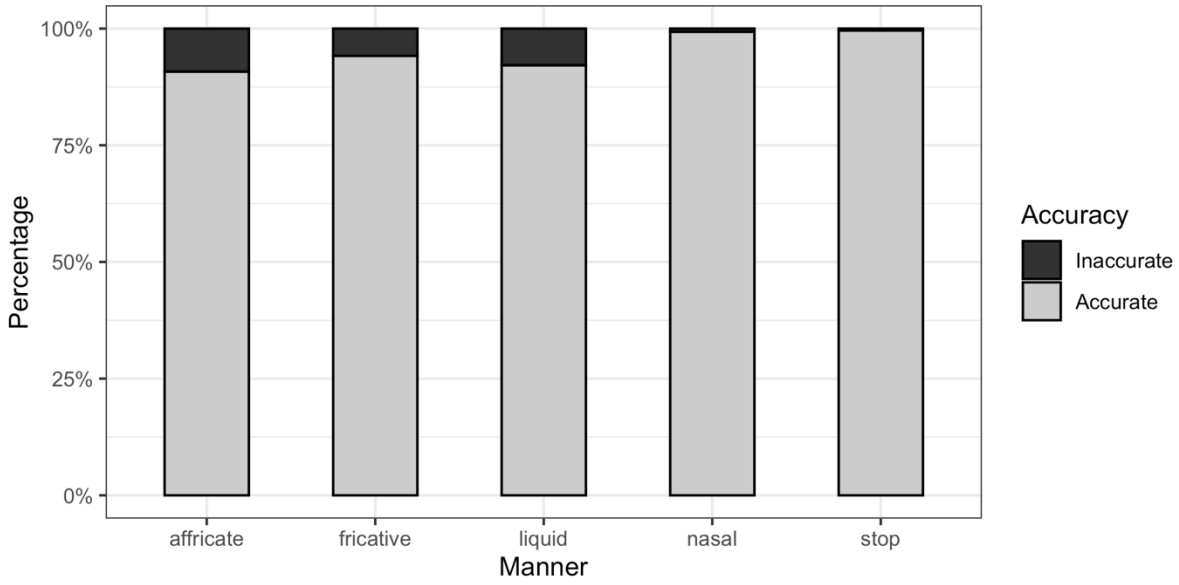

Figure 2. Percentage of accurate and inaccurate productions in English according to manner of articulation.

3.1.2. Sylheti data

Table 9 shows the results of best-fitting model conducted on the Sylheti data. The table shows that there was a significant interaction between *generation* and *overlap* in the data (β = -1.326, z = -4.859, p < 0.004).

Table 9. Regression coefficients with z- and p-values for the final model predicting production accuracy of consonants in Sylheti (n = 3669).

	β	***SE***	***z***	***p***
(Intercept)	**7.23894**	**0.6**	**12.065**	**< 0.001**
Generation (3)	0.01645	0.53083	0.031	0.975
Overlap (Unshared)	**-5.25725**	**0.61124**	**-8.601**	**< 0.001**
Generation*Overlap (3*Unshared)	**-1.32567**	**0.46368**	**-2.859**	**0.004**

The interaction between *generation* and *overlap* is shown in Figure 3. Both Table 9 and Figure 3 indicate that *generation* does not influence the likelihood of accurate production in Sylheti when phonemes appear in both English and Sylheti.

However, when phonemes are not shared between English and Sylheti, those in GENERATION 2 are more likely to produce accurate realizations compared to those in GENERATION 3.

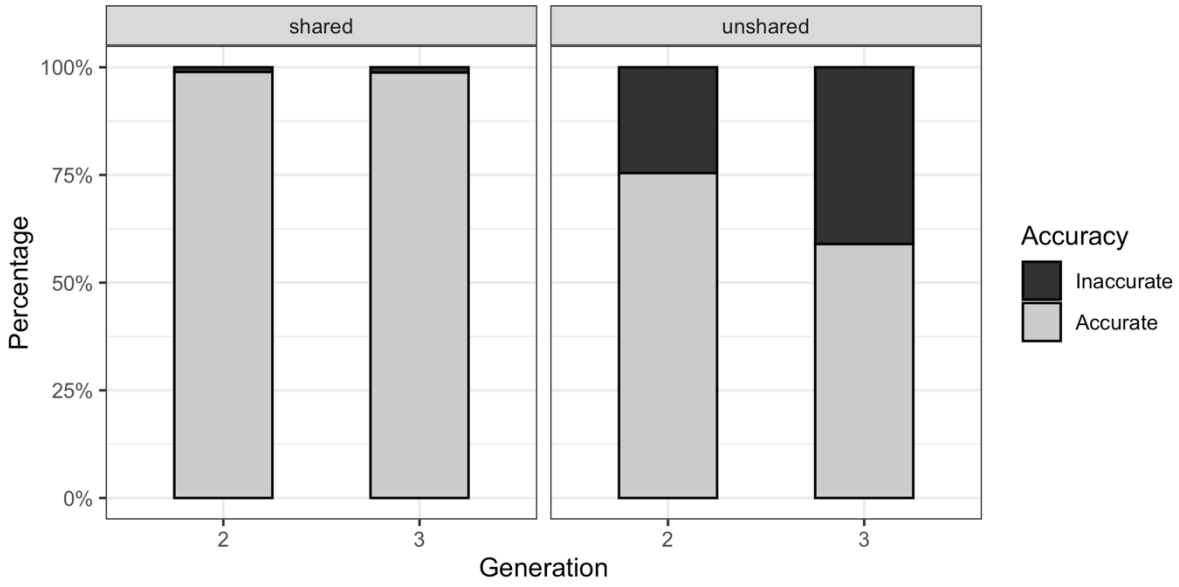

Figure 3. Percentage of accurate and inaccurate productions in Sylheti by generation and overlap.

3.2. Accuracy: vowel productions

A total of 5804 vowel tokens were analyzed in the children's word productions. This total comprised 2426 English tokens and 3378 Sylheti tokens across GENERATION 2 and GENERATION 3. Tables 10 and 11 show the percentage and number of accurate and inaccurate productions by generation in English and Sylheti, respectively.

Table 10. Number and percentage of vowel tokens produced accurately by generation in English (n = 2426).

	GENERATION 2 (%)	**GENERATION 3 (%)**	**GENERATION 2 (*n*)**	**GENERATION 3 (*n*)**	***Total***
Accurate	98.09	99.29	*1132*	*1263*	*2395*
Inaccurate	1.91	0.71	*22*	9	*31*
Total	-	-	*1154*	*1272*	*2426*

Table 11: Number and percentage of vowel tokens produced accurately by generation in Sylheti (*n* = 3378).

	GENERATION 2 (%)	**GENERATION 3 (%)**	**GENERATION 2 (*n*)**	**GENERATION 3 (*n*)**	***Total***
Accurate	98.57	98.08	1588	1733	*3321*
Inaccurate	1.43	1.92	23	34	*57*
Total	-	-	1611	1767	*3378*

As Tables 10 and 11 show, there is very little variation in the vowel data and the vast majority of tokens were accurate in both English (98.09%, n = 1132) and Sylheti (98.57%, n = 1588). Given the low number of inaccurate tokens across languages and generations, further statistical testing was not advisable and the results clearly show that vowel productions are overwhelmingly accurate in both languages regardless of generation.

3.3. Error patterns

Tables 12 and 13 depict the error patterns observed in the English and Sylheti consonant productions, respectively. The results confirm the comparatively greater accuracy overall on English than Sylheti consonants, as discussed for the accuracy measures. Thus, not only were there few errors on the English words, but there were also no noticeable differences across the generations. The most common error types on the English words were liquid gliding and fronting, the latter largely due to one GENERATION 3 child consistently realizing palato-alveolar fricatives and affricates as alveolars. There were also a number of fricative simplification errors, e.g. /θ/ > [s]. They conform to the typical developmental patterns seen in monolingual English-learning children. There was only one example of a Sylheti-specific sound being used in an English word, which could be attributed to cross-linguistic interactions: one GENERATION 2 child produced ‘red’ with an initial voiced alveolar trill.

Table 12. Error patterns in English by generation

	GENERATION 2	**GENERATION 3**	***Total***
	n (%)	*n* (%)	*n* (%)
Liquid gliding	14 (28.57%)	18 (18%)	32 (21.48%)
Fronting	-	30 (30%)	30 (20.13%)
Fricative simplification	10 (20.41%)	18 (18%)	28 (18.79%)
Devoicing	11 (22.45%)	4 (4%)	15 (10.07%)
Stopping	3 (6.12%)	6 (6%)	9 (6.04%)
Final consonant deletion	2 (4.08%)	3 (3%)	5 (3.36%)
Affrication	1 (2.04%)	2 (2%)	3 (2.01%)
Deaffrication	2 (4.08%)	1 (1%)	3 (2.01%)
Spirantisation	-	1 (1%)	1 (<1%)
Other	6 (12.24%)	17 (17%)	23 (15.44%)
Total	49	100	149

Table 13. Error patterns in Sylheti by generation

	GENERATION 2	**GENERATION 3**	***Total***
	n (%)	*n* (%)	n (%)
De-retroflexion	76 (44.19%)	83 (26.60%)	159 (32.85%)
De-dentalisation	33 (19.19%)	126 (40.38%)	159 (32.85%)
Deaspiration	22 (12.79%)	39 (12.50%)	61 (12.60%)
Deaffrication	27 (15.70%)	24 (7.69%)	51 (10.54%)
Liquid gliding	-	10 (3.21%)	10 (2.07%)
Trill deviation	-	10 (3.21%)	10 (2.07%)
Fronting	5 (2.91%)	3 (<1%)	8 (1.65%)
Final cons deletion	1 (<1%)	6 (1.92%)	7 (1.45%)
Devoicing	-	3 (<1%)	3 (<1%)
Stopping	-	2 (<1%)	2 (<1%)
Spirantisation	1(<1%)	1 (<1%)	1 (<1%)
Other	7 (4.07%)	5 (1.60%)	12 (2.48%)
Total	172	312	484

Analysis of the error patterns in Sylheti, in turn, revealed a greater number of error types by the GENERATION 3 than the GENERATION 2 children as well as a substantially higher number of de-dentalizations. The latter involved realizations of the Sylheti dental plosives /t̪/ and /d̪/ as alveolars. Note also that both sets of children made a fair amount of errors on retroflex plosives, which they invariably realized as alveolars. There were also a number of instances of *deaffrication*, with the Sylheti affricates /cʃ/ and /ɟʒ/ realized as fricatives, and of *de-aspiration*, with Sylheti /bʰ/ and /gʰ/ realized without aspiration. Most errors in Sylheti hence affected language-specific sounds.

4. Discussion

The purpose of this study was to examine consonant and vowel accuracy and error patterns in the heritage language (Sylheti) and the host language (English) of typically developing second- and third-generation British Bengali children. The results revealed significant differences in the Sylheti productions of the two groups of children on language-specific sounds but virtually equal performance in English. In what follows, we will first discuss the children's Sylheti production patterns and their implications for the maintenance and transmission of heritage languages across generations. We will then review the English results and their meaning for the acquisition of the host language in heritage language contexts. We then conclude by discussing the clinical implications of our findings.

4.1. Intergenerational differences in segmental accuracy and error patterns in Sylheti as a heritage language

The second- and third-generation heritage Sylheti-speaking children in this study showed both similarities and differences in their Sylheti consonant and vowel productions. First, as we hypothesized, the GENERATION 3 children were less accurate in their consonant productions than the GENERATION 2 children, however only on Sylheti-specific sounds. In line with this finding, the error analysis showed that GENERATION 3 children made more de-dentalization errors than GENERATION 2 children, producing Sylheti dental stops as alveolars. At the same time, however, GENERATION 3 children also displayed similarities to GENERATION 2 children in Sylheti phonological development. Indeed, both groups of children produced near-perfect accuracy on vowel production. Moreover, no differences were observed across the generations for sounds shared between Sylheti and English and neither word position nor manner class predicted accuracy or mediated accuracy patterns across the

generations. With the exception of de-dentalization errors, errors were also comparable across groups and showed an influence from English (i.e. retroflex stops were produced as alveolars and aspirated voiced stops were unaspirated). It is also important to point out that, despite the attested differences between GENERATION 2 and GENERATION 3 children, accuracy levels were relatively high for all participants, suggesting that even the third generation was on its way to develop the Sylheti phonological system.

These findings can be taken as evidence that third-generation heritage language-speaking children in the Bengali community in Wales may either need more time than second-generation children to develop the Sylheti phonological development or might never acquire language-specific sounds in a native manner. As discussed in Author 1 and Author 2 (2018), third-generation children tend to come from homes where English is the predominant language and they tend to live in ethnically heterogeneous areas with few Sylheti speakers. These factors significantly decrease the amount of Sylheti the children hear in their immediate environment, possibly prolonging or limiting speech sound development in the heritage language. Clearly, without longitudinal data we cannot determine whether phonological development in the heritage language of third-generation children is simply protracted or reaches a developmental plateau. As shown in the accuracy analysis, limited input appears to be inconsequential for sounds shared by the two languages, but it might be detrimental for the acquisition of sounds that are specific to Sylheti. Indeed, while children may successfully rely on English sounds for their production of phonetically similar Sylheti consonants (/b k ɡ m n ŋ f s z ʃ h j l/), production errors and comparatively lower accuracy may be typical for Sylheti-specific sounds that do not have phonetically similar equivalents in English (for example, aspirated voiced stops, retroflex and dental alveolars, the affricates /cʃ/ and /ɟʒ/, and /r/)

because, overall, these sounds are less frequent in the input and as such, they result in weaker and less reliable speech cues (see also Author 4, Author 1 and Other, 2018).

It is also possible that third-generation children are more likely than GENERATION 2 children to hear Sylheti from speakers whose input contains non-native features. This means that, in the context of overall reduced Sylheti input, GENERATION 3 children may actually be only hearing non-native Sylheti, with consequences on the Sylheti accent they will ultimately attain (see also Author 1 & Author 4, 2015). Indeed, Author 1 and Author 2 (2018) found non-native features in the Sylheti of GENERATION 3 children's mothers, suggesting that these were possibly partially responsible for the children's non-native realizations in the heritage language. Moreover, it cannot be excluded that such non-native features have become part of a new contact variety of Sylheti in Cardiff, a context characterized by a limited influx of new arrivals from Bangladesh, and, hence, by decreased contact with homeland norms. In this context, as in the case of Russian and Ukrainian immigrants in Canada (Nagy & Kochetov, 2013), heritage language speakers may consciously or unconsciously drift away from homeland norms towards those of the host language, highlighting a shift towards a British Bengali – rather than just Bengali – identity (e.g. Sevinç, 2016; Al-Azami, 2014). Taken together, these results suggest that both environmental (i.e. input-related) and psychological (i.e. identity-related) factors may limit the likelihood of heritage Sylheti speakers in Wales to attain a native accent in the heritage language.

4.2. Segmental accuracy and error patterns in English as the societal language

Contrary to the Sylheti findings, the results for English showed that neither GENERATION 2 nor GENERATION 3 British Bengali children face challenges or delays in developing the English phonological system, with *generation* not constituting a significant predictor in the best-fitting model. Thus, despite both sets of children

performing better on shared than unshared sounds, and on stops than affricates, as expected developmentally, they displayed high levels of accuracy on both English vowel and consonant production, with mean values on all measures above 90%, in line with the accuracy patterns reported for monolingual English-speaking children (cf. Dodd, Holm, Hua & Crosbie, 2003). Errors were also limited and comparable between groups, and mostly conformed to the typical developmental patterns seen in monolingual English-learning children.

These results suggest that speech sound development in the societal language is not compromised nor delayed in second- and third-generation British Bengali children growing up in Wales. Sylheti exposure at home and in the community does not significantly limit the quantity and quality of English exposure children have in the environment and has no consequences on English phonological development. Recall that the Bengali community in Wales is small, with only some 0.3% of the population (i.e. 10,687 individuals) considering themselves British Bangladeshis (ONS, 2011). Although the Bengali community in Cardiff has a close-knit social structure with shops, restaurants and community centers predominantly in the inner-city wards of Riverside and Grangetown, the number of Sylheti speakers remains limited as compared to the general English-speaking population. This means that the children in this community, despite having Sylheti-speaking parents, have sufficient opportunities in the environment (through school, media, etc.) to have contact with native English speakers and develop native English patterns. These findings may not be generalizable, however, to British Bengali children living in larger Bengali communities, as attested by the studies by McCarthy and colleagues that are set in London (McCarthy et al., 2013, 2014; Kirkham & McCarthy, 2020), or to other heritage language contexts with large communities (cf. Stangen, Kupisch, Proietti Erguen & Zielke, 2015).

5. Conclusions, Limitations and Directions for Future Research

Since we did not carry out a formal clinical intervention study and our sample was limited, the clinical implications from this study remain speculative. Nonetheless, the results suggest that 1) British Bengali children growing up in families where Sylheti is spoken to various extents can be expected to develop phonological skills in both languages during early childhood; 2) speech sound development in English, the societal language, should not be compromised nor delayed in second- and third-generation British-Bengali children; 3) third-generation British-Bengali children can be expected to display more difficulty with Sylheti-specific sounds than second-generation children. Overall, these findings indicate that while bilingual children of different generations may develop phonological skills in the societal language that are comparable to those of monolinguals, generational status might be an important variable to consider when evaluating bilingual children's phonological development in the heritage language, especially when it comes to language-specific sounds that are particularly dependent on amount of input for their development.

This work has several limitations that should be considered when interpreting the results. First, the sample in this study was small, and phonological skills were only assessed auditorily. Future studies should include larger samples of British Bangladeshi children, and given the heterogeneous nature of bilingual populations, children with different proficiency levels in English and Sylheti and different onsets of English acquisition. Future work should also complement the findings obtained in this study with acoustic measures as well as include native control groups. Moreover, while our approach to phonetic transcription took account of sub-phonemic information and made use of diacritic symbols, we recognize that an even more fine-grained approach might have identified more subtle differences in pronunciation accuracy. Finally, while the

present study showed differences in phonological accuracy across two generations of heritage language children, and as such is in line with previous work documenting an inter-generational drift away from homeland norms in the minority language (Author 1 & Author 2, 2018; Nagy & Kochetov, 2013), it is important to acknowledge that generational status coincides with language exposure. Future research is hence needed that disentangles these factors systematically.

This study complements previous work by Author 1 and Author 2 (2018) showing differences in the heritage language's speech patterns of second- and third-generation children, with the latter exhibiting an increasing drift towards the patterns of the host language. These findings have important implications for the maintenance, transmission and long-term survival of heritage languages, highlighting a need for future studies to go beyond second-generation speakers, especially in small heritage language communities that do not see a steady influx of new migrants. Future work should build on the present study and examine systematically what factors contribute to successful transmission and maintenance of speech patterns in heritage language settings.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Author 1 et al. (2014)

Author 1 et al. (2015)

Author 1 et al. (2019)

Author 1 & Author 2 (2018)

Author 1 & Author 4 (2015)

Author 4, Author 1 & other (2018)

Al-Azami, S. (2014). From Bengali to English: Sequential bilingualism of a second-generation British Bangladeshi. *International Journal of Bilingual Education and Bilingualism, 17*, 484-497. https://doi.org/10.1080/13670050.2013.831025.

Amengual, M. (2019). Type of bilingualism and its effect on the acoustic realization of allophonic variants: Early sequential and simultaneous bilinguals. *International Journal of Bilingualism*, 23, 954-970. https://doi.org/10.1177/1367006917741364.

Anderson-Mejías, P. L. (2005). Generation and Spanish language use in the Lower Río Grande Valley of Texas. *Southwest Journal of Linguistics, 24,* 1-12.

Ball, M.J., Howard, S. & Miller, K. (2018). Revisions to the extIPA chart. *Journal of the International Phonetic Association, 48*, 155-164. https://doi.org/10.1017/S0025100317000147

Barlow, J. A., Branson P. E., & Nip, I. S. B. (2013). Phonetic equivalence in the acquisition of /l/ by Spanish-English bilingual children. *Bilingualism: Language and Cognition, 16,* 68–85. https://doi.org/10.1017/S1366728912000235.

Bates D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1-48. https://doi.org/10.18637/jss.v067.i01.

Bayram, F., Kupisch, T., Pascual y Cabo, D., & Rothman, J. (2019). Terminology matters on theoretical grounds too!: Coherent grammars cannot be incomplete. *Studies in Second Language Acquisition*, 41, 257-264. https://doi.org/10.1017/S0272263119000287.

Benmamoun, E., Montrul, S., & Polinsky, M. (2013). Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics*, 39, 129–181. https://doi.org/10.1515/tl-2013-0009.

Blackledge, A., Creese, A., Baraç, T., Bhatt, A., Hamid, S., Wei, L., Lytra, V., Martin, P., Wu, C.-J., & Yağcioğlu, D. (2008). Contesting 'language' as 'heritage': Negotiation of identities in late modernity. *Applied Linguistics*, 29, 533-554. https://doi.org/10.1093/applin/amn024.

Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech, 46*, 217-243. https://doi.org/10.1177/00238309030460020801

Burrows, L., Jarmulowicz, L., & Oller, K. D. (2019). Allophony in English language learners: The case of taps in English and Spanish. *Language, Speech, and Hearing Services in Schools, 50,* 138–49. https://doi.org/10.1044/2018_LSHSS-17-0081.

Chalmers, R., & Miah, N. (1996). *Learning Sylheti.* London: The Centre for Bangladeshi Studies.

Chang, C. B., Yao, Y., Haynes, E. F., & Rhodes, R. (2011). Production of phonetic and phonological contrast by heritage speakers of Mandarin. *Journal of the Acoustical Society of America*, 129, 3964–3980. https://doi.org/10.1121/1.3569736.

Chatterji, S.K. (1926). *The origin and development of the Bengali language*. London: George Allen and Unwin.

Choi, J., Cutler, A., & Broersma, M. (2017). Early development of abstract language knowledge: Evidence from perception–production transfer of birth-language memory. *Royal Society Open Science*, 4. https://doi.org/10.1098/rsos.160660.

Cooperson, S. J., Bedore, L. M., & Peña, E. D. (2013). The relationship of phonological skills to language skills in Spanish-English-speaking bilingual

children. *Clinical Linguistics and Phonetics, 27*, 371–389. https://doi.org/10.3109/02699206.2013.782568.

Dodd, B., Holm, A., Hua, Z., & Crosbie, S. (2003). Phonological development: A normative study of British English-speaking children. *Clinical Linguistics and Phonetics*, 17, 617-643. https://doi.org/10.1080/0269920031000111348.

Evans, B., Mistry, A., & Moreiras, C. (2007). An acoustic study of first- and second-generation Gujurati immigrants in Wembley: Evidence for accent convergence? *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS XVI),* Saarbrücken, Germany, 6-10 August 2007, pp. 1741-1744.

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

Fennell, C.T., Sin-Mei Tsui, A., & Hudon, T.M. (2016). Speech perception in simultaneously bilingual infants. In E. Nicoladis & S. Montanari (eds.), *Bilingualism across the lifespan: Factors moderating language proficiency* (pp.43-62). Berlin: DeGruyter.

Fishman, J. (1964). Language maintenance and language shift as fields of inquiry: A definition of the field and suggestions for further development. *Linguistics, 9,* 32-70. https://doi.org/10.1515/ling.1964.2.9.32.

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

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

Gildersleeve-Neumann, C.E., Peña, E.D., Davis, B.L., & Kester, E.S. (2009). Effects on L1 during early acquisition of L2: Speech changes in Spanish at first English contact. *Bilingualism: Language and Cognition, 12,* 259–272. https://doi.org/10.1017/S1366728908003994.

Goldstein, B.A., & Bunta, F. (2012). Positive and negative transfer in the phonological systems of bilingual speakers. *International Journal of Bilingualism, 16,* 388–401. https://doi.org/10.1177/1367006911425817.

Goldstein, B.A., Fabiano, L., & Iglesias, A. (2004). Spontaneous and imitated productions in Spanish-speaking children. *Language, Speech, and Hearing Services in Schools, 35,* 5–15. https://doi.org/10.1044/0161-1461(2004/002).

Gope, A. (2018). The phoneme inventory of Sylheti: Acoustic evidences. *Journal of Advanced Linguistic Studies*, 7, 7-37.

Gope, A., & Mahanta, S. (2015). An acoustic analysis of Sylheti phonemes. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS XVIII).* Glasgow, ISBN 978-0-85261-941-4. Paper number 630, retrieved from http://www.icphs2015.info/pdfs/Papers/ICPHS0630.pdf

Grosjean, F. (2001). The bilingual's language modes. In Nicol, J. (ed.), *One mind, two languages: Bilingual language processing* (pp. 1-22). Oxford: Blackwell.

Hamid, S. (2007). *A study of language maintenance and shift in the Sylheti community in Leeds. A cross-generational survey of language behaviour, attitudes and perceptions on Sylheti ethnocultural migrants from rural backgrounds*. PhD dissertation: University of York.

Hyltenstam, K., Bylund, E., Abrahamsson, N., & Park, H. S. (2009). Dominant-language replacement: The case of international adoptees. *Bilingualism: Language and Cognition*, 12, 121–140. https://doi.org/10.1017/S1366728908004008.

International Phonetic Association (2005). *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

Johnson, D. E. (2009). Getting off the GoldVarb standard: Introducing Rbrul for mixed- effects variable rule analysis. *Language and Linguistics Compass, 3,* 359-383. https://doi.org/10.1111/j.1749-818X.2008.00108.x.Keffala, B., Scarpino, S., Scheffner Hammer, C., Rodriguez, B., Lopez, L., & Goldstein, B. (2020). Vocabulary and phonological abilities affect dual language learners' consonant production accuracy within and across languages: A large-scale study of 3- to 6-year-old Spanish–English dual language learners. *American Journal of Speech-Language Pathology, 29,* 1196-1211. https://doi.org/10.1044/2019_AJSLP-19-00145.

Kehoe, M. & Girardier, C. (2020). What factors influence phonological production in French-speaking bilingual children, aged three to six years? *Journal of Child Language, 47*, 945-981. https://doi.org/10.1017/S0305000919000874.

Khan, S. (2010). Illustrations of the IPA: Bengali (Bangladeshi Standard). *Journal of the International Phonetic Association*, *40,* 221-225. https://doi.org/10.1017/S0025100310000071.

Kirkham, S., & McCarthy, K. M. (2020). Acquiring allophonic structure and phonetic detail in a bilingual community: The production of laterals by Sylheti-English bilingual children. *International Journal of Bilingualism.* First published August 4, 2020. https://doi.org/10.1177/1367006920947180

Kupisch, T., & Rothman, J. (2018). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, 22, 564-582. https://doi.org/10.1177/1367006916654355.

Kupisch, T., Barton, D., Hailer, K., Klaschik, E., Stangen, I., Lein, T., & Weijer, J. van de (2014). Foreign accent in adult simultaneous bilinguals. *Heritage Language Journal*, *11,* 123-150.

Kuznetsova, A., Brockhoff, P.B. & Christensen, R.H.B. (2017). LmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*, 1-26. https://doi.org/10.18637/jss.v082.i13.

Lohndal, T., Rothman, J., Kupisch, T., & Westergaard, M. (2019). Heritage language acquisition: What it reveals and why it is important for formal linguistic theories. *Language and Linguistics Compass*, 13, 1–19. https://doi.org/10.1111/lnc3.12357.

MacWhinney, B. (2005). A unified model of language acquisition. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). Oxford, United Kingdom: Oxford University Press.

McCarthy, K., Evans, B., & Mahon, M. (2011). Detailing the phonetic environment: A sociophonetic study of the London Bengali community. *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII),* Hong Kong, 17-21 August 2011, pp. 1354-1357.

McCarthy, K., Evans, B., & Mahon, M. (2013). Acquiring a second language in an immigrant community: The production of Sylheti and English stops and vowels by London-Bengali speakers. *Journal of Phonetics, 41,* 344–358. https://doi.org/10.1016/j.wocn.2013.03.006.

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

McLeod, S. & Masso, S. (2019). Screening children's speech: The impact of imitated elicitation and word position. *Language, Speech, and Hearing Services in Schools, 50*, 71-82. https://doi.org/10.1044/2018_LSHSS-17-0141.

Montrul, S. (2016). *The acquisition of heritage languages*. Cambridge: Cambridge University Press.

Mora, J. C., & Nadeu, M. (2012). L2 effects on the perception and production of a native vowel contrast in early bilinguals. *International Journal of Bilingualism, 16,* 484–500. https://doi.org/10.1177/1367006911429518.

Munro, S., Ball, M.J., Müller, N., Duckworth, M., & Lyddy, F. (2005). Phonological acquisition in Welsh-English bilingual children. *Journal of Multilingual Communication Disorders*, 3, 24-49. https://doi.org/10.1080/14769670410001683467.

Nagy, N., & Kochetov, A. (2013). VOT across the generations: A cross-linguistic study of contact-induced change. In Siemund, P., Gogolin, I., Schulz, M. & Davydova, J. (eds.), *Multilingualism and language diversity in urban areas: Acquisition, identities, space, education* (pp. 19-38). Amsterdam/Philadelphia: John Benjamins.

Oh, J., Jun, S., Knightly, L., & Au, T. (2003). Holding on to childhood language memory. *Cognition*, *86,* 53–64. https://doi.org/10.1016/S0010-0277(02)00175-0.

Office for National Statistics [ONS] (2011). Key statistics for Wales, retrieved on 22 December 2015 from http://www.ons.gov.uk/ons/rel/census/2011-census/key-statistics-for-unitary-authorities-in-wales/stb-2011-census-key-statistics-for-wales.html

Pierce, L. J., Chen, J.-K., Delcenserie, A., Genesee, F., & Klein, D. (2015). Past experience shapes ongoing neural patterns for language. *Nature Communications*, 6, 1–11. https://doi.org/10.1038/ncomms10073.

Polinsky, M. (2018). *Heritage languages and their speakers.* Cambridge: Cambridge University Press.

Prezas, R.F., Hodson, B.W., & Schommer-Aikins, M. (2014). Phonological assessment and analysis of bilingual preschoolers' Spanish and English word productions. *American Journal of Speech-Language Pathology, 23,* 176–185. https://doi.org/10.1044/2013_AJSLP-12-0132.

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available: https://www.R-project.org/

R Studio Team. 2016. *RStudio: Integrated Development for R.* RStudio, Inc., Boston, MA. Available: http://www.rstudio.com/

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 and Communication Disorders, 51,* 368–383. https://doi.org/10.1111/1460-6984.12214.

Saito, K. (2017). Effects of sound, vocabulary, and grammar learning aptitude on adult second language speech attainment in foreign language classrooms. *Language Learning, 67*, 665-693. https://doi.org/10.1111/lang.12244.

Saito, K., Sun, H., & Tierney, A. (2018). Explicit and implicit aptitude effects on second language speech learning: Scrutinizing segmental and suprasegmental sensitivity and performance via behavioural and neurophysiological measures. *Bilingualism: Language and Cognition, 22*, 1123-1140. https://doi.org/10.1017/S1366728918000895.

Scarpino, S. E. (2011). *The effects of language environment and oral language ability on phonological production proficiency in bilingual Spanish-English speaking children.* PhD dissertation: Pennsylvania State University.

Scarpino, S. E., Hammer, C. S., Goldstein, B., Rodriguez, B. L., & Lopez, L. M. (2019). Effects of home language, oral language skills, and cross-linguistic phonological abilities on whole-word proximity in Spanish–English-speaking children. *American Journal of Speech-Language Pathology, 28*(1), 174–187. https://doi.org/10.1044/2018_AJSLP-18-0050.

Sevinç, Y. (2016). Language maintenance and shift under pressure: Three generations of the Turkish immigrant community in the Netherlands. *International Journal of the Sociology of Language,* 242, 81-117. https://doi.org/10.1515/ijsl-2016-0034.

Shriberg, L.D., Austin, D., Lewis, B.A., McSweeny, J.L., & Wilson, D.L. (1997). The Percentage of Consonants Correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research, 40,* 708–722. https://doi.org/10.1044/jslhr.4004.708.

Shriberg, L. D., & Kwiatkowski, J. (1980). *Natural Process Analysis: A procedure for phonological analysis of continuous speech samples.* New York: Macmillan.

Siegel, R., Winitz, H., & Conkey, H. (1963). The influence of testing instruments on articulatory responses in children. *Journal of Speech and Hearing Disorders*, 28, 67–76. https://doi.org/10.1044/jshd.2801.67.

Smit, A. B. (1993). Phonologic error distributions in the Iowa-Nebraska articulation norms project: Consonant singletons. *Journal of Speech and Hearing Research, 36,* 533–547. https://doi.org/10.1044/jshr.3603.533.

Smith, M.W., & Ainsworth, S. (1967). The effects of three types of stimulation on articulatory response of speech defective children. *Journal of Speech and Hearing Research*, 10, 333–338. https://doi.org/10.1044/jshr.1002.333.Stangen, I., Kupisch,

T., Proietti Ergün, A.L., & Zielke, M. (2015). Foreign accent in heritage speakers of Turkish in Germany. In H. Peukert (ed.), *Transfer effects in multilingual language development* (pp. 87-108). Amsterdam: John Benjamins.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behaviour and Development, 7,* 49–63. https://doi.org/10.1016/S0163-6383(84)80022-3.

Appendix: *English and Sylheti target words*

English target word	**Transcription**	**Sylheti target word**	**Transcription**	**English gloss**
mouse	maus	আম	ɑm	‘mango’
cheese	tʃiz	মুখ	muk	‘face’
sun	sʌn	নাক	nɑx	‘nose’
day	deɪ	কান	xɑn	‘ear’
pig	pɪg	টুট	ʈuʈ	‘lips’
mud	mʌd	দাত	ɖɑʈ̪	‘teeth’
rain	ɹeɪn	চোখ	seɔk	‘eyes’
mum	mʌm	গলা	ˈgɔlɑ	‘neck’
girl	gɜl	গাল	gʰɑl	‘cheek’
sing	sɪŋ	চুল	sul	‘hair’
song	sɒŋ	চাবি	ˈsɑbi	‘keys’
moon	mun	তালা	ˈʈ̪ɑlɑ	‘lock’
night	naɪt	রানী	ˈɾɑni	‘queen’
watch	wɒtʃ	রাজা	ˈɾɑɟʒɑ	‘king’
time	taɪm	সোনা	ˈʃunɑ	‘gold’
lip	lɪp	বোতল	ˈbuʈ̪ɔl	‘bottle’
lips	lɪps	দুধ	ɖuɖ	‘milk’
hair	heə	চোর	sur	‘thief’
neck	nek	গাড়ি	ˈgʰɔri	‘clock’
nose	nəʊz	সময়	ˈʃɔmɔi	‘time’
cheek	tʃik	শিপ	sif	‘spoon’
mouth	maʊθ	শাপ	sɑf	‘clean’
teeth	tiθ	মেগ	meg	‘rain’
jam	dʒam	মাছ	mɑs	‘fish’
fat	fat	হাপ	hɑf	‘snake’
thin	θɪn	বন্দুক	ˈbɔnduk	‘gun’
kick	kɪk	লাট	lɑʈ̪	‘kick’
ball	bɔl	কাষ্ঠ	ruk	‘wood’
leg	leg	আত	ɑʈ̪	‘hand’
shoes	ʃuz	নোক	neɔk	‘fingernails’
sock	sɒk	পাঁচ	fɑs	‘five’
bum	bʌm	জাল	zɑl	‘spicy’
nine	naɪn	ঘাস	gʰɑʃ	‘grass’

fish	fɪʃ	গাছ	gʰas	‘tree’
hat	hat	ফুল	ful	‘flower’
scarf	skɑf	রৌদ	rɔid	‘sunshine’
glove	glʌv	পাখি	ˈfaki	‘bird’
knife	naɪf	বলা	ˈbʰɔla	‘wasp’
fork	fɔk	ঘর	gɔr	‘house’
horse	hɔs	নীছ	nis	‘downstairs’
leaf	lif	পাঔ	fau	‘foot’
shave	ʃeɪv	টেং	ʈeŋ	‘leg’
fridge	fɹɪdʒ	রং	rɔŋ	‘paint’
pin	pɪn	লাল	lal	‘red’
tap	tap	পেক	fex	‘mud’
water	ˈwɔtə	কেলা	ˈxela	‘toys’
pigs	pɪgz	হুনের	ˈhuner	‘listen’
duck	dʌk	পুক	fuk	‘insect’
book	bʊk	জারু	ˈzaru	‘sweeping brush’
read	ɹid	পেঁয়াজ	ˈfiaiɟʒ	‘onion’
page	peɪdʒ	শাড়ি	ˈʃari	‘saree’
bat	bat	ঘোড়া	ˈgura	‘horse’
dog	dɒg	মাথা	ˈmaʈa	‘head’
goal	gəul	মগজ	ˈmɔgɔz	‘brain’
chair	tʃeə	নাগা	ˈnaga	‘Naga chilli’
table	ˈteɪbəl	মাপ	maf	‘measure’
bath	bɑθ	টাকা	ˈʈaxa	‘money’
towel	ˈtaʊəl	পেট	feʈ	‘stomach’
wash	wɒʃ	হাড্ডিসার	ˈhalax	‘skinny’
jug	dʒʌg	তেল	ʈel	‘oil’
catch	katʃ	দুনিয়া	ˈɖunia	‘world’
sad	sad	চিরুন	ˈsirɔn	‘comb’
chips	tʃɪps	দাম	ɖam	‘price’
juice	dʒus	ভারী	ˈbʰari	‘heavy’
cub	kʌb	তাপ	ʈaf	‘fever’
V	vi	রস	rɔʃ	‘juice’
king	kɪŋ	জামির	ɟʒamir	‘orange’
orange	ˈ ɒɹɪndʒ	কুততা	ˈkuʈʈa	‘dog’
hand	hand	নুন	nun	‘salt’
five	faɪv	তিন	ʈin	‘three’
thumb	θʌm	চশমা	ˈcʃɔʃma	‘glasses’
ring	ɹɪŋ	ব্যাং	beŋ	‘frog’
house	haʊs	মানুশ	ˈmanuʃ	‘people’
path	pɑθ	লেমবু	ˈlembu	‘lemon’
roof	ɹuf	ঠিবি	ˈɖibi	‘containers’
cup	kʌp	ঢালের	ˈɖaler	‘pour out’
van	van	হুখা	ˈhukar	‘dry’
web	web	গান	gan	‘sing’
bird	bɜd	বিলই	ˈbilai	‘cat’
dice	daɪs	কাট	xaʈ	cut’
red	ɹed	বালিশ	ˈbaliʃ	‘pillow’
thigh	θaɪ	জুতা	ˈzuʈa	‘shoe’
jelly	ˈdʒeli	গরম	ˈgɔrɔm	‘hot’
whale	weɪl	মোজা	ˈmuza	‘socks’

yacht	jɒt	বাঘ	bɑg	‘tiger’
badge	badʒ	ঈট	iʈ	‘brick’
shirt	ʃɜt	কাপ	kɑf	‘cup’
laugh	lɑf	গাড়ী	ˈgɑri	‘car’
yellow	ˈjeləʊ	রাখ	rɑx	‘keep’
yolk	jəʊk	কলম	ˈxɔlɔm	‘pen’
sugar	ʃʊgə	কাগজ	ˈxɑgɔz	‘paper’
shop	ʃɒp	লেখ	lex	‘writing’
zip	zɪp	কলা	ˈxɔlɑ	‘banana’
Z	zed	হীআল	ˈhiɑl	‘fox’
brush	bɹʌʃ	ঘুম	gʰum	‘sleep’
car	kɑ	চাঁদ	cʃɑɖ̲	‘moon’
crash	kɹaʃ	আগুন	ˈɑgun	‘fire’
drive	dɹaɪv	জলে	ˈzɔle	‘burn’
goat	gəʊt	পাল	fɑl	‘jumping’
garden	ˈgɑdən	টানের	ˈʈɑner	‘pulling’
job	dʒɒb	চল্লিশ	ˈcʃɑliʃ	‘forty’
vine	vaɪn	ডাক	ɖɑx	‘call out’
thorn	θɔn	জুড়ে	ˈzure	‘loud’
foot	fʊt	বালতি	ˈbʰɑlʈi	‘bucket’
cherry	ˈʧeɹi	বালু	ˈbʰɑlu	‘sand’
crab	kɹab	বাঁশি	ˈbʰɑʃi	‘whistle’
beach	biʧ	ভয়	ɖɔr	‘afraid’
globe	gləʊb	শর	ʃɔr	‘pig’
earth	ɜθ	নাচ	nɑs	‘dancing’
bed	bed	ঠিকানা	ˈʈikɑnɑ	‘address’
witch	wɪʧ	নাম	nɑm	‘name’
flush	flʌʃ	বাড	bɑɖ̲	‘bad’
zebra	ˈzebɹə	টাড়	ʈ̲ar	‘wire’
zoo	zu	বাদা	ˈbɑɖ̲ɑ	‘rubbish’
think	θɪŋk	নানি	ˈnɑni	‘grandmother’
itch	ɪʧ	সাত	ʃɑʈ̲	‘seven’
dance	dɑns	দশ	ɖ̲ɑʃ	‘ten’
bridge	bɹɪdʒ	খোলের	ˈkuler	‘open’
yes	jes	কীচা	ˈkisɑ	‘bag’
you	ju	ভাত	bɑʈ̲	‘rice’
vampire	ˈvampaɪə	সুখী	ˈkuʃi	‘happy’
wave	weɪv	ডীল	ɖɑil	‘lentils’
wrong	ɹɒŋ	-	-	-
bag	bag	-	-	-
gun	gʌn	-	-	-