



© 2021. This is the accepted manuscript, made available under the CC-BY 4.0 license

<https://creativecommons.org/licenses/by/4.0/>. Please note that some minor changes in the final article are not reflected in this version; refer to the main article <https://doi.org/10.1017/S0305000921000593>.

---

# Variation in quality of maternal input and development of coda stops in English-speaking children in Singapore

**Author names:** Jasper Hong, SIM; Brechtje POST

**Affiliations:** University of Cambridge

**Address for correspondence:**

Phonetics Laboratory  
University of Cambridge  
Sidgwick Avenue  
Cambridge CB3 9DA  
United Kingdom.

**Acknowledgements**

Our deepest appreciation to the families who participated in this study and the mothers who helped in checking the translation of the vocabulary checklist. We would like to thank Katrina Li Kechun and statisticians from the University of Cambridge Mathematics Department for their help and comments on the statistical analyses in an earlier version of this paper. We also thank the two anonymous reviewers for their careful reading of the paper and their invaluable comments and suggestions. This work was supported by fieldwork grants from the Faculty of Modern and Medieval Languages and Linguistics, and Jesus College, University of Cambridge. All mistakes are our own. The first author would also like to thank National Institute of Education, Nanyang Technological University, Singapore for the fieldwork grants and PhD funding that made the study possible.

**Abstract**

This study examines the effects of input quality on early phonological acquisition by investigating whether interadult variation in specific phonetic properties in the input is reflected in the production of their children. We analysed the English coda stop release patterns in the spontaneous speech of fourteen mothers and compared them with the spontaneous production of their preschool children. The analysis revealed a very strong positive input–production relationship; mothers who released coda stops to a lesser degree also had children who tended to not release their stops, and the same was true for mothers who released their stops to a higher degree. The findings suggest that young children are sensitive to acoustic properties that are subphonemic, and these properties are also reflected in their production, showing the importance of considering input quality when investigating child production.

## 1. Introduction

Individual variation in early language development and language outcomes of both monolingual and bilingual children may be attributed to differences in input quantity and quality. A growing number of studies have shown that greater access to linguistic information through a larger amount of language input generally leads to faster development in various linguistic domains (Hoff, 2006; Unsworth, 2016). In terms of phonological development, studies of bilingual children, for instance, have shown that the child's dominant language, which is typically defined as the language the child hears and uses the most frequently with significant others, is associated with higher rate of phonological development (e.g. Ball, Müller, & Munro, 2001; Law & So, 2006) and also phonological accuracy (En, Brebner, & McCormack, 2014; Goldstein, Bunta, Lange, Rodriguez, & Burrows, 2010; Wrembel, Marecka, Szewczyk, & Otwinowska, 2019). What is often overlooked, however, is that variability in phonological development or outcomes can also be a result of differences in the quality of input, that is, the specific phonetic and phonological properties in the language models. Monolingual children may be raised in mixed-accent or bi-dialectal families (e.g. Stanford, 2008; Thomas & Scobbie, 2015), or by bilingual caregivers who may modify their child-directed speech according to the age and gender of child, and situational context (e.g. Foulkes & Docherty, 2006; Smith, Durham & Fortune, 2007). Language input in bilingual communities can be even less homogenous, given the possible variation in the language background of caregivers (Lleó, 2016). Caregivers who speak a majority language and a heritage language, for example, may speak the native language with non-nativelike phonetic characteristics, and depending on their L2 use and length of residence outside of their L1 community, phonetic properties of both languages may also be qualitatively different from others (e.g. Fish, García-Sierra, Ramírez-Esparza, & Kuhl, 2017; Flege, Frieda, & Nozawa, 1997; Guion, Flege, & Loftin, 2000; Mayr & Montanari, 2015; Post & Jones, 2020). Further, even when both languages are acquired early, bilingual caregivers may differ in their language dominance, and so does the extent of cross-linguistic interactions, which affects production and perception (e.g. Amengual, 2018; Amengual & Chamorro, 2015). Such variation is commonplace in sociocultural contexts like Singapore and Malaysia, where speakers are all native speakers of their dialect but may differ in some properties of their accents according to their language background and various sociolinguistic factors (e.g. Phoon, Abdullah, & Maclagan, 2013; Sim, 2019). While much is known about variation in adult production, relatively fewer studies have examined the effects of such qualitative differences in the input on phonological acquisition in children. The present study focuses on this underexplored area of child phonological acquisition by examining whether interadult variation is reflected in the production of their children.

Studies that examined this input-production relationship have shown that speech properties of child production reflect specific properties of the caregiver input, especially in the early developmental years. Thomas and Scobbie (2015), for example, examined the FACE and GOAT vowels of a Glasgow boy aged 3;1 raised by parents with different British English accents; his father spoke Scottish Standard English (SSE), while his mother's accent closely resembled Southern Standard British English (SSBE). For the FACE lexical set, the boy used the SSBE [eɪ] predominantly, reflecting the accent of his mother. The vowel for the GOAT lexical set, however, was more mixed, but the boy used SSE [o] in a slight majority of the time. Studies of bilingual children also revealed how differential properties in the speech of caregivers are reflected in their children's speech. Khattab (2003), for instance, studied the voice onset time (VOT) production of two English-Arabic siblings aged seven and ten years, who were acquiring Arabic as a heritage language in England from their parents. She found that the idiosyncratic use of nasals and implosives in the production of voicing lead of the younger child was similar to the patterns found in the mother's pre-voiced stops. A recent study by Stoehr, Benders, van Hell, and Fikkert (2019) examined more directly the effects of non-native and attrited maternal input on children by investigating the production of VOT by Dutch-German bilingual pre-schoolers. These children acquired German as a heritage language predominantly from their mothers who spoke German as an L1. They acquired the majority language, Dutch from their fathers who were L1 speakers of Dutch, and also from their mothers who were L2 speakers. They found that individual variation in the VOT production of these child bilinguals was associated with individual variation of VOT in their mothers' non-native speech in Dutch and their mothers' attrited speech in the heritage language German. Effects of quality of input on bilingual outcomes can sometimes be difficult to ascertain, because differential features learned from the input can also resemble effects of cross-linguistic interactions in the bilinguals.

This study furthers the investigation of the input–production relationship and differs from these past studies in the following ways. The phonetic feature of interest in this study is English word-final oral stop release. Compared to segments and VOT, the presence or absence of coda stop release is much more variable and less predictable. For example, while /p, t, k/ are aspirated when they occur in the onset of a stressed syllable but not in a cluster after /s/, the same stops at the word-final position are not always (audibly) released, even if they precede the same phonetic environment. Therefore, this study also tested to see whether very young children are sensitive to differences in the statistical distribution of a variable feature in the input. Another difference is that instead of heritage languages, this study looked at a contact variety of English, Singapore English. There is therefore less variability and better comparability than when comparing between monolinguals and bilinguals, or native and non-native speakers, because all parents and children in this study were locals and native speakers of Singapore English and

Singapore Mandarin, and were living in the same broader linguistic community, but they differed in how frequently they released their English coda stops. Although these dyads were bilinguals, cross-linguistic interaction (CLI) is unlikely to present as a confounding factor. This is because, in addition to the children being highly English dominant, Mandarin does not allow coda oral stops (Hua, 2006), and the variable feature of coda stop non-release is a feature of Singapore English, the children's L1.

### *1.1. Coda stop release in English*

While coda stops are always unreleased in some languages (e.g. Korean, Cantonese), stop release is optional in many varieties of English. Speakers of established standards of English such as British and American English, for example, do not release coda stops all the time, and even less so in spontaneous speech. The release of stops is further modulated by factors such as place of articulation (PoA) and the position of the stop within the utterance. Fabricius (2002), in her sociolinguistic examination of t-glottaling (the pronunciation of syllable-final /t/ as glottal stop [ʔ]) in Received Pronunciation, reported that in interview style, t-glottaling at word-final position occurred 36% of the time before pauses, 40% before vowels, and on average 74% of the time before consonants, but they did so generally less frequently in the more careful reading passage style. That there is variability in how coda stops are released was also reported for American English by Song, Demuth, and Shattuck-Hufnagel (2012), who examined the development of acoustic cues to coda contrasts in monolingual children by analysing the spontaneous speech productions of six mother-child pairs. They found that for mothers, the likelihood of stop release varied with both PoA and utterance-position. Specifically, velar stops were released more frequently than alveolar stops, and utterance-final stops were released more frequently than utterance-medial ones (which included pre-vocalic stops).

Contrastingly, word-final singleton stops in Singapore English tend to be unreleased (or inaudibly released) or replaced by a glottal stop, and unreleased stops are also often accompanied by glottal reinforcement. In addition, syllable-final voiced obstruents are often devoiced (Bao, 2003; Deterding, 2007; Gut, 2005). Bao (2003) further added that these features are widely attested in all social strata of the community and found in both formal and informal speech. Quantitative information on stop release patterns in Singapore English was reported by Gut (2005), who examined the realisations of coda stop in the spontaneous speech of 16 adult Singaporean speakers (mostly Chinese) with an average age of 29 years. They were reported to be fluent and dominant in English, but were mostly early sequential bilinguals who learnt English from an average age of 5;6. Gut (2005) found that overall, coda stops were more frequently realised as a glottal stop and unreleased than released, but did not find a significant difference in

the realisations between voiceless and voiced stops, nor between Chinese and non-Chinese speakers. She reported similar effects of phonetic environment on the likelihood of coda stop release. Before consonants, coda stops of Chinese speakers were found to be released only about 10% of the time (52% unreleased, 38% as a glottal stop). Stops that preceded vowels and pauses were released more frequently (about 38% and 41% respectively), but were also as likely to be replaced with a glottal stop. There was also an indication of weak effects of PoA on stop release, as she found that of all tokens, /k/ was released most often, at 37.2% (n=113), while /t/ and /d/ were released less frequently, at 23.1% (n=511) and 31.5% (n=124) respectively. Compared to /k/, alveolar stops /t/ and /d/ tended to be unreleased (34% and 33% respectively), and /t/ was also the most likely to be produced as a glottal stop (42.9%).

### *1.2. Acquisition of English coda stops*

Previous studies have shown that children produce coda structure early, usually by around the age of two. As early as 1;6, children also exhibit adult-like use of cues to coda voicing and place contrasts (Demuth et al., 2009; Song et al., 2012), but the degree of systematicity and range of values for these cues may be different from adults. Song et al. (2012), for instance, found that at 1;6, children had more frequent stop releases, a greater mean number of release bursts, and more frequent and longer post-release noise than mothers. Indeed, early production can be inconsistent, and shows great within-speaker variability, where the same child may produce some coda consonants but not others (Stites, Demuth, & Kirk, 2004). Their early use of coda consonants may also be influenced by linguistic properties, such as segment type, vowel length, stress, position within the word, and prosodic structure (Kirk & Demuth, 2006). However, normative studies have shown that English-speaking children produce most coda stops (/p, b, t, d, k, g/) by the age of three (Dodd, Holm, Hua, & Crosbie, 2003). Song et al. (2012) also found the cues to coda contrasts, such as the effects of PoA and phonetic environment on stop release, were generally adult-like by 2;6.

The phonetic realisation of coda stops may differ between children, in part due to phonetic qualities of the input. Phonological acquisition in some contexts may involve competing alternatives between caregiver input and local norms, and some are further associated with social meanings. British-born speakers of South Asian heritage in United Kingdom, for example, having been exposed to Indian English by their caregivers and others in the ethnic community, may use retroflex [ɖ] in their English speech even if they are English dominant or English monolinguals, and some use them variably with the mainstream alveolar variant depending on the interlocutor (Sharma, 2011). In other contexts that have experienced significant language shifts like Singapore, previous generations of speakers may differ greatly in their language backgrounds, and

so children of later generations may receive L1 input from caregivers who were L2 learners, or from L1 speakers who have retained features from previous generations of L2 learners, and consequently exhibit these features in their own speech, even if they are highly English dominant. Indeed, Bao (1998) posited that the feature of non-release of coda stops in Singapore English is likely due to influence of substrate languages including Malay and Chinese dialects such as Hokkien, which are major languages in the sociolinguistic history of Singapore. Unlike Mandarin, these languages allow final codas /p, t, k/ like English, but they are unreleased and their preceding vowel is also glottalised. En, Brebner, & McCormack (2014), who examined the English phonology of English-Mandarin bilingual preschool Singaporean children (ages 4;0–4;5) using the Phonology Assessment from the Diagnostic Evaluation of Articulation and Phonology (Dodd, Zhu, Crosbie, Holm, & Ozanne, 2002), found that Mandarin-dominant children used phonological processes that may indicate potential effects of CLI (e.g. cluster reduction) that were not found in English-dominant children. However, all 70 children in the study, regardless of whether they were Mandarin or English dominant, glottalised syllable-final stops and devoiced syllable-final obstruents (e.g. [eʔk̚] and [eʔ] for *egg*), which suggest that, rather than being effects of CLI, these two features were learned from the input. Similarly, in his examination of intra-ethnic variation in the English-Malay adult bilinguals in Singapore, Sim (2019) found that his Malay-dominant subjects exhibited features that may potentially be attributed to CLI, such as unaspirated word-initial stops and the use of clear [l] syllable finally, as these were not found in the speech of English-dominant English-Malay bilinguals. English-dominant subjects, by contrast, displayed features that were not typical of Singapore English. For example, they preserved all tense-lax vowel pairs and produced VOT comparable with speakers of other established standards of English, but still sounded essentially Singaporean. However, he noted that all participants were early or simultaneous bilinguals, and should have formed separate phonetic categories for their two languages. Sim posited that the use and maintenance of ethnic features could also be due to socio-indexical reasons; based on the results from the language background survey, his Malay-dominant subjects were associated with more Malay-dominant families and social circles, and identified more with a more Malay-speaking culture than an English-speaking one. The exposure to a dominantly Malay-accented English accent could potentially explain how these differential features were acquired.

In complex multilingual contexts like Singapore, therefore, individual variation may be attributed to qualitative differences in the input given by individual caregivers. This means that even if two children received an equally high amount of English input, phonetic features in their English accents may differ because of qualitative differences in the input, and this is what the study sets out to investigate. If indeed children's production reflects the between-speaker variation in stop release in adults, the findings will lend support to acquisition theories that pay greater

emphasis on the role of input and the learning of phonetic forms, highlighting the sensitivities of children to subphonemic variation. To this end, this study sets out to test these three hypotheses:

- H1 Children will exhibit adult-like patterns in the distribution of realisations of coda stops.
- H2 Some mothers will release coda stops more frequently than what is expected based on local norms.
- H3 There will be a positive association between the stop release patterns of children and their mothers.

H1 is based on past findings on coda stop development that by as early as 2;6, children's stop production was generally adult-like (e.g. Song et al., 2012). Any systematic variation in the realisation of coda stops observed in the adults in this study should also be observed in their children, who were at least 2;8. H2 is based on the previous discussion that the accents of Singaporeans are not homogenous and can differ in qualitative aspects, even between bilingual speakers of the same languages, due to factors such as their language history, background and attitudes (e.g. Sim, 2019). H3 is based on past studies that observed similarities in the phonetic aspects of the input and the speech of monolingual children (e.g. Foulkes & Docherty, 2006; Smith et al., 2007) and bilingual children (e.g. Mayr & Montanari, 2015; Stoehr et al., 2019). Therefore, not only do we predict that children will produce adult-like patterns as a group (as specified in H1), we also predict that individual variation in the stop release patterns between mothers will be observed in the production of their children. For both H2 and H3, the analysis will also attempt to ascertain other language-external factors that contribute to any variation in stop release.

## 2. Methodology

### 2.1. Participants

The mother/child corpus used in this study consists of 14 Singaporean Chinese dyads, and the children were aged between 2;8 and 4;8 ( $M = 3;7$ ). These participants were selected from a larger corpus of 60 Singaporean families based on responses in a child language experience survey. This ensured that the subjects were comparable across various language-external factors that could affect phonological production (Kehoe & Havy, 2018; Sorenson Duncan & Paradis, 2016), which included language background and language dominance (e.g. En et al., 2014; Goldstein et al., 2010), child's vocabulary size (e.g. Scarpino, 2011), and socioeconomic status (Campbell et al., 2003). Children below 2;6 were excluded, as their production patterns may still be stabilising



and input effects may not be evident (Song et al. 2012; Vihman et al. 1994). Each component of the survey is further described below. Table 1 presents a summary of the details of the child participants.

The children were first matched in their language background. The children (8 girls and 6 boys) were all firstborns, to eliminate potential influence from older siblings. They were typically developing simultaneous bilinguals of Singapore English and Singapore Mandarin, who were exposed to both languages by the age of three (Genesee & Nicoladis, 2007). Their parents also spoke Mandarin and English.

Language dominance was measured with reference to existing instruments that were developed for multilingual contexts (e.g. Tan, 2011). The language use of the child was calculated from an accumulated measurement of the type (i.e. variety) and estimated amount and proportion of time the language variety was used with the significant people in his/her immediate ecosystem. Specifically, parents were asked to report the languages and specific varieties that their child used with significant adults and children (both direct/indirect input and output), the estimated percentage of the time that each language/variety was used, and the time spent with these people in hours per week. The calculation also considered the child's language use in self-interaction and exposure to media. The children selected for this study were all English dominant, who used Singaporean English at least 70% of the time ( $M = 83$ ,  $SD = 7.28$ ), to also minimise confounding effects of potential CLI. The exposure to other established standard varieties of English, particularly American and British English, from media consumption, was unexpectedly low for all children (around 1 to 2% of all English input), and therefore the influence of these varieties was limited. Some children were also exposed to other varieties such as Indonesian English or Filipino English through their domestic helpers, but exposure to those varieties was also low, with the highest being 11%. There are several caveats that concern the percentage language use results in Table 1 and the bilingual status of the children in this study. Many studies have classified bilinguals as those who have received a minimum of 10–20% of input in one of their languages (e.g. Kehoe & Havy, 2018; Lauro, Core & Hoff, 2020). This would mean that some of the children in this study would be considered monolinguals and others, bilinguals. However, at least for this study, a dichotomous classification based on purely quantitative terms may disregard the pluralistic nature of language acquisition in such a multilingual context, where in fact the children may be more 'bilingual' than the cumulative scores indicate them to be. Child C5, who would be considered a monolingual, for example, was reported to use English 100% of the time with her peers, stay-home helper and paternal grandparents, but used almost exclusively Mandarin with her maternal grandparents, and Mandarin about 20-40% of the time with her parents. Similarly, children C18 and C47 have the lowest percentage Mandarin use because they used mainly English with family members, but received Mandarin input 30-40% of the time at the

preschool/childcare that they attended on weekdays. Moreover, the cumulative scores of percentage of input between studies are not always comparable because of, for instance, differences in calculations and the different contexts that were considered in the measurement. Since the aim of this present study is not to compare between monolinguals and bilinguals, the child participants are all here regarded as English-dominant English-Mandarin bilinguals, with some regarded as more English-dominant than others in terms of overall language use. The language dominance of the mothers was measured using the Bilingual Language Profile (BLP; Birdsong, Gertken, & Amengual, 2012), which is a self-reported measure of the mother's language history, proficiency, use and attitudes. The composite dominance scores were automatically tabulated, and possible scores range from  $-218$  (Mandarin-dominant) to  $+218$  (English-dominant). The BLP scores of the mothers in this study suggest that none was Mandarin dominant ( $M = 82.72$ ,  $SD = 43.97$ , range =  $3.46-150.84$ ), but were English dominant to varying degrees and in different ways according to the four components measured by the BLP.

The survey also included the Family Affluence Scale (FAS) (Currie, Elton, Todd, & Platt 2008), an established measure of socioeconomic status (SES), but modified<sup>1</sup> to fit the Singaporean context. The FAS assesses SES by aggregating information on material affluence based on the material condition of the household. This study also included education level and profession of the parents as part of the measure. These items in the survey generate a composite score, with the highest possible SES score being 35; the average SES score of the participants was 23.5 ( $SD = 3.15$ ; range =  $19-31$ ).

Finally, a parental vocabulary checklist to measure their child's lexicon size was administered. As there was no established way to measure productive/expressive vocabulary for children of this age range (Milton & Treffers-Daller, 2013), a checklist composed of two elements was created for this study. The first is a local variant of the standardised MacArthur-Bates Communicative Development Inventories (MCDI), adapted by the National University of Singapore in both English and Mandarin Chinese, which was suitable for children below 36 months. The second component consists of two sets of high-frequency words taken from the vocabulary lists of the international Cambridge English Qualifications assessments for children from kindergarten to upper primary levels. The latter sets were adapted to the Singapore context and translated by the first author, who is a speaker of Singapore Mandarin, and the items were checked by two mothers who were also native speakers of Singapore Mandarin to ensure that the translations were accurate and reflective of local usage. The final checklist contained a total of

---

<sup>1</sup> The question in the original FAS, 'Do you have your own bedroom for yourself?' was replaced with 'What type of house does this child live in?'. The question 'Do you pay people from outside the family to work at your home on a regular (that is, on a daily or weekly) basis?' was also added.

1226 items in the two languages. Linear regression performed using R statistical software (R Core Team, 2020) and the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) on the checklist scores administered to 59 families of children of ages between 2;1 and 6;4 (i.e. the larger corpus mentioned above) revealed that age and amount of exposure to English were statistically significant predictors of the English vocabulary scores (age:  $\beta = 0.60$ ,  $p < .001$ ; English use:  $\beta = 0.26$ ,  $p < .05$ ;  $R^2 = .42$ ). Ethnic mother tongue use and age as main effects were also significant predictors of ethnic mother tongue vocabulary (age:  $\beta = 0.27$ ,  $p < .05$ ; ethnic mother tongue use:  $\beta = 0.48$ ,  $p < .001$ ;  $R^2 = .26$ ). This suggests that the vocabulary checklist is adequately discriminating, at least for the purpose of controlling for lexicon size in this study.

**Table 1.** Description of the child participants including age, gender, age of acquisition (AoA), percent use of English and Mandarin, English vocabulary score (Eng. Vocab.), total vocabulary score (Total vocab.), socio-economic status (SES) score, and mother's Bilingual Language Profile (BLP) score.

Subject ID	Age	Gender	AoA English	AoA Mandarin	% English use	% Mandarin use	Eng. vocab.	Total vocab.	SES	Mother's BLP score
C5	45	F	0	0	78	18	734	1187	23	31.34
C9	39	M	0	0	84	14	1027	1180	21	82.28
C17	53	M	0	0	79	8	860	968	21	3.46
C18	32	F	0	0	96	3	843	1023	25	129.68
C20	56	M	0	18	83	6	1136	1833	28	91.09
C24	34	F	0	0	79	20	885	1043	22	57.22
C30	48	F	0	0	71	28	1226	1907	22	68.75
C31	36	F	0	0	74	25	932	1327	19	20.89
C35	47	M	0	24	89	10	966	1017	23	150.84
C39	45	F	0	0	85	8	811	1083	24	107.43
C46	37	F	0	32	85	7	854	870	21	129.14
C47	47	M	0	12	92	4	1098	1337	31	106.26
C55	32	M	0	0	76	14	954	1376	23	65.20
C74	54	F	0	0	91	8	946	1491	26	114.43

Note: Age and AoA are in months

## 2.2. Materials

Naturalistic data from unstructured play and semi-structured interaction between the mother and child were used in the analysis. Each interaction lasted approximately 30 to 40 minutes for each pair. Activities during unstructured play included but were not limited to playing with toys, puzzle play and sketching/drawing. Parents were also asked to take part in semi-structured interaction using a large picture card that featured a park scene with many animals, food, objects and people engaged in leisure activities. Only speech in the informal style was included in the

analysis, to control for potential stylistic variation (Smith et al., 2007). Words that were read, or mimicked/imitated were excluded. Spontaneous speech is more representative of child-directed speech and the variant used in day-to-day interactions between mother and child. Elicitation techniques such as picture naming or word list reading, although allowing better control over the materials and therefore higher comparability of results, would very likely elicit canonical forms that might not reflect natural speech or local dialectal norms. An example of an interaction between a mother (M) and child (C) during a drawing activity is provided below, with words that were included in the analysis in bold and coda stops underlined, according to criteria that are described in the later section.

- (1) C. **Look**!
- M. What's this supposed to be?
- C. It's supposed to be a **shark**!
- M. A **shark**?!
- C. With fins, and one fin on **top**.
- M. Yeah, the dorsal fin. You **forgot**? It's called the dorsal fin.
- C. Dorsal fin.
- M. How does the dorsal fin **shape** like?

### 2.3. *Recording procedures*

The recordings took place in a quiet room with minimal reverberation in the participants' homes, without the presence of the researcher or any other person. To ensure that the recordings are of adequate quality for acoustic analysis of fine phonetic details, the mother and child each had pinned on their collar an omni-directional lapel microphone, which was connected to a NAGRA ARES-MII recorder recording at a sampling rate of 44.1 kHz at 16 bit. The mothers were also given instructions to ensure a good recording; they were instructed on the optimal position of the microphones if adjustments were needed, and were made aware of potential noise that could arise from the activities that would affect the recording. They were also reminded to avoid talking at the same time as the child. Noise from various sources such as traffic and electric fan was attenuated and kept to a minimum. Parents were also instructed to use only English to interact with their children, in order to avoid a bilingual mode (Grosjean, 2011), and to speak as they would normally with their child; minimal use of Mandarin, if any, was found in their interactions in the recordings.

### 2.4. *Auditory and acoustic analysis*

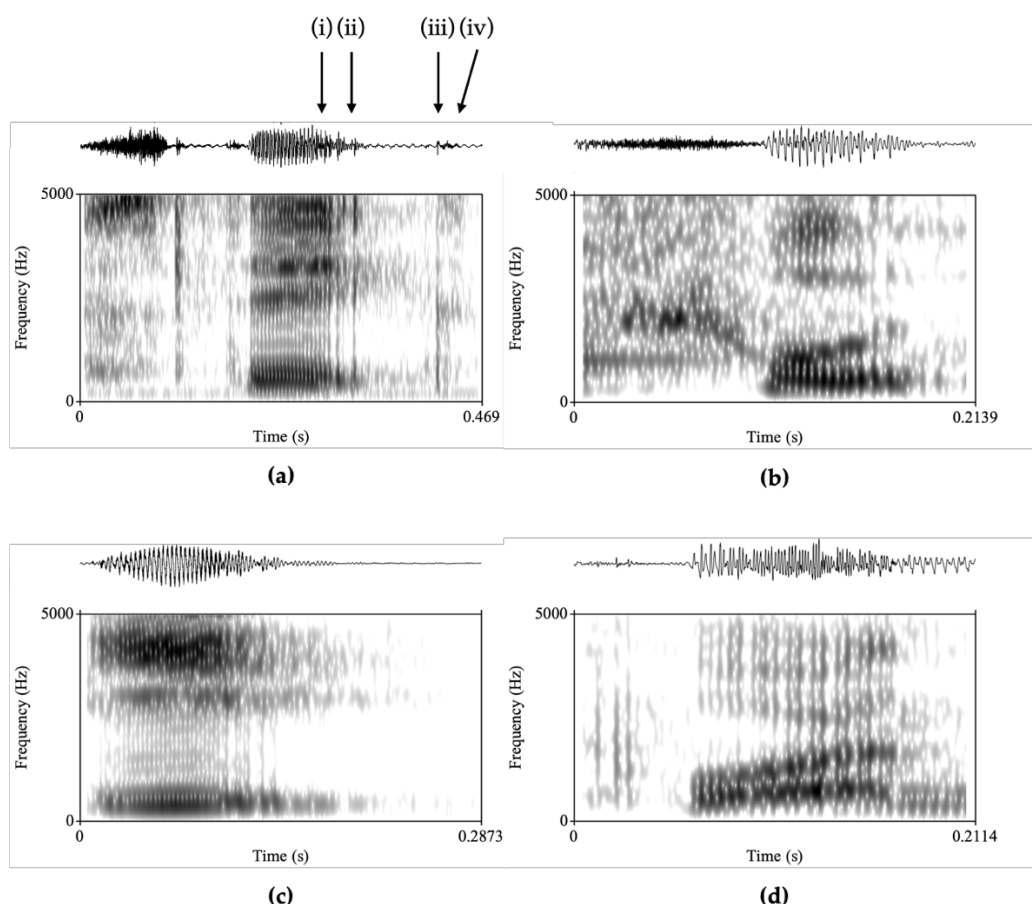
All word-final singleton oral stops in monosyllabic and stress-final disyllabic content words in the corpus were extracted, but bilabial stops /b, p/ were subsequently removed from further analysis due to their small number ( $n=71$ ) and unequal distribution according to phonetic environments between mothers, and thus the analysis comprised of only alveolar /t, d/ and velar /k, g/ stops. The target stops were also categorised according to their following phonetic environment: they occurred either before pauses (e.g. *that cat.*), vowels (e.g. *cat is*) or before consonant-initial words (e.g. *cat fell*). Homorganic stops were excluded. Table 2 shows a breakdown of tokens produced by each subject, categorised by their following phonetic environment. Since the materials yielded an inadequate number of pre-vocalic stops for statistical analysis, and since pre-vocalic (PV) and pre-pausal (PP) coda stops have been found to be released equally frequently in previous studies as mentioned above, they were grouped together in the analysis, to be compared with pre-consonantal stops (PC). Unusable tokens such as those of poor acoustic quality or those with ambiguous stop bursts were discarded (see below for the acoustic cues that were used in the analysis). The final set of data contained 700 adult tokens ( $M = 50$ ,  $SD = 8.06$ , range = 40–66) and 339 child tokens ( $M = 24$ ,  $SD = 7.45$ , range = 15–39).

**Table 2.** Number of tokens analysed according to mother-child pairs and phonetic environments, including pre-vocalic and pre-pausal (PV+PP), and pre-consonantal (PC) positions

Pair ID	Mother			Child		
	PV + PP	PC	Total	PV + PP	PC	Total
C5	34	20	54	20	10	30
C9	30	23	53	15	8	23
C17	26	18	44	20	19	39
C18	38	15	53	9	12	21
C20	21	19	40	7	8	15
C24	38	26	64	23	13	36
C30	35	31	66	13	7	20
C31	23	22	45	13	12	25
C35	15	25	40	13	6	19
C39	25	18	43	14	10	24
C46	39	14	53	5	17	22
C47	30	14	44	11	7	18
C55	36	13	49	9	6	15
C74	30	22	52	19	13	32
Total			700			339

The various realisations of the coda stops were identified manually by the first author using both aural cues and acoustic cues in the waveforms and spectrograms on Praat (v. 6.1.6; Boersma & Weenink, 2019). In connected speech, stops may be dropped entirely by adults, but their

omission could also be developmental in the case of very young children, and therefore stops that were dropped were categorised separately for the initial analysis. Representative spectrograms and waveforms of the possible realisations of coda stops in Singapore English (i.e. released, unreleased, and glottal stop replacement), as well as those that were dropped, are shown in Figure 1, (a)–(d) respectively. The relevant acoustic events that were used in the identification were those defined in Miles, Yuen, Cox, & Demuth (2016) and Song et al. (2012), which are also shown in Figure 1: (i) presence of a coda stop: observable formant transitions at the end of the vowel from the vowel steady state, based on the effects the different stops have on the F2 and F3 of the vowels (Thomas, 2011); (ii) glottalisation: presence of creaky voice, shown by glottal irregularity near the end of vowel as indicated by aperiodicity in the spectrogram and irregular spikes of energy in the waveform; (iii) coda burst: characterised by an abrupt spike in the waveform and a strong energy transient on the spectrogram; and (iv) post-release noise: high energy aperiodic frication in waveform and on the spectrogram. The identification of the different realisations was first done aurally and then confirmed by the absence or presence of key acoustic events: if there was at least one release burst with or without post release noise, or frication if the stop was replaced by an affricate, or if it was replaced by an ejective (i.e. [t][t̚][t̚s][t̚ʰ]), the token was labelled as ‘released’. If there was an absence of burst/noise that indicated a release but with formant transitions indicating the presence of a stop (i.e. [t̚]), it was coded as ‘unreleased’. Stops with flat periodicity and formant structure were coded as a ‘glottal’ stop [ʔ], or ‘dropped’, based on the presence of creaky voice. A second rater, a phonetician who was not involved in this project, was trained in the coding and asked to rate 100 randomly selected tokens (about 10% of all tokens). As the cues for released stops were reliable and their identification was straightforward, tokens that were coded as ‘released’ were excluded from the random selection of the 100 tokens. The rater was therefore asked to rate whether the 100 tokens were dropped, unreleased, or replaced by a glottal stop. 88% of the tokens were in agreement. Cohen’s kappa was computed to assess the agreement; there was substantial agreement between the raters,  $\kappa = .77$  (95% CI, .65 to .89),  $p < .0001$ .



**Figure 1** Representative waveforms and spectrograms taken from adult speech data for **(a)** released coda stop in the word *steak* (with monophthongisation of /eɪ/) [steʔk], **(b)** unreleased coda stop in the *food* [fuʔd̪], **(c)** glottal stop replacement in the word *eat* [iʔ], and **(d)** dropped coda stop in the merged words *put on* [pʊʔn]. Acoustic cues: (i) coda stop transition, (ii) glottalisation, (iii) burst, and (iv) post-release noise

## 2.5. Statistical analysis

Mixed-effects logistic regression analyses were conducted using R statistical software (R Core Team, 2020) and the lme4 package (Bates et al., 2015). The specific response variable and the fixed and random effects included in the models are described below. For all models, to evaluate the contribution of each predictor, and to arrive at a more restricted model, pairwise model comparisons between a full model that included all the explanatory variables and a more restricted model that excluded the predictor under consideration were performed using likelihood ratio tests.

## 3. Results

### 3.1. Overall means

The distribution of the four realisations for each mother and children is first examined. Table 3 presents the overall means for each mother and child. Based on the gross averages, for mothers, there were more stops that were not released: 44% of the stops were unreleased, 15% were replaced by glottal stops, and 33.7% were released. Children, contrastingly, produced more released stops (52.5%) than unreleased stops (35.1%) and glottal stops (7.4%). In addition, mothers dropped 7.3% ( $n=51$ ) of all stops, and children dropped 5% ( $n=17$ ). As expected, many cases of final consonant deletion (48 for mothers and 5 for children) were elisions due to connected speech processes, resulting in the merger of words (e.g. [smn] *sit on*). No more than three stops per child were dropped, and thus the other 12 child tokens that were dropped were likely speech errors rather than due to developmental delay. All children in this study can therefore be regarded to have acquired the full coda structure. As predicted, individual results in Table 3 show that both mothers and children vary considerably in how frequently coda stops were released; the average stop release for mothers ranged from 5.8% to 69.4% and for children, from 21.1% to 90.9%.

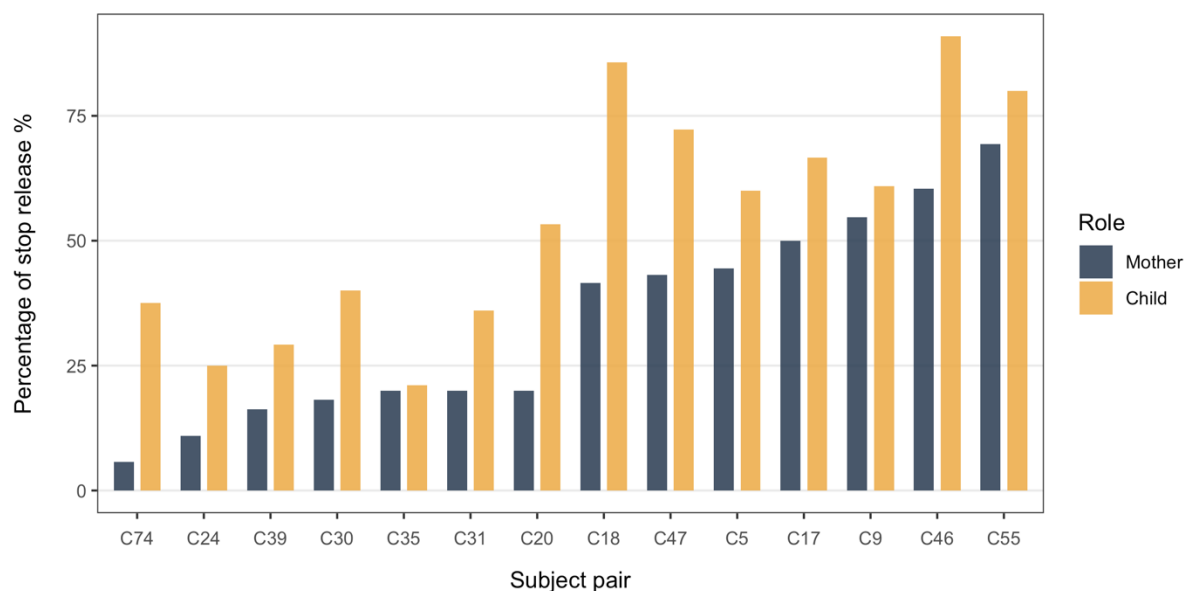
**Table 3.** Overall percentages of coda stops that were released, unreleased, produced as glottal stop and dropped by each mother-child pair

Pair	Released		Unreleased		Glottal stop		Dropped	
	Mother	Child	Mother	Child	Mother	Child	Mother	Child
C5	44.4	60.0	42.6	30.0	7.4	6.7	5.6	3.3
C9	54.7	60.9	26.4	30.4	15.1	8.7	3.8	0
C17	50.0	66.7	36.4	17.9	11.4	10.3	2.3	5.1
C18	41.5	85.7	32.1	9.5	22.6	4.8	3.8	0
C20	20.0	53.3	65.0	33.3	12.5	0	2.5	13.3
C24	10.9	25.0	56.2	41.7	18.8	22.2	14.1	11.1
C30	18.2	40.0	57.6	25.0	24.2	25.0	0	10.0
C31	20.0	36.0	71.1	56.0	4.4	8.0	4.4	0
C35	20.0	21.1	42.5	63.2	25.0	5.3	12.5	10.5
C39	16.3	29.2	46.5	66.7	14.0	0	23.3	4.2
C46	60.4	90.9	18.9	9.1	15.1	0	5.7	0
C47	43.2	72.2	27.3	27.8	20.5	0	9.1	0
C55	69.4	80	24.5	20.0	2.0	0	4.1	0
C74	5.8	37.5	67.3	53.1	13.5	0	13.5	9.4

The overall means of stop release by mother-child pairs are further presented graphically in Figure 2, in increasing order of mothers' production. A positive association between mother and child overall production patterns can also be observed in the figure; mothers who released coda stops to a lesser degree also had children who tended to not release their stops, and the same is true for mothers who released their stops to a higher degree. A correlation test was performed on



the means of overall stop release between children and parents. Due to the small sample size, a non-parametric correlation, Kendall's tau, was used. The overall percent release of coda stops of the children significantly correlated to the percent release of coda stops of the mothers,  $\tau = .58$ ,  $p = .004$ .



**Figure 2** Distribution of overall percentages of stop release by caregiver-child pairs.

### 3.2. Realisations of coda stops according to phonetic environment and place of articulation

As the realisations of coda stops in Singapore English are also influenced by phonetic environment and PoA, the percentages of the three main realisations (i.e. excluding dropped tokens) as a function of these factors are presented graphically in box plots and violin plots in Figure 3. Individual observations of all subjects were included. Visual inspection of Figure 2 revealed two groups of mothers, with the division falling between participants C20 and C18; one group of mothers released coda stops less frequently, below 25% of the time, while mothers in the other released coda stops more frequently. For the sole purpose of visual comparison in Figure 3, the individual observations were grouped into two groups: '(L)ower' for mothers (and their children) that released coda stops less frequently, and '(H)igher' for the group that released coda stops more frequently. The outlines of the violin plots illustrate the kernel probability density, which is the proportion of the data located at a particular point, with thicker parts representing higher frequency of sample points. To assess the effects of role (mother or child), PoA and phonetic environment on the realisations of coda stops, three separate mixed-effects generalised regression models, one for each of three main realisations, were run on all tokens. Each model

included role, PoA and phonetic environment and all their two-way interactions as fixed effects, subjects and tokens as random effects, and the binary outcome of the realisation of focus as the response variable.

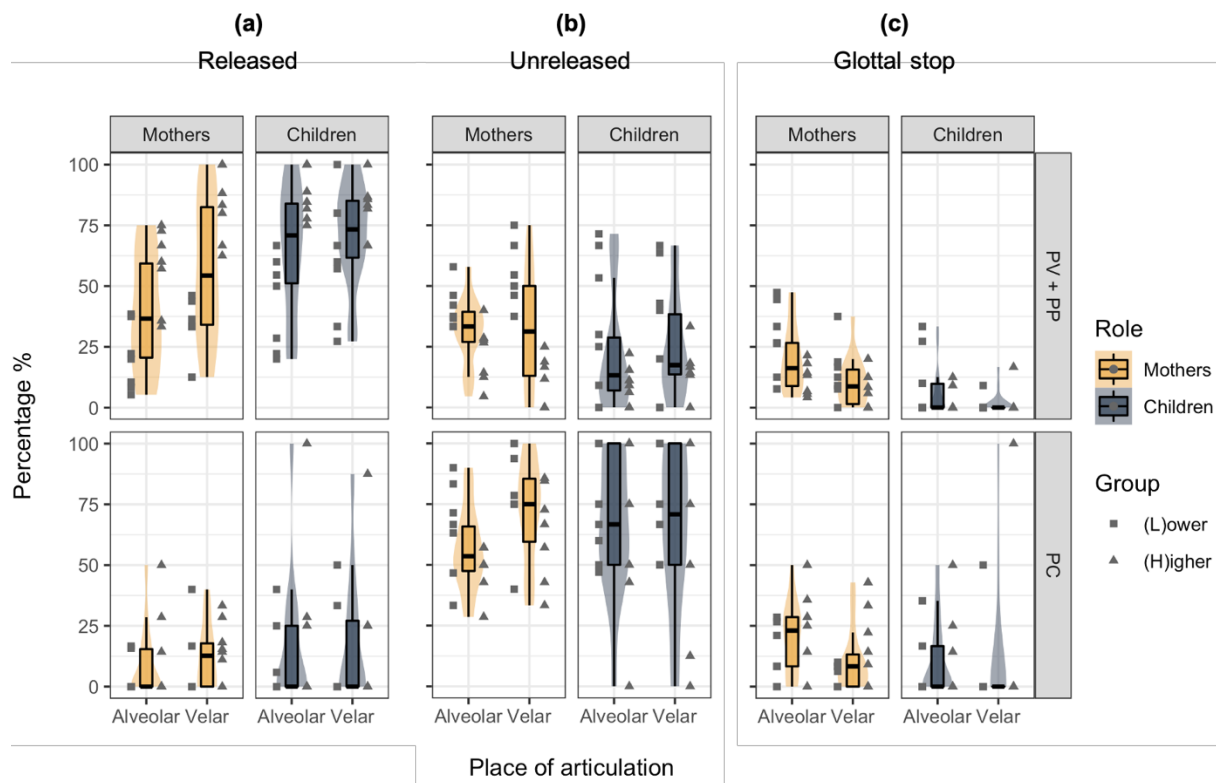
Stops that were released were first examined. In the best-fitting model that performed significantly better than an intercept-only baseline model ( $\chi^2(3) = 227, p < .001$ , marginal  $R^2 = .29$ , conditional  $R^2 = .47$ ), the three main effects were significant predictors; PV+PP stops were more likely to be released than PC stops,  $B = 2.41$ ,  $OR = 11.12, p < .001$ , 95% CI [7.41, 16.69], and velar stops were more likely to be released than alveolar stops,  $B = 0.88$ ,  $OR = 2.42, p < .001$ , 95% CI [1.52, 3.85]. Children were also more likely to release their stops than mothers,  $B = 0.99$ ,  $OR = 2.70, p < .001$ , 95% CI [1.90, 3.82]. In Figure 3, it can be observed that the inter-speaker variation in the release of stops mentioned above is most pronounced for PV+PP stops, evinced by the large interquartile ranges and long whiskers of the boxplots, as well as the relatively uniform widths of the violin plots that indicate large spread. By visual inspection, the differences between (L) and (H) groups are consistent after effects of PoA and phonetic environment are considered, although less categorical than when comparing global averages. Across contexts, some mothers, mostly belonging to (H), still released more stops on average than other mothers, mostly belonging to (L). This is evinced by how, especially for PV+PP stops, the individual observations of (H) mothers cluster within the upper quartile of the boxplots, with many at or near the maximum; the converse is true for those in (L), with more falling below the median, and at or near the minimum of the range. Child production generally reflects this pattern, and the differences between (L) children and (H) children are also most evident in their production of PV+PP stops. While subjects may not fall neatly into the (L)/(H) groups across all contexts, it is evident that there is considerable inter-speaker variation with regard to the frequency of stop release, even after effects of PoA and phonetic environment were considered.

Unreleased stops were then examined. The best-fitting model with unreleased stops as the response variable performed significantly better than an intercept-only baseline model ( $\chi^2(3) = 117, p < .001$ , marginal  $R^2 = .14$ , conditional  $R^2 = .29$ ). The main effect of phonetic environment and its interaction with PoA were significant predictors. Compared to PV+PP stops, PC stops were significantly more likely to be unreleased,  $B = 1.29$ ,  $OR = 3.63, p < .001$ , 95% CI [2.42, 5.46]. That the interaction between PoA and phonetic environment is significant but not the main effect of PoA indicates that the difference between alveolar and velar stops is only significant when phonetic environment is considered. Specifically, velar PC stops were significantly more likely to be unreleased than alveolar PC stops,  $B = 0.70$ ,  $OR = 2.02, p = .025$ , 95% CI [1.09, 3.74].

Finally, glottal stops were analysed. A caveat is that not all children in this study produced glottal stops and only 25 of such tokens were recorded, and so the results may not be indicative of

patterns of Singapore children. In the best-fitting model that performed significantly better than an intercept-only baseline model ( $\chi^2(4)=33.69, p < .001$ , marginal  $R^2 = .13$ , conditional  $R^2 = .29$ ), the main effects of role, PoA, and the interaction between role and phonetic environment were significant predictors. Alveolar stops were significantly more likely to be replaced by glottal stop than velar stops,  $B = 1.08$ ,  $OR = 2.95$ ,  $p < .001$ , 95% CI [1.63, 5.35]. The significant interaction between role and phonetic environment reveals that, compared to mothers, children were more likely to replace PC stops with glottal stops than they did with PV+PP stops,  $B = 1.11$ ,  $OR = 3.05$ ,  $p = .03$ , 95% CI [1.12, 8.31].

In summary, the analysis revealed some systematicity in the distribution of realisations of coda stops in Singapore English in the adults, and children's production generally reflected these patterns. Effects of phonetic environment and PoA on stop release were found; PV+PP stops were released more often than PC stops, and velar stops were released more often than alveolar stops. Stops that were not released were mostly unreleased rather than replaced by glottal stops. This is especially so for PC stops, which were mostly unreleased. There was also a PoA effect on whether stops that were not released were unreleased or replaced by a glottal stop; alveolar stops were more likely to be replaced by glottal stops, while velar stops were more likely to be unreleased. There were however two main differences between the production of children and mothers. Firstly, children released significantly more stops than mothers. Second, while phonetic environment did not influence the likelihood of glottal stop replacements for mothers, children replaced more PC stops with glottal stops than they did for PV+PP stops. The analyses also revealed considerable inter-speaker variation in both mothers and children in how frequently stops were released.



**Figure 3** Box and violin plots of percentages of **(a)** released stops, **(b)** unreleased stops and **(c)** glottal stops as a function of role (left and right panels of each plot), phonetic environment (top and bottom rows) and place of articulation (left and right of each panel), with the inclusion of individual observations, grouped by (L) and (H). The outlines of the violin plots illustrate the kernel probability density (the proportion of the data located at a particular point).

### 3.3. Predictors of inter-speaker variation in stop release

Finally, the predictors of inter-speaker variation in stop release patterns between mothers and children were explored. Mixed-effects generalised linear regression analyses were run to model the binary outcome of stop release (i.e. released or not released) in the mother and child data, which were analysed separately. Subjects and tokens were added as random effects. Language-internal fixed-effects factors included the two previously explored factors: PoA and phonetic environment. Although Gut (2005) did not find effects of phonological voicing on the likelihood of stop release in her adult subjects, it was added into the two models as a fixed effect to confirm the findings. Language-external or lexical fixed-effects factors that may potentially condition the release of stops were included in the saturated models. For the mothers, these fixed effects included their age, the age of their children, SES, and their language dominance measured by the BLP. For the children, the language-external or lexical fixed-effects factors included their age, percentage English use, gender, SES, English vocabulary score and total vocabulary score. To

ascertain the input–production relationship, the mean stop release of their respective mothers was included in the child model. As PoA and phonetic environment influence stop release as shown above, the mean stop release of each mother specific to PoA (i.e. velar or alveolar) and phonetic environment (i.e. PV+PP or PC) was calculated, generating four averages per mother. Each individual child token was then compared with the respective specific mean of their mother (Mother\_%) according to the PoA and the phonetic environment of the child token, rather than the global average.

The results for the full models for mothers and children are presented in Table 4 and Table 5 respectively. In the model for mothers, PoA ( $\chi^2(1) = 10.3, p = .0013$ ), phonetic environment ( $\chi^2(1) = 127, p < .001$ ), and age of mothers ( $\chi^2(1) = 6.29, p = .012$ ) yielded significant improvement of model fit. The best-fitting mothers-only model confirms that velars were more likely to be released than alveolar stops,  $B = 1.15$ ,  $OR = 3.15, p < .001$ , 95% CI [1.66, 5.99], and PV+PP stops were more likely to be released than PC stops,  $B = 2.68$ ,  $OR = 14.56, p < .001$ , 95% CI [8.34, 25.40]. It also revealed a positive association between mother's age and stop release,  $B = 0.17$ ,  $OR = 1.19, p = .01$ , 95% CI [1.04, 1.36]. The BLP score was further broken down into its four individual components (i.e. language use, history, attitudes, and proficiency) and analysed in a separate model, but none of the components was a significant predictor. In the model for children, phonetic environment ( $\chi^2(1) = 5.1, p = .024$ ) and their mother's production ( $\chi^2(1) = 9.89, p = .0017$ ) contributed significantly to model fit. The best-fitting children-only model confirms that PV+PP stops were more likely to be released than PC stops,  $B = 1.34$ ,  $OR = 3.80, p = .004$ , 95% CI [1.55, 9.35], but the effect of PoA was not significant, due to the almost equally frequent release of alveolar stops, especially by children in the (H) group. The children-only model also revealed a positive association between mother's production and the likelihood of stop release,  $B = 0.03$ ,  $OR = 1.03, p < 0.001$ , 95% CI [1.01, 1.05].

**Table 4.** Regression coefficients of a saturated mixed-effects logistic regression model fit to the coda stops of mothers with stop release as response and subject and token as random effects

Fixed factors	<i>B</i>	<i>SE</i>	Odds Ratio	[95% CI]	<i>p</i>
(Intercept)	-6.43	2.56	0.00	[0.00 – 0.24]	<b>.0012</b>
PoA	0.91	0.30	2.48	[1.37 – 4.49]	<b>.003</b>
Phonetic environment	2.48	0.26	11.98	[7.17 – 20.03]	<b>&lt; .001</b>
Voicing	0.58	0.32	1.79	[0.96 – 3.33]	.07
Age (child)	-0.03	0.03	0.97	[0.91 – 1.03]	.29
Age (mother)	0.17	0.06	1.18	[1.04 – 1.34]	<b>.009</b>
SES	-0.04	0.09	0.96	[0.81 – 1.14]	.64
BLP	-0.002	0.006	1.00	[0.99 – 1.01]	.67
Observations	700				
Marginal $R^2$	0.35				

Conditional $R^2$	0.52
-------------------	------

Note: CI = confidence interval. Reference category for PoA is alveolar, phonetic environment is PC, and voicing is voiced.

**Table 5.** Regression coefficients of a saturated mixed-effects logistic regression model fit to the coda stops of children with stop release as response and subject and token as random effects

Fixed factors	<i>B</i>	<i>SE</i>	Odds Ratio	[95% CI]	<i>p</i>
(Intercept)	-4.28	2.97	0.01	[0.00 – 4.67]	.15
PoA	-0.15	0.38	0.86	[0.41 – 1.83]	.70
Phonetic environment	0.96	0.42	2.61	[1.14 – 5.98]	<b>.02</b>
Voicing	0.24	0.39	1.27	[0.60 – 2.73]	.53
Age (child)	-0.01	0.03	0.99	[0.94 – 1.04]	.70
% English use	-0.01	0.04	0.99	[0.91 – 1.07]	.81
English vocab.	0.003	0.003	1.00	[1.00 – 1.01]	.28
Total vocab.	-0.002	0.001	1.00	[1.00 – 1.00]	.24
SES	0.15	0.10	1.16	[0.95 – 1.42]	.15
Gender	-0.80	0.55	0.45	[0.15 – 1.33]	.15
Mother_ %	0.04	0.009	1.04	[1.02 – 1.05]	<b>&lt; .001</b>
Observations	339				
Marginal $R^2$	0.31				
Conditional $R^2$	0.42				

Note: CI = confidence interval. Reference category for PoA is alveolar, phonetic environment is PC, voicing is voiced, and gender is female.

#### 4. Discussion

This study examined the coda stop release patterns of 14 Singaporean mother and child dyads, in order to uncover inter-speaker variation in the adults and to investigate the effects of such qualitative differences in the input on the development of coda stops of their children. The three hypotheses set out earlier predicted that while children as a group would exhibit adult-like patterns with regard to the distribution of realisations of coda stops, there would be individual variation in the frequency of stop release in the children that could be attributed to variation that would also be observed in their mothers. The findings of this study support all three hypotheses, which are summarised and discussed in turn.

Our findings support the first hypothesis, as the overall production patterns of both mothers and children in this study generally reflected the local adult norms reported in Gut (2005). A caveat is that the specific quantitative information in Gut's study is not directly comparable with the findings of this study. This is because her sample comprised a different number of stops according to their PoA and phonetic environment, and further in her analysis

these two effects were analysed separately. Therefore, only general patterns reported in her study are discussed. In terms of stop release, gross averages revealed that, like most Singaporeans, mothers and children in this study released coda stops relatively less frequently than speakers of other standard varieties of English. However, children in this study were found to release stops more frequently than their mothers, which has also been reported in previous studies (e.g. Song et al., 2012). One reason could be the children's syntactically less complex utterances; children often produce words in isolation, and indeed children in this study produced 13.9% more pre-pausal tokens than mothers. Another reason that is pointed out by other studies could be biological or physiological, where the higher rate of stop release in children is attributed to the immature motor development and their smaller laryngeal airway, which results in greater subglottal and intraoral pressures (Imbrie, 2003; Song et al., 2012). A third possible reason (suggested by an anonymous reviewer) was that once the children in our study had begun to produce stops in an adult-like way, they overproduced/over-articulated them; it took longer for them to fully match adult models. This is similar to reports that children fail to reduce English unstressed vowels in an adult-like way until age six or later (Payne et al., 2012). Effects of PoA and phonetic environment reported by Gut (2005) were also observed in the regression models that included all mother and child tokens; pre-vocalic and pre-pausal stops were released more often than those before consonants, and velar stops were released more often than alveolar stops. The effect of PoA was, however, not observed in the children-only model, which is likely due to the almost equally frequent release of alveolar stops, especially by the children in the (H) group. No effect of voicing was found. In terms of the distribution of unreleased and glottal stops, stops that were not released were mostly unreleased (or inaudibly released), and compared to velar stops, alveolar stops were more likely to be replaced by glottal stops, and these patterns are also aligned with those found in Gut (2005). The findings here show that children's production patterns generally reflect those in the input. While some of these patterns could potentially be explained by other factors (e.g. for instance aerodynamics, where velar stops are released more often because of the smaller occlusion that results in a larger pressure build-up), features such as the predominance of unreleased stops are largely attributed to patterns in the input, as these are dialect-specific features.

The second hypothesis predicted variation in the frequency of stop release between mothers. We found that some mothers matched the rate of coda stop release of American and British adults and children reported in the mentioned studies, while others released the coda stops much less frequently, and to a degree similar to local norms, even after effects of PoA and phonetic environment were considered. Interadult variation in the speech of caregivers, however, can sometimes arise due to differences in the modifications made to their child-directed speech (CDS). For instance, mothers of much younger children or infants may exaggerate certain

phonological contrasts or use more canonical forms. In some bilectal contexts, more standard variants are used when interacting with younger children and girls, while more vernacular forms or local variants are used towards older children and boys (Foulkes & Docherty, 2006). Such effects of age and gender, however, were not found in this study. This is likely because the nonrelease of coda stops is an invariable feature and one without much sociolinguistic salience, given that the nonrelease of coda stop is a pervasive feature of Singapore English that is widely attested in all social strata of the community (Bao, 2003). In other words, for many Singaporeans, the release and nonrelease of stops are not alternative forms. Furthermore, in casual conversations, some mothers have been found to use predominantly nonstandard variants or local dialects even with young children of the same age group (e.g. Smith et al., 2007). Likewise, we expect mothers in this study to pay less attention to their speech in casual play with their children, and not adopt an alternative variant that deviates from their informal register. Therefore, the interadult variation observed in this study is very likely to be due to individual differences in the phonetic realisations of coda stops. A preliminary analysis performed to uncover potential determinants of the variation was inconclusive. The only significant language-external predictor was the age of mothers. However, the adults in this study were mostly from the same age group and therefore the differences are unlikely to be due to them belonging to different phases of the bilingual education policy or exhibiting age-graded language variation, nor are they a result of differences in length of exposure to English. The effect of age that is observed here may be contributed by factors at a more micro-level that were not considered in the analysis. For example, a factor in this study that correlated with age that may offer some explanation, interestingly, is their seniority in their jobs. All six of the oldest mothers, who were above 35 years old, held managerial positions in the middle to upper management that also involved frequent interactions with clients. Of the six, five belonged to the (H) group. The communicative demands of their jobs may perhaps have made them more aware of their speech features, and may have also motivated them to adopt phonetic features of exonormative standards that index positive meanings and stances that are crucial for their roles, such as standardness, education, or attention to detail. The other language-external factor, language dominance as measured by the BLP, was not found to be significant predictor of stop release, nor were the individual component scores. However, inspection of individual questions in the BLP revealed some differences in their language history that could be explored in future studies. It was mentioned previously that, due to language shifts, previous generations of Singaporeans differed considerably in their language backgrounds, and thus the input that later generations received may differ in both quantitative and qualitative ways. The responses in the BLP may suggest that (H) mothers were raised in a more English dominant environment; three (H) mothers only started learning Mandarin after 3;0, and four of them only started to feel comfortable using Mandarin in their teenage years.



The final hypothesis of the study tested the effects of such qualitative differences in the maternal input on their children's production. The analysis revealed a very strong statistically significant positive input–production relationship, even after effects of PoA and phonetic environment were considered; mothers who released coda stops to a lesser degree also had children who tended to not release their stops, and the same was true for mothers who released their stops to a higher degree. The variation observed is unlikely to be due to age-related effects, as supported by the regression analysis. Children as young as 1;6 have been found to exhibit adult-like cues in coda stop production, and by 2;6, production patterns closely approximate those of the adults (e.g. Demuth et al., 2009; Song et al., 2012). That children C18 and C55, who were 2;8 and the youngest in the group, released their stops frequently suggests that the nonrelease of stops of other children in this study was unlikely to be a result of biological or physiological constraints. Similarly, older children who released their stops less frequently than others, such as C74, C30 and C35, show that the nonrelease of stops was unlikely to be due to developmental delay or differences in the length of exposure to English input. Language dominance is also an unlikely determinant, as children in this study were all highly English dominant and matched in their amount of English use. In addition, effects of CLI are not expected, as Mandarin lacks coda oral stops, and previous studies have shown that both early and late L2 English learners were able to produce English singleton coda stops without much difficulties (e.g. Xu Rattanasone & Demuth, 2014; Xu & Demuth, 2012). Other language-external or lexical effects, such as SES and vocabulary sizes, were also non-significant predictors. The findings therefore strongly suggest that the main contributor of the variation observed in the children was the qualitative differences in the maternal input, corroborating the strong input–production relationship attested in previous studies (e.g. Stoehr, et al., 2019).

The findings of this study lend support to acquisition theories that focus on the role of the input. Previous studies have shown that infants are sensitive to within-category variation and that fine-grained distinctions are retained, based on their speech perception (e.g. Cristià, 2011; McMurray & Aslin, 2005), suggesting that the variation in acoustic realisations that are irrelevant to category membership is not ignored in the acquisition process, contrary to the assumption of more traditional theories of language acquisition. Given that children in this study acquired the same phonemes and phonological rules but differed in the phonetic implementation based on their mother's production, the findings suggest that young children are indeed sensitive to fine acoustic properties that are nonphonemic, and further these properties are also reflected in their production, suggesting that the source of input matters.

One question that arises is how variability in other sources of input such as that of their father, peers and other significant adults may have an influence on the phonological acquisition in children, and how they negotiate variable input. Variation at the societal level is commonplace in

multilingual contexts. In the case of Singapore, apart from the inter-ethnic and intra-ethnic variation that exists in the speech of peers and other significant adults such as teachers in the child care or nursery, children also hear foreign accents from the consumption of media. Previous studies have shown that it is important to consider the relative significance of the various models of input to the children. Parents are the primary sources of input in the formative years, and thus in this study we see strong correspondences between properties of the input and properties of the children's output. In mixed-dialects environments, children may adopt accent features of both fathers and mothers (e.g. Thomas and Scobbie, 2015). However, this input–production relation is often overridden by peer effects or community norms as children get older. In their adolescence, they gain a deeper awareness of sociocultural and appropriateness norms, and the speech models of their peers and the community become more significant to the children (e.g. Stanford, 2008). In situations where multi-lingualism or multi-dialectism is the norm, however, the individual may choose to retain their accent acquired from earlier models because of its value as a marker of a certain identity (e.g. Sharma, 2011; Sim, 2019). The study on the VOT production of two English-Italian-Spanish simultaneous trilingual sisters in the United States by Mayr & Montanari (2015) exemplifies this point. The two children heard English from their native English father and other native speakers from the larger native English-speaking community, Italian from their Italian-speaking mother and teachers and accented Italian from their English-dominant peers in the Italian school the two children attended, and Spanish from their monolingual nanny. Due to the 'major language effect', their English production was target-like, but their VOTs in Italian were not, perhaps due to the accented speech of their English-dominant peers. The children produced Spanish VOTs that were similar to the adult model, as the nanny was their significant input model. Such studies on multiple accents or foreign accent and their social pressures and influence on child phonological acquisition are sparse, and this clearly is an avenue for future research.

Taken together, the considerable between-speaker variation, as well as the strong input–production relationship attested in this study, echo the conceptual and methodological implication that a complete, accurate depiction of child production cannot be achieved by averaging group behaviours. While this is especially so for studies on multilingual populations, one must be equally cautious to assume group homogeneity by virtue of the adults being 'native' speakers, given that there can also be considerable individual variation (e.g. Cristià, 2010). We propose that, at the very least, child production studies should take the production patterns of the significant caregiver into account.

To conclude, the variability in the stop release of mothers contributes to the understanding of the complex linguistic environment in which children in multilingual contexts acquire their phonological representations. Through the use of a variable property in Singapore

English, this study has demonstrated the direct role of maternal input in phonological development, and has shown that the input effects extend to specific phonetic details. More importantly, the findings of this study show that variable production in children is not only due to differences in the quantity of input; qualitative aspects of the input also play a significant role.

## References

- Amengual, M. (2018). Asymmetrical interlingual influence in the production of Spanish and English laterals as a result of competing activation in bilingual language processing. *Journal of Phonetics*, 69, 12–28. <https://doi.org/10.1016/j.wocn.2018.04.002>
- Amengual, M., & Chamorro, P. (2015). The Effects of Language Dominance in the Perception and Production of the Galician Mid Vowel Contrasts. *Phonetica*, 72(4), 207–236. <https://doi.org/10.1159/000439406>
- Ball, M. J., Müller, N., & Munro, S. (2001). The acquisition of the Rhotic Consonants by Welsh-English bilingual children. *International Journal of Bilingualism*, 5(1), 71–86. <https://doi.org/10.1177/13670069010050010401>
- Bao, Z. (1998). The sounds of Singapore English. In J. A. Foley, T. Kandiah, Z. Bao, A. F. Gupta, L. Alsagoff, C. L. Ho, ... W. D. Bokhorst-Heng (Eds.), *English in new cultural contexts: Reflections from Singapore* (pp. 152–174). Singapore: Oxford University Press.
- Bao, Z. (2003). Social stigma and grammatical autonomy in nonnative varieties of English. *Language in Society*, 32(1), 23–46. <https://doi.org/10.1017/S0047404503321025>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Birdsong, D., Gertken, L. M., & Amengual, M. (2012). *Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism*. Retrieved from <https://sites.la.utexas.edu/bilingual/>
- Boersma, P., & Weenink, D. (2019). Praat: Doing phonetics by computer (Version 6.0.45). Retrieved from <http://www.praat.org/>
- Campbell, T. F., Dollaghan, C. A., Rockette, H. E., Paradise, J. L., Feldman, H. M., Shriberg, L. D., ... Kurs-Lasky, M. (2003). Risk factors for speech delay of unknown origin in 3-year-old children. *Child Development*, 74(2), 346–357. <https://doi.org/10.1111/1467-8624.7402002>
- Cristià, A. (2010). Phonetic enhancement of sibilants in infant-directed speech. *The Journal of the Acoustical Society of America*, 128(1), 424–434. <https://doi.org/10.1121/1.3436529>

- Cristià, A. (2011). Fine-grained variation in caregivers' /s/ predicts their infants' /s/ category. *The Journal of the Acoustical Society of America*, 129(5), 3271–3280.  
<https://doi.org/10.1121/1.3562562>
- Currie, C. E., Molcho, M., Boyce, W., Holstein, B., Torsheim, T., & Richter, M. (2008). Researching health inequalities in adolescents: The development of the Health Behaviour in School-Aged Children (HBSC) Family Affluence Scale. *Social Science & Medicine*, 66(6), 1429–1436. <https://doi.org/10.1016/j.socscimed.2007.11.024>
- Deterding, D. (2007). *Singapore English*. Edinburgh University Press.  
<https://doi.org/10.3366/edinburgh/9780748625444.001.0001>
- Demuth, K., Shattuck-Hufnagel, S., Song, J. Y., Evans, K., Kuhn, J., & Sinnott-Armstrong, M. (2009). Acoustic cues to stop coda voicing contrasts in 1 to 2-year-olds' American English. *The Journal of the Acoustical Society of America*, 125(4), 2570–2570.  
<https://doi.org/10.1121/1.4783744>
- Dodd, B., Holm, A., Hua, Z., & Crosbie, S. (2003). Phonological development: A normative study of British English-speaking children. *Clinical Linguistics & Phonetics*, 17(8), 617–643.  
<https://doi.org/10.1080/0269920031000111348>
- Dodd, B., Zhu, H., Crosbie, S., Holm, A., & Ozanne, A. (2002). *Diagnostic Evaluation of Articulation and Phonology (DEAP)*. London: Psychology Corporation.
- En, L. G. W., Brebner, C., & McCormack, P. (2014). A preliminary report on the English phonology of typically developing English-Mandarin bilingual preschool Singaporean children: English phonology of typically developing English-Mandarin Singaporean children. *International Journal of Language & Communication Disorders*, 49(3), 317–332.  
<https://doi.org/10.1111/1460-6984.12075>
- Fabricius, A. (2002). Ongoing change in modern RP: Evidence for the disappearing stigma of t-glottalling. *English World-Wide*, 23(1), 115–136.  
<https://doi.org/10.1075/eww.23.1.06fab>
- Fish, M. S., García-Sierra, A., Ramírez-Esparza, N., & Kuhl, P. K. (2017). Infant-directed speech in English and Spanish: Assessments of monolingual and bilingual caregiver VOT. *Journal of Phonetics*, 63, 19–34. <https://doi.org/10.1016/j.wocn.2017.04.003>
- Flege, J. E., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics*, 25(2), 169–186.  
<https://doi.org/10.1006/jpho.1996.0040>
- Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, 34(4), 409–438. <https://doi.org/10.1016/j.wocn.2005.08.002>

- Genesee, F., & Nicoladis, E. (2007). Bilingual First Language Acquisition. In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 324–342). Malden, MA ; Oxford: Blackwell Pub.
- Goldstein, B. A., Bunta, F., Lange, J., Rodriguez, J., & Burrows, L. (2010). The effects of measures of language experience and language ability on segmental accuracy in bilingual children. *American Journal of Speech-Language Pathology*, 19(3), 238–247.  
[https://doi.org/10.1044/1058-0360\(2010/08-0086\)](https://doi.org/10.1044/1058-0360(2010/08-0086))
- Grosjean, F. (2011). The Bilingual's Language Modes. In J. L. Nicol (Ed.), *Explaining linguistics. One mind, two languages: Bilingual language processing* (pp. 1–22). Malden: Blackwell Publishing.
- Guion, S. G., Flege, J. E., & Loftin, J. D. (2000). The effect of L1 use on pronunciation in Quichua–Spanish bilinguals. *Journal of Phonetics*, 28(1), 27–42.  
<https://doi.org/10.1006/jpho.2000.0104>
- Gut, U. (2005). The realisation of final plosives in Singapore English: Phonological rules and ethnic differences. In D. Deterding, A. Brown, & E. L. Low (Eds.), *English in Singapore: Phonetic Research on a Corpus* (pp. 14–25). McGraw Hill.
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26(1), 55–88. <https://doi.org/10.1016/j.dr.2005.11.002>
- Hua, Z. (2006). The Normal and Disordered Phonology of Putonghua (Modern Standard Chinese)-Speaking Children. In Z. Hua & B. Dodd (Eds.), *Phonological Development and Disorders in Children* (pp. 81–108). <https://doi.org/10.21832/9781853598906-007>
- Imbrie, A. K. K. (2003). Acoustical study of the development of stop consonants in children. *The Journal of the Acoustical Society of America*, 114(4), 2394–2395.  
<https://doi.org/10.1121/1.4778038>
- Kehoe, M., & Havy, M. (2018). Bilingual phonological acquisition: The influence of language-internal, language-external, and lexical factors. *Journal of Child Language*, 1–42.  
<https://doi.org/10.1017/S0305000918000478>
- Khattab, G. (2003). Age, Input, and Language Mode Factors in the Acquisition of VOT by English-Arabic Bilingual Children. In M.-J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (Vol. 15, pp. 3213–3216). Barcelona, Spain: Universitat Autònoma de Barcelona.
- Kirk, C., & Demuth, K. (2006). Accounting for Variability in 2-Year-Olds' Production of Coda Consonants. *Language Learning and Development*, 2(2), 97–118.  
[https://doi.org/10.1207/s15473341lld0202\\_2](https://doi.org/10.1207/s15473341lld0202_2)

- Lauro, J., Core, C., & Hoff, E. (2020). Explaining Individual Differences in Trajectories of Simultaneous Bilingual Development: Contributions of Child and Environmental Factors. *Child Development*, 91(6), 2063–2082. <https://doi.org/10.1111/cdev.13409>
- Law, N. C. W., & So, L. K. H. (2006). The relationship of phonological development and language dominance in bilingual Cantonese-Putonghua children. *International Journal of Bilingualism*, 10(4), 405–427. <https://doi.org/10.1177/13670069060100040201>
- Lleó, C. (2016). *Bilingualism and Child Phonology*. <https://doi.org/10.1093/oxfordhb/9780199935345.013.53>
- Mayr, R., & Montanari, S. (2015). Cross-linguistic interaction in trilingual phonological development: The role of the input in the acquisition of the voicing contrast. *Journal of Child Language*, 42(5), 1006–1035. <https://doi.org/10.1017/S0305000914000592>
- McMurray, B., & Aslin, R. N. (2005). Infants are sensitive to within-category variation in speech perception. *Cognition*, 95(2), B15–B26. <https://doi.org/10.1016/j.cognition.2004.07.005>
- Miles, K., Yuen, I., Cox, F., & Demuth, K. (2016). The prosodic licensing of coda consonants in early speech: Interactions with vowel length. *Journal of Child Language*, 43(2), 265–283. <https://doi.org/10.1017/S0305000915000185>
- Milton, J., & Treffers-Daller, J. (2013). Vocabulary size revisited: The link between vocabulary size and academic achievement. *Applied Linguistics Review*, 4(1). <https://doi.org/10.1515/applirev-2013-0007>
- Payne, E., Post, B., Astruc, L., Prieto, P., & Vanrell, M. del M. (2012). Measuring Child Rhythm. *Language and Speech*, 55(2), 203–229. <https://doi.org/10.1177/0023830911417687>
- Phoon, H. S., Abdullah, A. C., & Maclagan, M. (2013). The Consonant Realizations of Malay-, Chinese- and Indian-Influenced Malaysian English. *Australian Journal of Linguistics*, 33(1), 3–30. <https://doi.org/10.1080/07268602.2013.787902>
- Post, B., & Jones, S. (2020). Similarity and contrast in L1 pronunciation attrition in bilinguals. In C. Gabriel, A. Pešková, & M. Selig (Eds.), *Contact, variation and change in Romance and beyond*. Berlin: Schmidt, Erich Verlag.
- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Scarpino, S. E. (2011). *The Effects of Language Environment and Oral Language Ability on Phonological Production Proficiency in Bilingual Spanish-English Speaking Children*. ProQuest LLC.
- Sharma, D. (2011). Style repertoire and social change in British Asian English. *Journal of Sociolinguistics*, 15(4), 464–492. <https://doi.org/10.1111/j.1467-9841.2011.00503.x>

- Sim, J. H. (2019). "But you don't sound Malay!": Language dominance and variation in the accents of English-Malay bilinguals in Singapore. *English World-Wide*, 40(1), 79–108. <https://doi.org/10.1075/eww.00023.sim>
- Smith, J., Durham, M., & Fortune, L. (2007). "Mam, my trousers is fa'in doon!": Community, caregiver, and child in the acquisition of variation in a Scottish dialect. *Language Variation and Change*, 19(01). <https://doi.org/10.1017/S0954394507070044>
- Song, J. Y., Demuth, K., & Shattuck-Hufnagel, S. (2012). The development of acoustic cues to coda contrasts in young children learning American English. *The Journal of the Acoustical Society of America*, 131(4), 3036–3050. <https://doi.org/10.1121/1.3687467>
- Sorenson Duncan, T., & Paradis, J. (2016). English Language Learners' Nonword Repetition Performance: The Influence of Age, L2 Vocabulary Size, Length of L2 Exposure, and L1 Phonology. *Journal of Speech, Language, and Hearing Research: JSLHR*, 59(1), 39–48. [https://doi.org/10.1044/2015\\_JSLHR-L-14-0020](https://doi.org/10.1044/2015_JSLHR-L-14-0020)
- Stanford, J. N. (2008). Child dialect acquisition: New perspectives on parent/peer influence. *Journal of Sociolinguistics*, 12(5), 567–596. <https://doi.org/10.1111/j.1467-9841.2008.00383.x>
- Stites, J., Demuth, K., & Kirk, C. (2004). Markedness vs. Frequency Effects in Coda Acquisition. *Proceedings of the 28th Annual Boston University Conference on Language Development*, 2, 565–576. Somerville, Massachusetts: Cascadilla Press.
- Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2019). Bilingual Preschoolers' Speech is Associated with Non-Native Maternal Language Input. *Language Learning and Development*, 15(1), 75–100. <https://doi.org/10.1080/15475441.2018.1533473>
- Tan, S. H. (2011). Multilingual infant vocabulary development in Singapore. In *Multilingual Norms*. Frankfurt Peter Lang GmbH, Internationaler Verlag der Wissenschaften.
- Thomas, E. R. (2011). *Sociophonetics: An introduction*. New York: Palgrave Macmillan.
- Thomas, S., & Scobbie, J. M. (2015). Mixed Accents: Scottish Children with English Parents. In J. A. Mompean & J. Fouz-González (Eds.), *Investigating English Pronunciation: Trends and Directions* (pp. 83–104). London: Palgrave Macmillan UK. [https://doi.org/10.1057/9781137509437\\_4](https://doi.org/10.1057/9781137509437_4)
- Unsworth, S. (2016). Quantity and quality of language input in bilingual language development. In E. Nicoladis & S. Montanari (Eds.), *Bilingualism across the lifespan: Factors moderating language proficiency*. (pp. 103–121). <https://doi.org/10.1037/14939-007>
- Vihman, M. M., Kay, E., de Boysson-Bardies, B., Durand, C., & Sundberg, U. (1994). External sources of individual differences? A cross-linguistic analysis of the phonetics of mothers' speech to 1-yr-old children. *Developmental Psychology*, 30(5), 651–662. (1995-01251-001). <https://doi.org/10.1037/0012-1649.30.5.651>



- Wrembel, M., Marecka, M., Szewczyk, J., & Otwinowska, A. (2019). The predictors of foreign-accentedness in the home language of Polish–English bilingual children. *Bilingualism: Language and Cognition*, 22(2), 383–400.  
<https://doi.org/10.1017/S1366728918000044>
- Xu, N., & Demuth, K. (2012). Acoustic Analysis of English Coda by Mandarin Learners of English. *Proceedings of the 14th Australasian International Conference on Speech Science and Technology*, 41–44. Sydney, Australia.
- Xu Rattanasone, N., & Demuth, K. (2014). The acquisition of coda consonants by Mandarin early child L2 learners of English. *Bilingualism: Language and Cognition*, 17(3), 646–659.  
<https://doi.org/10.1017/S1366728913000618>