LINGUISTIC SOUND SYMBOLISM AND READING DEVELOPMENT: SOUND-SHAPE MATCHING AND PREDICTORS OF READING IN MULTILINGUAL SINGAPORE

FEI TING WOON

SCHOOL OF SOCIAL SCIENCES

NANYANG TECHNOLOGICAL UNIVERSITY

2018

LINGUISTIC SOUND SYMBOLISM AND READING DEVELOPMENT: SOUND-SHAPE MATCHING AND PREDICTORS OF READING IN MULTILINGUAL SINGAPORE

FEI TING WOON

SCHOOL OF SOCIAL SCIENCES

A thesis submitted to the Nanyang Technological University in partial fulfilment of the requirement for the degree of Masters of Arts

2018

ACKNOWLEDGEMENTS

It all started in high school.

I stumbled upon the profession of a speech therapist while volunteering at a school for children with cerebral palsy. The discovery led me to a degree programme in linguistics, during which I developed a mild obsession with how our brains process languages. How do English speakers read aloud when *ghoti* can be pronounced just like 'fish' and how do Singlish speakers use *leh* in the right contexts all the time? Eventually, I moved on and began asking different questions while remaining a linguist at heart. Now, I have come to the end of this season's exploration where I delved into crossmodal sensory processes that are involved in language.

I would like to express heartfelt gratitude to my supervisor, Dr Suzy Styles, for introducing me to this realm of psychology and for pushing me to discover new ways of looking at the language puzzle. I would like to thank her for her patience, humour, and encouragement over the last two years as I slowly broke away from my 'Chomskian' roots and brushed up on my statistical analysis skills.

To my loved ones who have showered love and support on me in various ways, words fail me at this point; I cannot thank you all enough.

Above all, I thank God, for, in Him, all things are held in its place.

It does not all end with graduate school.

Contents

Abstract

The evidence of sound-symbolism (sounds are linked systematically with other forms or referents) is well-attested with the bouba-kiki paradigm. Adults reportedly prefer matching sounds like voiceless consonants and high front vowels like /t, k, i, e/ with spiky shapes and voiced consonants and low back vowels like /b, m, l, o, u, a/ with curvy shapes. Drijvers et al (2015) tested adults with dyslexia using a bouba-kiki task and found that those adults made fewer congruent sound-symbolic choices; the researchers postulated that deficits in crossmodal processing underlie deficits in sound-symbol matching and in reading. In order to examine the developmental trajectory between sound-symbolism and reading, we tested a large sample of pre-schoolers (5 y 10 m) with a novel bouba-kiki task (the Alien Zoo task) and established norms. We then correlated the results with results from earlier measures of known predictors of reading that the same children completed at (24 m and 48m). We found no correlations and discuss our findings in relation to the research gap in reading development.

With the same group of children, we analysed archival data of the caregiver's language use with them. We present, for the first time, detailed descriptions of language environments of children growing up in Singapore and report different models of bilingualism in pre-verbal children tailored to the Singapore landscape. The mean language inputs of these models serve as a better guideline to classify bilingual child in Singapore. We found that these language input models explained the differences in the children's scores on standardised English language tests at age 4, but did not explain the differences in the Alien Zoo task at age 6. This suggests that the Alien Zoo task is not biased by the amount of English language exposure in Singaporean bilingual preschoolers.

Chapter 1: Introduction

Language: Iconic or Arbitrary?

Arbitrariness has long been recognised as a feature of human language, where there is no link between a combination of sounds put together (forming a word) and the item it refers to (Saussure,1916:1959). Iconicity in language, on the other hand, is where features of the form suggest meaning (Perniss & Vigliocco, 2014). For example, as English speakers, we may have an intuitive sense that the word ‘bowl’ goes better with the referent of a prototypical rounded shape container used to hold food than the word ‘KitKat’ and that the latter goes better with the referent of a light, crunchy chocolate wafer better than ‘bowl’. This ‘intuition’ is the basis of a form of iconicity found in spoken languages, known as sound-symbolism (for a review, see Lockwood & Dingemanse, 2015). In normal adults, the effect of sound-symbolism has been shown to be automatic (Spence & Parise, 2010), outside the realm of conscious processing (Hsieh, Hung, & Styles, 2016), and fast-acting (Lockwood et al., 2016).

Behavioural tests for sound-symbolism largely involve studying the mapping of consonants and vowels in relation to varied shapes and sizes of objects. It began with Köhler (1929:1947) and Sapir (1929) who examined the associations between speech sounds and perceived sizes of visual stimuli. Sapir reported the cross-modal association between the speech sounds /a/ and /i/ and size of objects. He observed that people associate the pseudo-word ‘mal’ with a larger exemplar of an imagined object and ‘mil’ with the smaller exemplar. More recently, these studies have been extended beyond shape and size to include different modalities like taste, texture and luminance (Deroy & Spence, 2016; Spence, 2011).

Sound-symbolism and its implications

Researchers of linguistic iconicity and sound symbolism suggest that a language system devoid of iconic links between word forms and human experience would be difficult to learn (Perniss, Thompson, & Vigliocco, 2010). Imai and colleagues designed a word-learning task using novel Japanese onomatopoeic words that described the manner of motions e.g. *chokachoka* to describe brisk walking with light steps (Imai, Kita, Nagumo, & Okada, 2008). They found that after receiving training, 25-month-old Japanese children could generalise these onomatopoeic words to new situations but could not do the same with non-sound-symbolic tokens. Imai and Kita (2014) also proposed the Bootstrapping Hypothesis, which they posit has implications for language acquisition and language evolution. According to them, humans have innate biases for sound-symbolic word-object mappings which scaffold later semantic and syntactic learning.

Neuroimaging research on sound-symbolism help provides insights into processing and integration across modalities. Using fMRI and fractional anisotropy (FA), Revill et al (2014) found increased activation in the left superior parietal cortex when participants responded to words they found sound-symbolic. In each trial, the participants were presented with a written antonym pair in a foreign language and a spoken word. They were asked to match the spoken word with one of the antonyms (which may or may not be sound-symbolic). Participants were not only able to pick up on sound-symbolic cues and match correct sound-to-form above chance, those who were more sensitive to sound-symbolic forms also showed increased FA in the left superior longitudinal fasciculus. The authors postulated that sound-symbolism activates areas of the brain that are engaged in cross-modal sensory processing.

The implications of sound-symbolism are important and research into this area would create more testable models for multisensory processing.

Sound-symbolism in behavioural studies: Bouba-Kiki paradigm

The original testing paradigm for sound-shape matching was developed by Köhler (1929:1947). Participants were presented with two shapes – one spiky and one curvy – and two pseudo-words – *maluma* and *takete* – and asked to decide which pseudo-word went better with which shape. This two-alternative forced choice paradigm was later adapted and popularised by Ramachandran and Hubbard (2001); replacing *maluma* with *bouba* and *takete* with *kiki*. Though the descriptions of the methods in their study were lacking, Ramachandran and Hubbard postulated that 95-98% of normal adults preferred matching words containing rounded vowels (/o/, /u/) and voiced bilabial consonants (/b/, /m/) with curvy shapes. Contrastively, adults preferred matching words containing voiceless stop consonants (/k/, /t/) and high front vowels (/i/, /e/) to spiky shapes. Given the high rates of agreement, this direction of preference is deemed 'congruent'. The effect of this sound-shape matching has since been found consistently with better-controlled studies, though the success rate of matching at 95% is not always replicated and perhaps only found when the speech sounds in the auditory stimuli are maximally distinct (Styles & Gawne, 2017a). More recent studies have attempted to identify the elements that may be driving this effect including particular phonemes, features e.g. prosody, orthography, and language backgrounds of participants (Cuskley, Simner, & Kirby, 2017; Fort, Martin, & Peperkamp, 2015; Nielsen & Rendall, 2011).

As it is, the graphemes 'b' and 'm' contain curved lines while 'k' is made up only of straight lines, and looks spiky. To examine if the sound-shape associations were conditioned due to exposure to the Roman script, Davis (1961) was the first to conduct a cross-cultural bouba-kiki paradigm with 8-14-year-olds Tanzanian children. More recently, Bremner et al. (2013) tested the paradigm with Himba people in Namibia. They found a significant effect for sound-symbol matching using the bouba-kiki paradigm even though the participants did not use written language and had little or no exposure to Western influences, hence ruling out the possibility of being influenced by Western orthography or culture.

However, there is no denying that specific linguistic influences and our experiences with language will influence our perception in a bouba-kiki task. Styles and Gawne (2017) found that speakers of Syuba in Nepal did not show a preference when asked to match pseudo-words with curvy and spiky shapes. They suggested that the bouba-kiki paradigm fails when auditory stimuli did not match the phonology inventory and phonotactics of the participants' native tongues.

Shang (2017) tested sound-shape matching using Mandarin Chinese tones on Mandarin, non-Mandarin and English-Mandarin bilingual speakers and found that patterns of sound-symbolic matches were influenced in ways that are consistent with the way the participants process Mandarin Chinese tones. For non-Mandarin speakers, they systematically matched Tone 1 with the spiky shape and Tone 3 with the curvy shape. For the Mandarin speakers, they matched Tone 1 with the curvy shape and Tone 4 the spiky shape. The English-Mandarin bilinguals in the experiment followed the same patterns of choices either as the Mandarin or the non-Mandarin speakers depending on the nature of the task. This shows that the effect of linguistic sound-symbolic matching is modulated sensory processing of one's language system. The sound-symbolic effect elicited with a bouba-kiki paradigm is a systematic effect that appears not be driven by orthography but is linked with language experience.

Bouba-kiki paradigm: atypical population and children

Individuals on the autistic spectrum make fewer congruent matches when tested with the bouba-kiki paradigm, suggesting impairment of cross-modal integration affecting the success rate of selecting congruent matches. Oberman and Ramachandran (2008) conducted the task with ten high-functioning individuals with autism and twenty age-matched neurotypical controls. The participants were presented with five pairs of audiovisual stimuli and were asked to point out which shape had which name e.g. *bouba/kiki, wow/bloop*. There was no significant difference between the groups on full-scale IQ and verbal IQ, and the autistic individuals showed no impairment in their perceptual processing and organisation in a singular modality (visual). However, for the bouba-kiki task, the autistic participants were guessing at chance while the neuro-typical children guessed significantly better than chance. The authors suggested that the lack of preference for the congruent audio-visual matches reflects impairment in cross-modal processing in autistic individuals.

In a separate study by Occelli et al (2013), high-functioning autistic participants made significantly fewer sound-symbolic matches than neuro-typical controls, although both groups made guesses above chance. Low-functioning autistic participants in the same study showed no preference for congruent sound-symbol matches by guessing at chance Occelli et al. (2013) posited that autistic individuals have global deficits in multisensory integration; affecting their choices in a bouba-kiki paradigm.

Studies with pre-literate children and pre-lingual infants posit that sound-symbol matching is unlearned and not solely dependent on experience with language. Maurer, Pathman and Mondloch (2006) replicated the bouba-kiki paradigm with 2.5-year-old children. They introduced a storyline to keep the children interested in the task. The experimenter began with, “My friend Mr Green Rabbit drew pictures of his favourite toys. He calls them funny names. One is called *Bamu* and the other is called *Kutey,"* and then showed the child the two line drawings side by side (one spiky and one curvy). The children were asked to pick which shape had which corresponding name. With their relatively little language exposure and small vocabulary size, the children still demonstrated a preference for congruent sound-shape matching. The authors theorised that the sound-symbolism effect was not driven by statistical learning from exposure to language solely, as the children guessed above chance just like adults whose lexicon stores are larger. Spector and Maurer (2013) also replicated the study using auditory stimuli with reduplicated sounds with no word-internal contrastive sounds e.g. *koko* instead of *kutey*, and also found that children preferred congruent sound-symbol matches.

With infants, the bouba-kiki paradigm has been modified with looking-time tasks and results have been mixed. Ozturk, Krehm and Vouloumanos (2012) conducted a study on 4-month old infants in Chile using a looking-time task. Infants were shown pictures of a spiky shape or a blobby shape while they listened to forty-second strings of *kiki* or *bubu* in congruent and incongruent conditions separately. The infants looked longer at the incongruent sound-shape pairs. The authors concluded that they preferentially matched sound and shape as they discriminated between congruent and incongruent pairs. Fort et al. (2013), however, found no preferential looking effects when they presented 4-month olds with two shapes side by side and played one auditory pseudo-word. The authors suggested that their complex experimental set-up was overwhelming for the infants and may have masked the sound-symbol effect.

The meta-analysis by Styles and Gawne (2017a) showed that the normed rate of congruent responses by typical adults is 84-94% when experimenters use pseudo-words made with canonical 'curvy' phonemes, /b, m, l, o, u/ and 'spiky' phonemes /k, t, i, e/. At present, however, there are no norms on the rate of congruent responses by children, while data from children do exist, the experiment paradigms are somewhat less controlled than in adult studies. For the two studies involving pre-literate toddlers, experimenters recited the pseudo-word for each trial (Maurer et al., 2006; Spector & Maurer, 2009). Nygaard, Herold and Namy (2009) presented evidence that adults exaggerate their infant-directed speech of pseudo-words with predictable prosodic cues to highlight differences in meaning e.g. higher pitch and shorter duration to convey the meaning of 'happy' compared to 'sad'. Hence, results from the studies with children could have been influenced by the prosodic cues given unintentionally by the experimenters. To reduce confounds in the bouba-kiki paradigm with children, bouba-kiki studies designed for children should use standard pre-recorded auditory stimuli. A norming study on a large sample of pre-reading children will help to establish the influence of language exposure on sound-symbolism.

Sound-symbolism and its relation to Dyslexia

For the majority of the population, the process of learning to read is accomplished with relative ease; a remarkable feat considering that writing systems are fairly recent inventions in human history as opposed to spoken communication (Lieberman, 2006). It is proposed that our brains are evolutionarily ready to acquire speech but not to learn to read (Lieberman & McCarthy, 2014). Learning to read an alphabetic language requires rehearsing and establishing associations between two low-level sensory processes – auditory (speech sounds) and visual (letters). Possible causes for reading deficits include deficits in auditory processing (Chung et al., 2008; Meyler & Breznitz, 2005; Ziegler & Goswami, 2005), disturbances to the visual pathway (Wang, Bi, Gao, & Wydell, 2010), and/or cross-modal auditory-visual processing (Blau, van Atteveldt, Ekkebus, Goebel, & Blomert, 2009; Blomert, 2011; Siok, Spinks, Jin, & Tan, 2009).

In the study by Blau et al. (2009), a group of dyslexic adults and normal controls were presented with letter-speech sound stimuli in congruent and incongruent conditions while in a fMRI scanner. Dyslexic adults showed significantly less activation of the superior temporal gyrus (STG) when presented with congruent letter-sound stimuli compared to the normal controls. More importantly, they also show little difference in STG activation when the stimulus was congruent and when it was not. This lack of suppression of STG activation when stimuli were incongruent is not found in normal adults, suggesting that for dyslexic adults who had equivalent activation for congruent and incongruent pairings, more effortful processing is required for successful reading. Results were replicated when the same task on was carried out with 9-year-old dyslexic children (Blau et al., 2009).

Froyen, Willems and Blomert (2011) measured event-related potentials (ERP) of 11-year-old dyslexic children when played speech sounds in isolation and with letters either simultaneously or 200 ms before speech. The dyslexic children did not show an early mismatched negativity, an effect which was shown in normal readers. The authors posit that for dyslexic readers, early influences of visual graphemes on speech sound are absent; suggesting that there was no automation of letter-sound integration despite four years of formal instruction in reading.

The study by Drijvers et al. (2015) remains the only one that has examined dyslexia with the bouba-kiki paradigm. The authors held a similar hypothesis to Occelli et al (2013): if auditory-visual abstraction underpins sound-symbolism, individuals with known low-level sensory deficits in auditory, visual, audio-visual linking processing will make significantly fewer sound-symbolic choices than normal controls. As hypothesised, adults with dyslexia in the study made significantly fewer congruent sound-symbol choices (61%) compared to the controls (73%). We note that the overall guessing rate for congruent choices by the controls was lower than studies on other normal adults. While the visual stimuli in the study were well-controlled for size and colour and were judged as distinctly curvy or spiky by separate participants, the auditory stimuli were pseudo-words containing phonemes that are not typical in bouba-kiki studies like /z, g, v, n, r/. These particular consonants have not been regularly used to elicit sound-symbolic effects. It is plausible that the lower rate of making congruent choices was due to the choice of speech sounds for the auditory stimuli.

While studies on dyslexic children have revealed deficits in auditory and auditory-visual processing, at present, we do not know much about the developmental trajectory of dyslexia in relation to sound-symbolism. Also, as dyslexia manifests as a difficulty in fluent word recognition and through weaker spelling and decoding abilities (Lyon et al., 1995), it is only diagnosed after some formal instruction of reading. While pre-literate children at risk for dyslexia have been shown to demonstrate difficulties in auditory processing (Goswami, Fosker, Huss, & Mead, 2011), an exploration into auditory-visual processing deficits with non-reading/letter based tasks have not been widely explored and should be conducted. The bouba-kiki paradigm adapted for children would serve as a good task for this purpose.

Alien Zoo task

While the bouba-kiki paradigm has been well-studied with adults, there are only a handful of studies regarding children's sound-symbol matching. Previous studies involving children reviewed earlier used rather uninteresting black and white line drawings as visual stimuli (Maurer et al., 2006; Oberman & Ramachandran, 2008; Occelli et al., 2013). We question the validity of some of the auditory stimuli e.g. 'Mmmmm' and 'Shhhhh' in Oberman & Ramachandran (2008) in their ability to elicit the expected bouba-kiki effect.

For the purposes of our study, we designed a novel sound-shape matching task – the Alien Zoo task – by adapting the bouba-kiki paradigm, we created exciting visual stimuli to capture and sustain the attention of young children. Auditory stimuli in the task were generated using canonically ‘curvy’ and ‘spiky’ speech sounds to maximise the sound-symbolic effect. We aim to norm the task on a large cohort of Singaporean pre-schoolers and to use this non-reading based task to explore the developmental trajectory of reading deficits and sound-symbolism.

Singaporean children: Language background and development

For our investigation into the developmental trajectory of sound-symbolism and its relations with dyslexia, we look at young children growing up in Singapore.

The children who participated in our study are from the largest longitudinal cohort study in Singapore, Growing Up in Singapore Toward (healthy) Outcomes (GUSTO). The GUSTO study is funded by the Singapore National Research Foundation (NRF) under its Translational and Clinical Research (TCR) Flagship Programme and administered by the Singapore Ministry of Health’s National Medical Research Council (NMRC) and its initial aim was to examine the diet of pregnant women and related developmental outcomes of their children (“Overview of GUSTO,” 2009). The researchers recruited pregnant women (range: 18-40, M = 30.4), who underwent their first-trimester antenatal dating ultrasound scan clinic at public maternity units, at the National University Hospital and KK Women's and Children's Hospital between June 2009 and September 2010. The women recruited are Singapore citizens or permanent residents who are of Chinese (55.9%), Malay (26.1%) or Indian (18%) ethnicity who had homogeneous parental ethnic background. A total of 1176 babies were delivered (Soh et al., 2014). In the years following birth, the mother and child dyads were invited for various physiological, neurological and behavioural tests before the child’s birthday annually.

As part of a collaborative effort with GUSTO, we processed and analysed previously collected and unprocessed archival data on language background of the children as reported by the parents in the Language Background Questionnaire (see Appendix 1). This analysis was conducted in exchange for research time with the children in the cohort who performed our sound-symbol matching task in the months leading to their sixth birthday (see Chapter 3).

In the archival dataset, parents were asked to provide information about language use with their child from all the caregivers at **6 months** and **18 months** in a Language Background Questionnaire. The questionnaire was designed by the GUSTO team, and the digital archive of the data collected had not been processed prior to our involvement.

At the moment, there are no published works on language backgrounds of Singaporean infants. There is a general assumption of bilingual home language environments (English and an official 'Mother Tongue') in accordance with the bilingual language policy in Singapore (Silver & Bokhorst-Heng, 2016). However, it is almost certain that not every child growing up in Singapore experiences an English-Mother Tongue input environment. It is fundamentally important to understand language backgrounds in home environments as language acquisition begins in the home. Thus, our analysis of the language backgrounds of the Singaporean children will not only help us characterise the children in our study, it will also provide broader insights into language backgrounds of Singaporean children in general.

Within the first year, auditory perception of the pre-lingual child tunes into, and biases towards, speech sounds in languages occurring in one's environment (Kuhl, 2011a). Studies on infant-directed speech show that people naturally hyper-articulate their vowels, speak in a higher pitch and with more positive affect when speaking to infants compared to adults (Cristia, 2013; McMurray, Kovack-Lesh, Goodwin, & McEchron, 2013). Speakers do so as they inherently pay attention to the linguistic ability and emotional needs of their audience (Burnham, Kitamura, & Vollmer-Conna, 2002).

There are even differences in the characteristics of infant-directed speech based on languages spoken by the mothers. Cantonese-speaking mothers have been shown to hyper-articulate tones and not vowels in their infant-directed speech to 3 and 6-month-olds (Xu Rattanasone, Burnham, & Reilly, 2013). Xu Rattanasone and her team postulated that because perceptual attunement for lexical tones – of which there are six in Cantonese – occurs before vowels (Mattock & Burnham, 2006), mothers alter their speech to highlight the variation among the tones instead of vowels to aid their children's acquisition of Cantonese. Mothers also have been shown to speak differently to infants at-risk and not at-risk for dyslexia; they do not hyper-articulate their vowels for the at-risk infants (Kalashnikova, Goswami, & Burnham, 2016). These infants showed poor sensitivity in an auditory perception task and the authors postulated that mothers may be modifying their speech to provide other cues for the at-risk infants who were not picking up on hyper-articulated vowels. These studies show that within the child's first year, caregivers' language use can vary considerably and this variation has an impact on language development. Therefore, it is crucial for us to examine the language use of the caregivers of the children in our study.

With the data from Language Background Questionnaire, we are able to describe in detail patterns of language use of caregivers with the children (in Chapter 2). We then used cluster analysis to find the most effective way to cluster the children, based on the pattern of exposure to their different languages.

For measures of language development of the children in the GUSTO cohort, parents were first asked to complete the Singapore version of the Communicative Developmental Inventories (CDI) when their child was 8 months and 24 months (Reilly et al., 1993; S. H. Tan, 2009). The CDI is a checklist of words through which parents indicated if their child spontaneously uttered those words. Each marked word is scored at 1 and total scores provided an indication of the size of expressive vocabulary of the children. The CDI tool has been normed and extensively studied with various populations for its validity and reliability (Arriaga, Fenson, Cronan, & Pethick, 1998; Law & Roy, 2008; Styles & Plunkett, 2009). Evidence suggests that high CDI scores correlate with better reading abilities at a later age.

At 48 months, the GUSTO children completed a variety of tests, including three tests which measure predictors of early reading abilities. Firstly, phonological awareness was measured via three subtests from the second edition of the Comprehensive Test of Phonological Processing (CTOPP) (R. Wagner, Torgesen, Rashotte, & Pearson, 2013). Phonological awareness refers to the ability to attend to and manipulate individual sounds of spoken words and it remains as one of the most powerful predictors of reading ability for young children (Frijters et al., 2011). Poor readers and young children at risk for dyslexia have been shown to have poorer phonological awareness (Snowling & Melby-lervåg, 2016; see Swanson et al., 2003 for a review). While deficits in phonological awareness in poor readers and children with dyslexia become less obvious after the first grade, weaker access to phonemes impinges on the children's ability in reading at a later age (De Jong & Van Der Leij, 2003). Secondly, in the archival dataset, receptive vocabulary was measured using the Peabody Picture Vocabulary Test IV (PPVT) (Dunn & Dunn, 2007). Children indicate their understanding of a target word by selecting an illustration which best represents the target word, read aloud to the child, out of four choices. An indication of word knowledge provided by the PPVT score is shown to be correlated with reading comprehension at a later age (Cunningham & Stanovich, 1997). Thirdly, in the archival data set, the children sat for the Lollipop test which was designed as a measure for school readiness, by testing children on their knowledge of colours, shapes, numbers, counting, position, spatial recognition, letters and writing (Chew, 1981).   sub-test of letter identification and writing has shown success in predicting early school achievement for reading (Chew & Morris, 1984; Forget-Dubois et al., 2007).

Tracking the same children from age 2 to 6, we have a unique combination of results from our novel Alien Zoo task and archival data for language backgrounds and predictors of reading. Our descriptive and clustering analysis and results of the language backgrounds at 6 and 18 months allow for the children in the GUSTO cohort to be characterised in detail. We are better able to understand what languages are used with the children and which languages are dominant. By clustering the children according to the patterns of language use that are most common in Singapore, we allow the data to inform us about what bilingual patterns of language use looks like with young children growing up in Singapore. This data-driven approach to classification is the first of its kind for bilingualism research in Singapore and provides a better basis for classification of Singaporean children. Because the GUSTO cohort is the largest longitudinal cohort study of children in Singapore, our analysis characterises Singaporean children in general.

For the first time, we present correlation analyses of sound-symbolism and predictors of reading and explore the gap in the literature regarding the developmental trajectory of dyslexia in relation to sound-symbolism. With the data from the same children tracked across different ages, we are also able to examine the possible effects of language backgrounds on the sound-symbolic choices.

Overview

In summary, the aims of this dissertation are to examine the following:

1. What are the patterns of language used by caregivers with children in Singapore as they were growing up?
2. Do children in Singapore demonstrate sound-symbolic effect with the bouba-kiki paradigm?
3. How does performance on the bouba-kiki task relate to known predictors of reading?

A detailed description of the language background of children in our study is presented in Chapter 2. The cluster analyses also presented in Chapter 2 reveals language use patterns of caregivers in Singapore and we gain insight into the language mix that children in Singapore are receiving during their first two years. Methods, results and analysis of our adaptation of the bouba-kiki paradigm – Alien Zoo task – is discussed in Chapter 3. Cross-examination of the children's scores on predictors of early reading abilities and the Alien Zoo task is also presented in Chapter 3 along with implications of cross-modal deficits and reading. We also present the analysis to examine group differences of language backgrounds on the Alien Zoo task. Limitations of the study and general discussion on overall findings are presented in the final chapter.

Chapter 2: Language Backgrounds of Children in Singapore

The official Singapore census identifies the Singapore population as consisting of three numerically significant ethnic groups: Chinese, Malay, and Indian. Since gaining independence in 1959, the state has adopted the use of four official languages: English; Mandarin Chinese; Bahasa Melayu; and Tamil. English is set as the official working language while the other languages are designated as ‘Mother Tongues’ of the respective ethnic groups (Sim, 2016).

Language planning policies in Singapore are largely upheld by the Ministry of Education and an English – Mother Tongue bilingualism is the foundation on which the nation’s language and education policies have been built upon since 1987 (Low & Brown, 2005). The national education policy defines an individual’s Mother Tongue on the basis of the official classification of one’s racial identity. The Mother Tongue Language Policy, adopted by mainstream schools, requires lessons to be conducted using these official languages. Students are required to take one Mother Tongue subject and have the rest of their subjects taught in English (Dixon, 2005). It is therefore unsurprising that success in the mainstream education system is known to be tied to a good command of the English language and literacy.

Among adult Singaporeans, English – Mother Tongue bilingualism is purported as commonplace, with English as the common medium for communication even though languages spoken in home environments may be entirely different. According to census data, for children between the ages of 5 – 14, about 50 percent of Chinese and Indians use English as their dominant home language. For the Malay population within the same age bracket, 26 percent use English as their home language; an increase from 2010 (Department of Statistics, 2015). Figure 1 shows the increase in the percentage of Singaporean children who use English as their home language from 2010 to 2015.

Figure 1. Increase in percentage of young Chinese, Malay and Indian Singaporeans (5-14 years old) using English as their dominant home language according to three largest racial groups.

Firstly, besides the information provided by the official population census, little is known about the home language environments of young children in Singapore. There has been no systematic study of languages used with infants growing up in Singapore. An individual's development of his/her language ability is heavily influenced by language exposure. A measure of home language environments and the amounts of exposure would help in our understanding of the language development of the children in our study (Bedore et al., 2012; Paradis, Emmerzael, & Sorenson Duncan, 2010).

Secondly, the proportion of languages heard by a child differs greatly from one household to another and broad definitions of bilingual vs. monolingual children may not be useful in our understanding and classification of Singaporean children. Presently, researchers of language development and deficits differ in their ways of classifying a Singaporean child as monolingual or bilingual. Singh and colleagues consider an infant as an English – Mandarin bilingual if the child has 30-70% exposure to Mandarin and the child is a Mandarin-speaking monolingual if exposure rate is 90% or above (Singh, 2017; Wewalaarachchi, Wong, & Singh, 2017) while other researchers differentiate bilingual children and determine their L1 by looking at the highest percentage of time of language use at home with parents and main caregiver, age of acquisition and proficiency levels instead of using a cut-off (Pua et al., 2017). For a better representation of bilingual vs. monolingual Singaporean child, we propose the use of data-driven classification.

We aim to examine the following:

1. What languages are used in the home environments by caregivers of pre-lingual and pre-literate children in Singapore?
2. What are the common patterns of language use by caregivers with these young children? Do the patterns of language exposure reflect the common patterns of language use in Singapore? What patterns of bilingual exposure are present?
3. Do differing language exposure patterns of English predict later English language ability?

In the following sections, we will present a detailed descriptive analysis of the language backgrounds in home environments of 6-month olds in Singapore. The children were tracked and their language backgrounds are analysed again at 18-months. These children are from the GUSTO cohort, and some of whom are also involved in our novel sound-symbolic experiment (the Alien Zoo) at 6 years old as outlined in the collaborative research agreement.

Data for this language background analysis was collected via a Language Background Questionnaire completed by parents. A description of the data handling and general trends in caregiver language use will be followed by clustering analyses carried out to identify common patterns of language use with Singaporean children at 6-months and 18-months. A discussion of the findings will conclude the chapter.

Study 1: Analysis of archival data

GUSTO Language Background Questionnaire

The children who participated in our study are from the largest longitudinal cohort study in Singapore, Growing Up in Singapore Toward (healthy) Outcomes (GUSTO). As we are interested in examining the developmental trajectory of reading acquisition (and its deficits) and sound-symbolism, we began with the examination of the language(s) used in home environments of the children. We analysed the data collected from their parents from a Language Background Questionnaire (see Appendix 1). The survey was designed by the GUSTO team and the data collected have since been archived in a digital format though it has not been further processed by researchers from GUSTO. As these children are tracked across ages, a large proportion of them goes on to be involved in our sound-symbol matching task at age 6.

As part of the ongoing birth cohort study, the questionnaire was sent to the parents of participating children at two time-points when the child was (i) 6-months old; and (ii) 18-months old. The aim of the survey was to examine the types and number of languages each child was exposed to as well as the amount of time the child was exposed to each language. The survey consisted of 13 questions and was completed by either a parent (usually the mother) or a caregiver of the child. The complete set of questions is presented in Appendix 1. Examples of questions include:

1. When you talk to your child, a. how often do you use English? ____% of the time
2. When your parents talk to your child, does your mother use: Mandarin ____%?

The questionnaires were sent home to be completed by the parents. Answers from completed and partially completed questionnaires were compiled and documented digitally by GUSTO researchers.

Information on the identity of the respondent was provided in the digital archive only for the 6-month time-point: 370 out of 444 responses (83%) were completed by the mother of the child, and 9 by the father, and the rest of the respondents left this question unanswered. Information regarding the identity of the respondent was not provided at the 18-month time-point. As the initial aim of the GUSTO research was set out to examine the health and diet of pregnant mothers and corresponding relations to development of their children, it is assumed that majority of surveys at the second time-point were also completed by mothers.

A better response rate for completed questionnaires was seen at 6 months; having a total of 444 responses. Figure 2 shows the number of responses collected for both time-points. A comparison of the language input of children across both time-points will be presented in a later section.

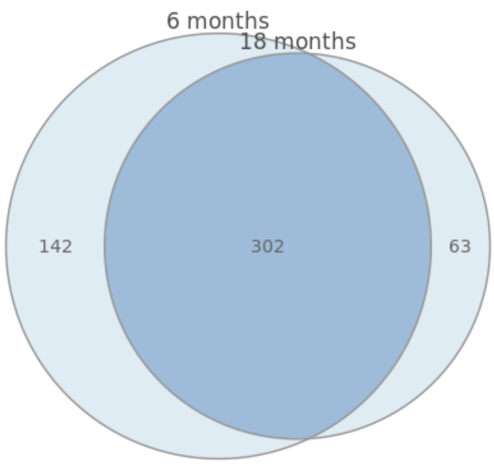


Figure 2. The number of participants at each time point is shown. Diagram was drawn using VennDiagram.net (Pires, 2017).

In Study 1, we will evaluate from the Language Background Questionnaire the following information:

1. How many languages are used with the child and what are the percentages of use?
2. Who are the caregivers and how much time does each caregiver spend with the child?
3. What is the parents' language use with 6-month olds like?
4. What is grandparents' language use with 6-month olds like?
5. What are the patterns of language use of domestic helpers and other caregivers with 6-month olds?

With the information listed above, we computed a Language Input Matrix for each child present in the analysis. This allows us to characterise the children to examine possible individual and group differences.

Methods

Data Handling

Language Backgrounds. Due to the nature of the structure of the questionnaire (Appendix 1), it may have allowed people to answer the specific questions in different ways based on their interpretation of the questions. Hence, manual processing of the data was necessary before it could be used for further analyses.

In the questionnaire, respondents were presented with a list of languages: English; Mandarin; Dialect; Malay; Tamil and Other, and were asked **when speaking to their child how often would they use the above languages** (answer to be given in percentage). The respondents were asked to answer identical questions about language use by **their parents** and **in-laws**, **domestic helpers** and any **other caregiver(s)** who were involved in taking care of the child.

We labelled both sets of grandparents with numbers (1-4). According to the question structure in the questionnaire, Grandparent (GP) 1 and 3 are grandmothers of the child while Grandparent 2 and 4 are grandfathers. As mentioned, we were not provided with the identity of the parent who completed the form at the 18-month time point. Anecdotally, the GUSTO researchers recount that most respondents (Parent 1 in the questionnaire) were the child's mother. Therefore, the corresponding GP 1 and 2 refers to maternal grandparents while GP 3 and 4 refers to paternal grandparents. We note that there is a certain degree of error with this classification based on our assumption. However, as subsequent analyses pool the data from different caregivers into a time-weighted input matrix according to the amount of time each person spends with the child and the amount of time they spend speaking each of their languages, the discrete identity of each caregiver is less relevant than the overall pattern of language input.

Language(s) used. If the percentage of 'Dialect' and 'Other' language were provided, respondents were asked to state specifically which dialects or languages they were. 'Dialect(s)' in the context of Singapore and in this questionnaire refer to common Chinese varieties spoken by Chinese Singaporeans e.g. Hokkien, Teochew and Cantonese (Sim, 2016). Languages listed in the archival questionnaire, therefore, reflect the aforementioned policies on language use in Singapore. To process the data when respondents listed Chinese dialects under 'Others' or other Indian languages like Hindi under 'Dialects', data for these cases were moved to the correct category.

Upon further inspection of the languages reported under "Others", we noted that they were predominantly either an Indian language (usually of South Indian origin) e.g. Kannada, Bengali, or Malayalam, or Bahasa Indonesia; a regional variant of the Malay language spoken in Singapore. The questionnaire was also designed to only elicit a maximum usage of two languages (one being English) by the domestic helper with the child (see Appendix 1). The second language (if any) used by the domestic helpers was also reported under "Others". However, according to the parents' report, most domestic helpers used Malay, Bahasa Indonesia, or an Indian language with the child. As these languages are largely considered to be related to official Mother Tongues in Singapore – originating from the same language family or sharing similar linguistic structures – we decided to create new categories and list these languages as separate from the label “Others”.

Therefore, for language use with the child, we recoded the information provided into the following categories: English, Mandarin, Dialect (Chinese), Malay, Malay (Other), Tamil, Indian (Other), and Others. Malay (Other) refers to ‘Bahasa Indonesia’ or ‘Indonesia Malay’ as reported by some parents. Indian (Other) refers to any Indian languages reported that are not Tamil. The new categories allowed for an accurate reflection of truly “Other” languages; that is, languages not natively spoken by Singapore. Any languages reported as ‘Other’ that does not fall into these two categories remain under that label e.g. Dutch.

The questionnaire further required the respondents to give an estimation of the percentage of a particular language used in proportion to the total number of languages used. However, not all respondents gave answers which totalled up to a 100%, so no matter the total value, the proportion was recomputed based on the individual’s total.

English	Mandarin	Dialect	Malay	Tamil	Total	**Proportion of English used**
55	40	1	10	-	106	55/106 = 0.52

Table 1. An example of the calculation of language use by proportion for one child.

Next, from the proportions, we identified the dominant language used by each caregiver. Dominance was calculated by dividing the proportion of each language with the sum of the proportion of all other languages used. Caregivers are considered dominant in one language (e.g. English) if the proportion of English use is larger than the sum of all other languages used.

Parent	English	Mandarin	Dialect	Malay	Tamil	Other(s)	Dominant language
1	0.9	0.1	-	-	-	-	English
2	0.25	-	-	0.75	-	-	Malay
3	0.5	-	-	-	0.5	-	"Balanced"
4	0.125	-	0.375	-	0.125	0.375	"Balanced"

Table 2. An example of language dominance of parents determined in this section. Parents 1 – 4 represent parents from different households.

In Table 2, we see an example of how language dominance of caregivers is determined. Some parents are classified to be "Balanced" speakers, i.e. they have do not have a dominant language which they use with their child. We note that the category of "Balanced" speakers may not be specific enough. As the example in Table 1 shows, a parent who uses English and Malay equally with their child would pattern differently from a parent who uses four different languages with Dialect and an 'Other' language taking up an equal but larger proportion of use though both parents are considered "Balanced" speakers. However, for the current level of description, further differentiation of the types of 'balanced' speakers will not be presented.

Family types. All households were classified into sixteen family types based on the information about which caregiver routinely spoke which language to the child (see Table 3). Parents were considered involved in contributing to the language environments of their child if the total number of languages spoken by both parents was more than or equals to one. Grandparents were considered involved if it was reported that they saw the child at least once a day or once a week. For domestic helpers and any other caregiver(s), if the questions regarding them were left unanswered, it was assumed that these households did not engage a domestic helper or any other caregivers to help care for the child, or that the household judged their linguistic input to the child to be minimal.

Family types	
1	Parents only
2	Parents, GP 1&2
3	Parents, GP 1&2, GP 3&4
4	Parents, GP 3&4
5	Parents, GP 1&2, Helper
6	Parents, GP 1&2, Helper, Someone else
7	Parents, GP 1&2, Someone else
8	Parents, GP 1&2, GP 3&4, Helper
9	Parents, GP 1&2, GP 3&4, Helper, Someone else
10	Parents, GP 1&2, GP 3&4, Someone else
11	Parents, GP 3&4, Helper
12	Parents, GP 3&4, Helper, Someone else
13	Parents, GP 3&4, Someone else
14	Parents, Helper
15	Parents, Helper, Someone else
16	Parents, Someone else

Table 3. The classification of family types according to language use.

Care time. Parents were asked to provide details on the amount of time in any given week each caregiver spends with the child in hours and as a percentage (See Appendix 1 for the questions). For our analysis, we considered first the data given in percentages. If the percentage of time spent was not provided, the information from the number of hours given was considered.

We classified the households into the same 16 family types but now based on the information on the proportion of **time spent** and calculated an average time spent by each caregiver in each family type. All caregivers were considered involved in caring for the child if the input for time spent was provided. For some households, the classification of family types based on **language use** and **time spent** may differ due to missing information.

However, this does not affect the final computation of the language input matrix for each child as the computation is performed with the raw information on time spent reported by the parents. 57 families did not provide any information for the time spent caring for the child. For these households; the missing values were replaced with the average time shown in Table 4, according to the family type they were listed under based on **language usage**.

Language Input Matrix. With information on language use and proportion of time spent by each caregiver, the proportion of all the languages spoken to the child is calculated. For instance, the total amount of English language input a child receives was calculated by the following equation:

$$\Sigma \text{ of all caregivers } (Caregiver's\ care\ time \times amount\ of\ English\ used\ by\ each\ caregiver)$$

The sum of input from **each language** used was calculated. Subsequently, a language input matrix was tabulated, with the proportion for all the languages spoken to the child by all the caregivers. Figure 3 displays an example of the child's language input matrix visualised with a pie chart. All plots and pie charts in this chapter were drawn using Tableau unless otherwise stated (Tableau Software Inc., 2017).

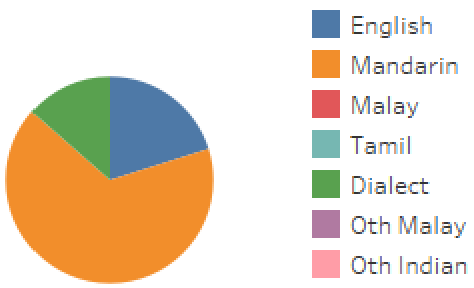


Figure 3. An example of a child's language input matrix (6-months).

For the language input matrix, the sum of the total proportion of languages for each individual should be one. There were 65 cases where missing data from language usage resulted in a total less than one. For these cases, instead of attempting to guess the proportions of language use, a recalculation of the proportion of time spent looking after the child was done after removing time spent by the individual whose language usage information is missing. Consequently, for these 65 households, the proportion of language use of the remaining caregivers increased. The missing information on language use is usually that of the grandparents and domestic helpers.

Results

Family types. While official national census collects information on the languages used by Singaporeans, publically available information rarely reports the use of more than two languages. There is also an absence of information about languages found in the home environments of Singaporean children given the linguistic diversity found in Singapore.

Table 4 shows the distribution of households in our study according to the family type they were classified under. From Table 4, we see that the care of the 6-month olds in our study was taken on mostly by their parents only, or with help from one set of grandparents.

	Family types	Number of families (6-months)
1	Parents only	109
2	Parents, GP 1&2	59
3	Parents, GP 1&2, GP 3&4	48
4	Parents, GP 3&4	50
5	Parents, GP 1&2, Maid	17
6	Parents, GP 1&2, Maid, Someone else	3
7	Parents, GP 1&2, Someone else	15
8	Parents, GP 1&2, GP 3&4, Maid	16
9	Parents, GP 1&2, GP 3&4, Maid, Someone else	5
10	Parents, GP 1&2, GP 3&4, Someone else	20
11	Parents, GP 3&4, Maid	13
12	Parents, GP 3&4, Maid, Someone else	1
13	Parents, GP 3&4, Someone else	7
14	Parents, Maid	49
15	Parents, Maid, Someone else	2
16	Parents, Someone else	30

Table 4. The classification of family types and number of households per type (6-months).

Table 5 shows the mean time spent looking after the child by each caregiver according to the sixteen family types.

Family Type	P1	P2	GP 1	GP 2	GP 3	GP 4	Helper	Someone
1	0.80	0.20						
2	0.40	0.18	0.36	0.07				
3	0.33	0.19	0.24	0.06	0.15	0.03		
4	0.48	0.15			0.32	0.05		
5	0.30	0.10	0.23	0.04			0.32	
6	0.19	0.14	0.08	0.05			0.11	0.44
7	0.58	0.14	0.07	0.03				0.17
8	0.37	0.14	0.11	0.05	0.04	0.01	0.27	
9	0.21	0.18	0.08	0.08	0.08	0.08	0.08	0.21
10	0.33	0.12	0.03	0.03	0.09	0.07		0.34
11	0.34	0.17			0.15	0.04	0.30	
12	0.30	0.06			0.28	0.03	0.30	0.03
13	0.46	0.22			0.06	0.02		0.24
14	0.41	0.11					0.42	
15	0.31	0.18					0.33	0.18
16	0.54	0.15						0.31

Table 4. The proportion of time spent with the child by each caregiver in each family type (6-months). Mothers tend to provide 2-4 times more care time compared to fathers (emphasised in red).

When reviewing the average time spent caring for the child shown in Table 5, an interesting phenomenon is observed. The average amount of time mothers spent looking after their 6-month-olds was about 2-4 times more than their spouses. This was the case even if grandparents or domestic helpers were involved. The only family type with a similar proportion of care time by both parents is when both maternal and paternal grandparents are involved as well as a domestic helper and an external caregiver (see Family type 9 in Table 5).

What languages are Parents using with their 6-month-olds?

Number of languages used with the child (6-months)	Parent 1	Parent 2	Both Parents
1	48	100	36
2	333	298	326
3	58	35	73
4	5	4	9
5	0	0	0
6	0	0	0
Total	444	436	444

Figure 4. The number of languages used with the child by the parents at 6-months. When calculating the number of languages that both parents use, overlaps were not counted. Visual representation of the spread is presented with the bar chart.

The total number of languages used by the parents is presented in Figure 4. When calculating the total number of languages used by both parents, common languages spoken by both parents were only counted once. For example, if a child's mother used English and Mandarin, and his father used English and Malay, the total tally would be three. It is important to note that while most parents in our sample spoke two languages to their child, 18.5% of parents reported using three or more languages with their 6-month-old child.

Language Dominance (6-months)	Parent 1	Parent 2
English	140	133
Mandarin	86	107
Dialect	2	4
Malay	65	75
Other Bahasa	1	1
Tamil	28	32
Other Indian	15	15
Other(s)	1	0
"Balanced"	106	69
Total	444	436

Table 6. Distribution of parents based on the dominant language used with the child (6-months).

According to the 2015 General Household Survey, 73.2% of Singaporeans aged 15 and above are literate in two or more languages and 36.9% who are aged 5 and older speak English most often at home (Department of Statistics, 2015). Table 6 shows the number of parents with corresponding language dominance. It is unsurprising, therefore, to find that when the child is 6 months of age, the majority of parents (71%) in the GUSTO cohort reported using two languages with their children. The most common language usage pattern among parents is the use of two languages with English as the dominant language.

Grandparents' language use with the child

How often do your parents/parents-in-law see your child? (6-months)	Grandparents (Parent 1)	Grandparents (Parent 2)
Once a day	154	144
Once a week	97	114
Once a month	54	57
Less often	111	101

Table 7. The frequency of a child's interaction with his/her grandparents at 6-months.

Table 7 shows the number of grandparents grouped according to the frequency of interaction of interaction with their grandchild. About half of the children (57.3%) see their paternal and/or maternal grandparents at least once a week. Therefore, it is important to identify the patterns of language use of grandparents when they are interacting with the child.

Number of languages used with grandchild (6-month)	GP 1	GP 2	GP 3	GP 4	Total
1	196	203	217	200	95
2	179	139	153	113	190
3	23	9	12	10	138
4	1	1	0	0	4
5	0	0	0	0	1
6	0	0	0	0	0
Total	399	352	382	323	428

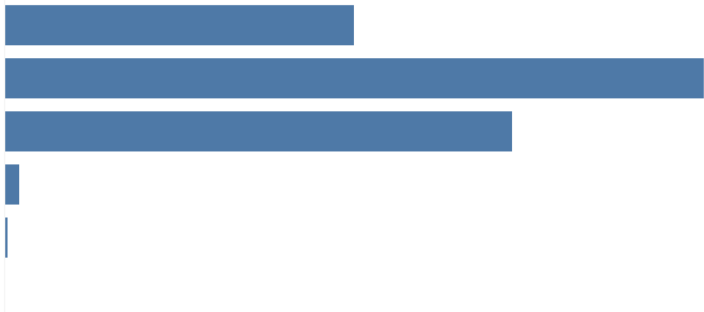


Figure 5. The number of languages used by grandparents with their grandchild at 6-months. When calculating the number of languages that both grandparents use, common languages were only counted once. Visual representation of the spread is presented with the bar chart.

Figure 5 presents the number of languages used by the grandparents with their 6-month old grandchild. Compared to parents' use of language which is overwhelmingly bilingual (71%); 44% of grandparents use two languages while 32.2% are using three languages with their grandchildren.

Language dominance	GP 1	GP 2	GP 3	GP 4
English	37	45	35	41
Mandarin Chinese	136	122	139	110
Dialect	33	25	20	15
Malay	102	95	102	88
Other Bahasa	4	3	1	0
Tamil	30	24	30	25
Other Indian	15	13	14	13
Others	1	2	3	4
"Balanced"	41	22	38	28

Table 8. Grandparents' dominant language when speaking with their grandchild.

Ethnicity	Language	% spoken
Chinese	Mandarin	38.5
	Dialect	37.2
Malay	Malay	93.2
Indian	Tamil	43.6

Table 9. Percentage of languages spoken most frequently at home for Singaporeans aged 55 and over according to 2015 General Household Survey (Department of Statistics, 2015).

Table 8 shows the dominant language of the grandparents when talking to their 6-month old grandchild. Majority of Singaporeans, aged 55 and over, use ethnic languages as their dominant language at home as presented in Table 9 (Department of Statistics, 2015). Consistent with data from the population census, the majority of the grandparents in our study use an ethnic language as their dominant language with their grandchild. We note that the ethnic languages, including the Other Bahasa language and other Indian languages, may be the native tongues of the grandparents.

Domestic helper's language usage with the child at 6 months

Out of the 130 families who responded to the question on language use of domestic helpers, 96 respondents (73.8%) of reported that English was the dominant language used by the helper. The number of languages used by the domestic helpers with the child under their care shown in Figure 6, reportedly a majority of helpers use only one language. The distribution of helpers according to language dominance is shown in Table 9.

Number of languages used with child	Helper
1	100
2	30
3	0
4	0
5	0
6	0

Figure 6. The number of domestic helpers and the corresponding number of languages they use. Visual representation of the spread is presented with the bar chart.

Language dominance	Helper
English	96
Chinese Mandarin	1
Dialect	0
Malay	74
Other Bahasa	9
Tamil	3
Other Indian	3
Others	2
"Balanced"	2

Table 10. The dominant language used by the domestic helpers with the child.

The questionnaire was designed only to include the use of two languages by the domestic helpers (See Appendix 1) and parents reported that majority of the helpers used only one language with the 6-month olds in our study. Most foreign domestic workers in Singapore come from Indonesia (Phua, Hui, Nodzenski, & Bacolod, 2012). It is therefore unsurprising that Malay (and Bahasa Indonesia) is the most common dominant language used by domestic helpers after English.

Other caregivers' language use with the child at 6-months

Beyond the above-mentioned caregivers, parents were asked for language use of anyone else who spends time looking after their child, for example, teachers at child-care centres, and nannies. Assuming that there may be more than one individual involved, the information provided for the language use was taken as an estimation of all other parties involved e.g. all the teachers at the childcare centre and not for individuals. 233 households (52.5%) reported engaging someone else other than grandparents and domestic helpers to help care for their 6-month-olds.

Number of languages used with child	Someone else
1	93
2	128
3	12
4	0
5	0
6	0

Figure 7. The number of other caregivers(s) and the corresponding number of languages they use with the child. Visual representation of the spread is presented with the bar chart.

Language dominance	Someone else
English	64
Chinese Mandarin	47
Dialect	2
Malay	42
Other Bahasa	2
Tamil	18
Other Indian	3
Others	2
“Balanced”	53

Table 11. The dominant language used with the child by other caregivers.

Figure 7 shows the number of languages external caregiver(s) uses with the children in our study and Table 11 shows the distribution of these external caregivers according to language dominance. Majority of these caregivers use two languages when speaking to the child and English is the most common dominant language.

Language Input matrix at 6-months

A child’s language input matrix represents all the languages he/she was hearing in his/her environment at 6-months; languages used intentionally by his/her caregivers to talk to him/her.

Number of languages in child’s environment at 6 months	
1	31
2	284
3	117
4	12
5	0
6	0

Table 12. The number of languages used by caregivers when the child is at 6 months

Table 12 shows the distribution of children according to the number of languages used by their caregivers at 6 months. In accordance with the prevalence of bilingualism amongst Singaporean adults, the majority of the children are receiving input of two languages from their adult caregivers. Table 13 shows the distribution of children according to the dominant languages present their environments at 6-months.

Language dominance	
Language Input matrix (6-month)	
English	140
Chinese Mandarin	117
Dialect	5
Malay	90
Other Bahasa	1
Tamil	34
Other Indian	15
Others	1
"Balanced"	41

Table 13. The distribution of children according to dominant languages in child's environment at 6 months.

Most 6-month olds in the GUSTO cohort (60.1%) are hearing at least two languages in their home environments, but the bilingual input is not balanced. English is the most common dominant language that caregivers use (33%) when speaking with their 6-month-olds. The language input matrixes of all 444 children are visualised as pie charts and presented in Figure 8. The children are arranged in a random fashion and variation and individual differences of language input are apparent.

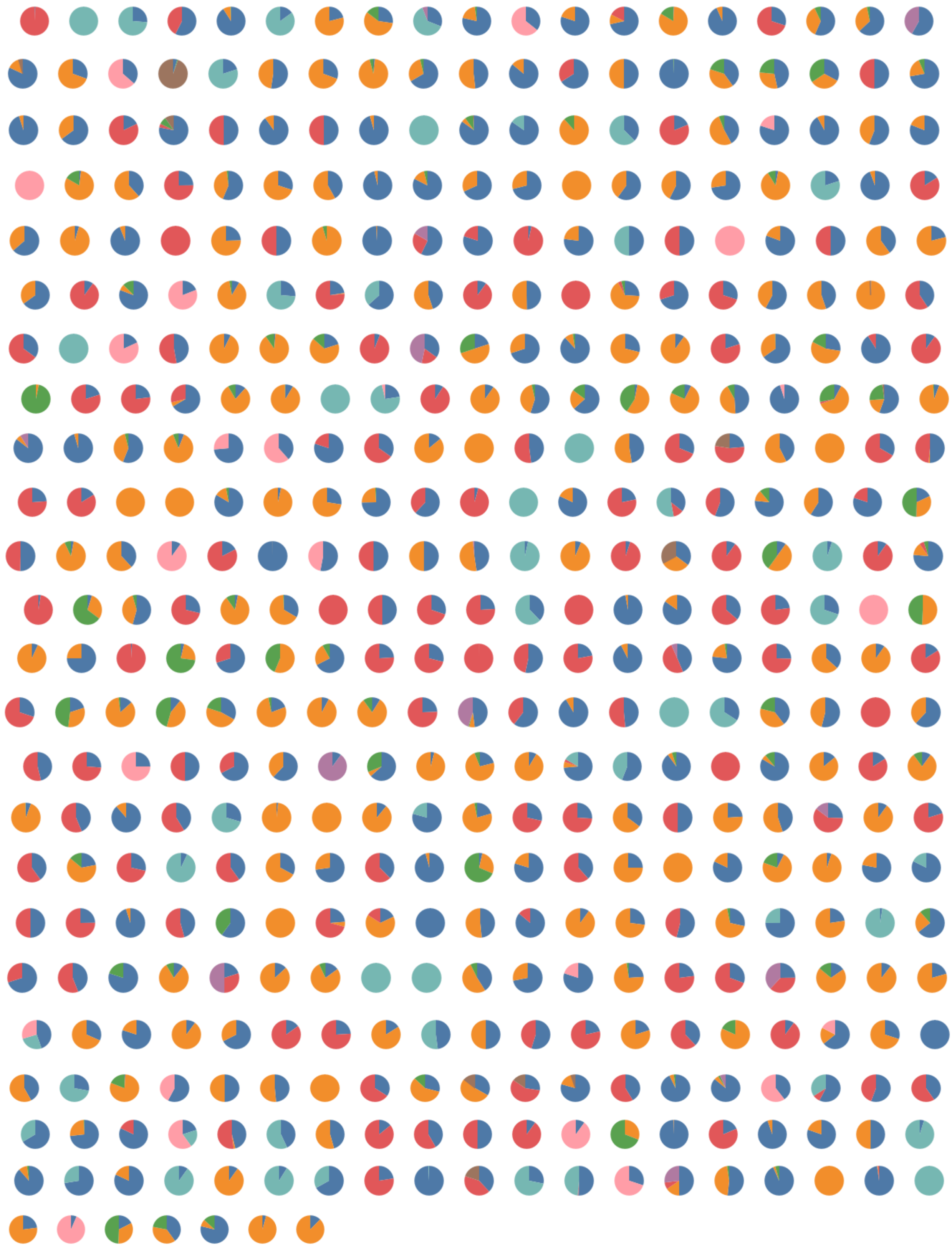
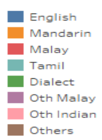


Figure 8. The language input matrixes of all children according to parent-reported data at 6-months.



Descriptive overview of Language Use of Caregivers with 6-month olds

With regard to language use, the Singapore population census published once a decade and the General Household Survey conducted once every five years only reports on Singaporeans' use of English and ethnically indexed languages. We have very limited data on the use of three languages or more and are uninformed about languages used with young children growing up in Singapore.

Here, we have, for the first time, presented a detailed descriptive analysis of language use by caregivers with 6-month olds growing up in Singapore. The data used was information reported by parents (mostly mothers) via the GUSTO Language Background Questionnaire from 2011 to 2012 and this data has not analysed by the GUSTO team. While the design of the survey may not be optimal to elicit comprehensive information about language use, our analysis provided fresh insights into the language backgrounds of pre-lingual Singaporean children and extracted several key descriptors:

i. Parents: Bilingual input is the most common (71%); English is the most common dominant language.
ii. Grandparents: Bilingual input may be most common (44%) but a large proportion (33%) use three languages with their grandchildren. Ethnic languages are used predominantly.
iii. Overall, children are mostly receiving bilingual input (English and another language) but the inputs are not balanced.

Beyond this descriptive analysis and the noting individual differences as seen from the pie charts in Figure 8, we also aimed to identify common patterns and groupings of children according to their language exposure at 6-months and 18-months. For this purpose, the calculated language input matrixes were be used for further computation and statistical analyses in the following sections.

Analysis A: Clustering Analysis for Language Input matrix at 6-months

From the previous section, we gained understanding about the languages used in the home environments of pre-lingual 6-month olds in Singapore. In this section, we examine observable patterns of language exposure among these children through cluster analysis. We also investigate if later language ability is predicted by differing exposure patterns of English at 6 and 18 months of age.

As mentioned, the research on language exposure of pre-lingual children in Singapore is scarce and there is hardly any fixed criterion for classification of monolinguals, bilinguals or multi-linguals that are based on normalised data of Singaporean children. Instead of merely classifying them according to prescribed ethnic/language groups, we would like our data to inform us on the best way to characterise patterns of multilingualism in Singapore according to the nuanced details of the language exposure for children in the GUSTO sample and what proportions of exposure to various languages would lead a child to be classified as a "dominant" in a particular language or as a "balanced" bilingual. We investigated the possibility of various types of bilingual exposure and explore patterns of language input among them. The data-driven classifications would allow us to characterise Singaporean children for their language background. Furthermore, as most standardised measures of language abilities test for English language proficiency, the classifications would allow for future comparisons of between-group differences.

Methods

While configurations of language input by respective caregivers will differ from one household to another, there are expected similarities and we aim to find an optimal number of groups in which similarities among children in a group are high while similarities among children of different groups are low. Thus, as a means of exploration, *K*-means clustering analysis on the computed language input matrixes was conducted using SPSS (IBM Corp, 2016). *K*-means clustering process begins with an initial grouping of *K* clusters and new partitions are generated by assigning each data point to its closest cluster centre. The metric for measuring distance is the Euclidean metric. The process is repeated until the cluster membership stabilises (Jain, 2010).

In our study, the *K*-means clustering method forms a designated number of clusters (decided by the researcher) using an iterative process, in which every child in our study is assigned to a cluster based on their position in a multi-dimension statistical space. Each dimension corresponds to a variable in the language input matrix. There is a maximum of eight dimensions corresponding to the eight different language variables: English; Mandarin; Dialect; Malay; Other Bahasa; Tamil; Other Indian; Others. When the clustering is complete, each cluster centre reflects the mean language input of each language variable involved in the said cluster based on the language input of all the children assigned to that cluster. That is to say, each child occupies a position in multidimensional space (according to their percentage of input in each of their languages), and the clustering process tells us which ‘group’ the child belongs to, on the basis of how close the other children are in the multidimensional space.

The cluster centres are visualised with the language input matrix pie charts of the centroid exemplars of each cluster. Centroid exemplars of each cluster are individual children whose language input is the closest in Euclidean distance to the means reflected by the cluster centres and thus they are highlighted in this section as the best representations of language input variation of that cluster.

The clustering process was iterated with a gradual increase of clusters by one for each round. To determine an optimal number of clusters for our data, we made two calculations using the elbow method and average silhouette method using an R algorithm (Kassambara, 2017). The elbow method calculates the total within-cluster sum of squares while the average silhouette method measures how well each object lies within the cluster. The results are shown in plots in Figure 9.

Characteristic of the clustering process in the SPSS programme, the first *K* clusters are chosen as the initial clusters when *K* clusters are requested (IBM Corp, 2016). The selection of a case as a new centre is governed by rules based on the distance from the case to existing cluster centres, relative to the distances among the current centres. Hence, the order of cases may affect whether cases are selected when evaluated in the first pass and whether they are replaced by a later case after an iteration. In order to ensure that the model we present is the most representative solution and to assess the stability of the clusters, for each *K*, we randomly sorted our data twenty times according to twenty sets of random numbers generated and ran the *K*-means analysis for each set. The results we report for each K is the analysis that had the highest reoccurrence out of the twenty runs. For the ease of comparison, the clusters for all *K* values are given consistent letter names in each cluster analysis, using the following code: A – English-dominant, B – Mandarin-dominant, C – Malay-dominant, D – Tamil-dominant, E – Dialect & Mandarin, F – Other Indian-dominant, G – Others, H – English & Mandarin, I – English & Malay, J – English & Tamil, K – English & Other Indian, L – English & Other Malay, M – English & Dialect, and N – Dialect-dominant. All the figures reported about the cluster make-up are arranged according to the size of the cluster.

Results

The optimal number of clusters for language backgrounds of the 444 6-month olds in the GUSTO cohort as determined by the elbow method and average silhouette method are 4 and 9 respectively. The elbow method selects the optimal number of clusters where the total within-cluster sum of square is minimised and increasing the number of clusters will not further reduce the variation within the cluster while the average silhouette method identifies the optimal number of clusters by calculating how close a point in a cluster is to other points in its neighbouring clusters. Figure 9 shows the plots generated in R following the algorithm from Kassambara (2017).

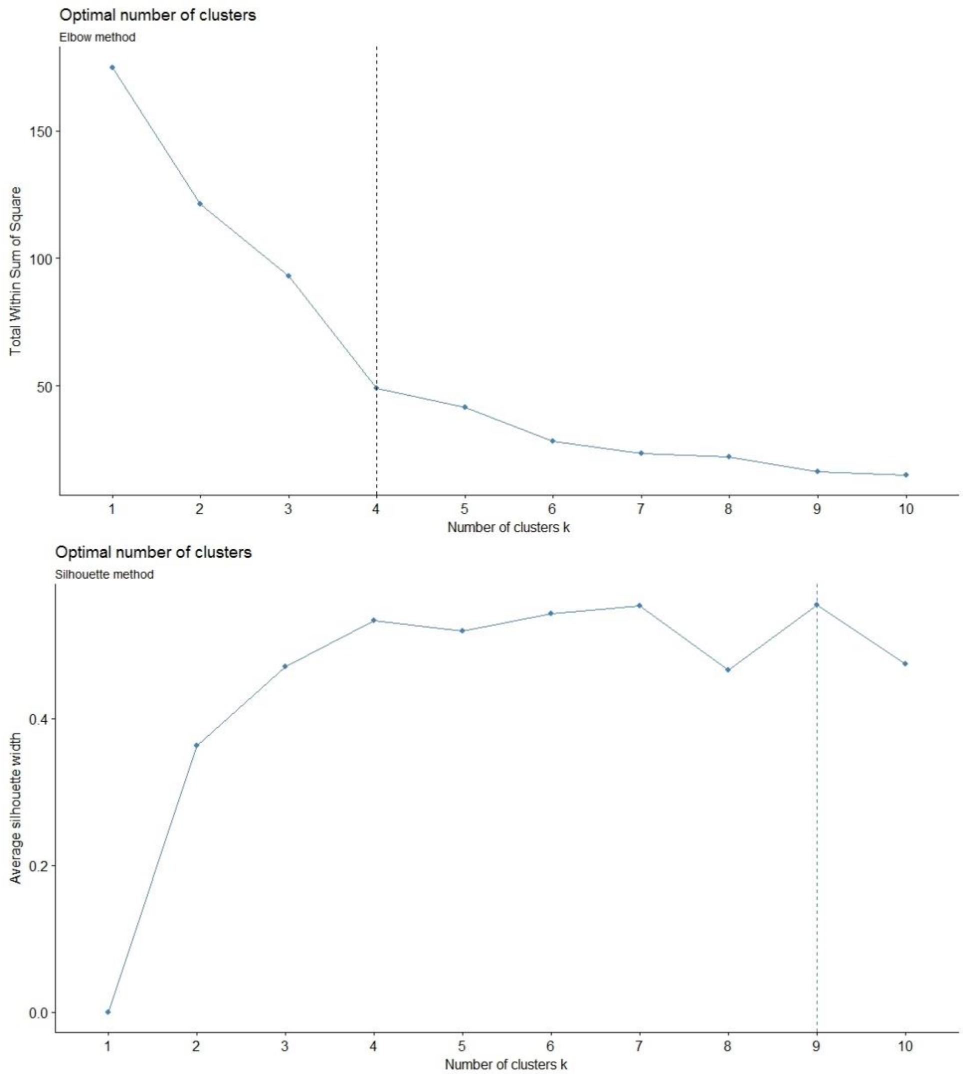


Figure 9. Plots for an optimal number of clusters calculated with elbow and average silhouette method. The algorithm for the calculation and plots were from Kassambara (2017)

4 clusters

K-means clustering was performed on the language input matrix of 444 6-month olds in the GUSTO cohort. We started with setting the number of clusters to 4 to reflect the main language groups from the language input matrix.

When the number of clusters is set to four, the final cluster centres are set as follows:

A: English-Dominant, B: Mandarin-Dominant, C: Malay-Dominant, D: Tamil-Dominant. This solution for a 4-cluster model is the most stable. It occurred in 11 out of the 20 different randomly seeded runs. Each child remained in the same cluster 71% of the time within the 20 runs of the cluster analysis.

Cluster Centres

(4-cluster Lang Input matrix 6-months)

	English (A)	Mandarin (B)	Malay (C)	Tamil (D)
English	.69	.19	.29	.18
Mandarin	.12	.72	-	-
Dialect	.02	.09	-	-
Malay	.03	-	.70	-
Other Bahasa	.02	-	.01	-
Tamil	.02	-	-	.81
Other Indian	.09	-	-	-
Others	.01	-	.01	-
N	156	138	114	36
Range of Distances	.122 – 1.148	.040 – 1.127	.017 – .638	.022 – .462

Figure 10. Means of language input for each cluster when the number of clusters is set at 4 (Lang Input matrix at 6-months) and centroid exemplars for each cluster; the means for the dominant languages are highlighted. N represents the number of children in each cluster; The Range

represents the Euclidean distance of the individuals in the cluster from the cluster centre. At the base of the table, a pie chart represents the centroid exemplar (the most typical child) from each cluster.

Figure 10 shows the means of language input of each language within an individual cluster when the number of clusters was set at 4, as well as the number of members for each cluster, and visual representations of language input matrixes of all the centroid exemplars.

At first glance, this clustering may align with commonly reported statistics on language use in Singapore, with language dominance as a key differentiator. It is, however, uninformative. The reported means of the variables which make up the largest cluster, Cluster A (English-dominant), consisted of contributions from all the other languages: English (.69), Mandarin (.12), Dialect (.02), Malay (.03), Other Bahasa (.02), Tamil (.02), Other Indian (.09) and Others (.01). It is unlikely that any individual child would have such a varied language input from their caregivers. Data collected from the questionnaire also does not include children whose language input exceeded four languages. The range of Euclidian distances of cluster membership (how close an individual is to the cluster centre) is extremely large, ranging from .122 – 1.148.

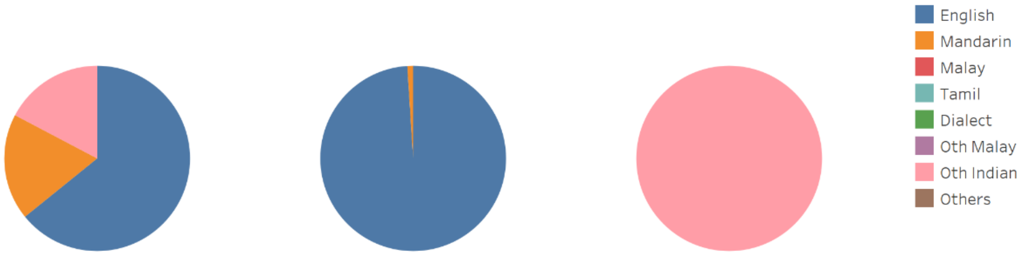


Figure 11. Variation of language input within individuals from the English-Dominant cluster (left to right): individual closest to the cluster centre; individual at the mean distance from the cluster centre; individual furthest away from the cluster centre.

Figure 11 shows the language input matrix of the centroid exemplar of the largest cluster and the language input matrix of the child who is at the mean Euclidian distance from the cluster centre and the language input matrix of the child at the furthest Euclidian distance from the cluster centre. The child who is the centroid exemplar of Cluster A (English-dominant) has language input of mostly English (.64) and some Mandarin (.19) and an Other Indian language (.17) while the child at the furthest away from the cluster centre is only exposed to an Other Indian language (1.00). Although the allocation of children into four groups is mathematically parsimonious, it is clear that the children in this particular cluster do not belong to a coherent language usage group and an increase in the number of clusters is necessary to better classify the children into meaningful clusters. An increase of the number of clusters to 9 as suggested by the average silhouette method accounted for variances from the other languages besides English and the three official Mother Tongues.

9 clusters

With the number of clusters set at nine as suggested by the average silhouette method, clearer distinctions of the various groupings of children are observed. The children were further separated corresponding to dominance in various languages.

The final cluster centres are as follows: A: English-dominant, B: Mandarin-dominant, C: Malay-dominant, D: Tamil-dominant, E: Dialect/Mandarin-dominant, F: Other Indian-dominant, G: Others-dominant, H: English-Mandarin balanced and L: English-Other Malay balanced. This solution is the most stable; occurring 14 out of the 20 times we ran the analysis on different randomly-ordered sets of the same data. A child also remained in the same cluster at a rate of 98.4% across all twenty clustering analyses.

Figure 12 provides cluster centre means when the number of clusters was set at 9, the size of the clusters is indicated along with the visualisations of the language input matrix of the centroid exemplars.

	Malay (C)	English (A)	Mandarin (B)	English & Mandarin (H)	Tamil (D)	Other Indian (F)	Dialect & Mandarin (E)	English & Other Malay (L)	Others (G)
English	.29	.80	.12	.48	.18	.21	.17	.35	.05
Mandarin	-	.09	.85	.48	-	-	.37	.03	-
Dialect	-	.02	.04	.03	-	-	.54	-	.01
Malay	.70	.04	-	-	-	-	-	.13	-
Other Bahasa	.01	-	-	-	-	-	-	.48	-
Tamil	-	.03	-	-	.81	.01	-	-	-
Other Indian	-	.01	-	-	-	.77	-	-	-
Others	.01	-	-	.01	-	0	-	-	.94
N	112	109	88	60	36	16	15	7	1
Range of distances	.008 - .411	.075 - .515	.025 - .339	.041 - .381	.022 - .462	.045 - .436	.096 - .554	.061- .505	.000

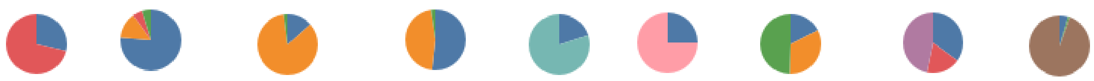


Figure 12. Means of language input for each cluster when the number of clusters is set at 9 (Lang Input matrix at 6-months) and centroid exemplars for each cluster; the means for the dominant languages are highlighted. N represents the number of children in each cluster; The Range represents the Euclidean distance of the individuals in the cluster from the cluster centre. At the base of the table, a pie chart represents the centroid exemplar (the most typical child) from each cluster. The clusters have been arranged according to cluster size.

When K is set at nine, the largest cluster changes from the English-dominant cluster to the Malay-dominant cluster (Cluster C). Cluster A (English-dominant) is the second largest cluster with means of the contributing variables remained largely the same as when there were 4 clusters: English (.80), Mandarin (.09), Malay (.04), Tamil (.03), Dialect (.02), Other Indian (.01). While the mean amount of English input for the English-Dominant cluster has increased from .69 to .80 when K was increased from 4 to 9 – effectively reassigning some children to another cluster – this has not lead to a further and clearer distinction of children in this cluster. The range of distances from cluster centre constituting cluster membership remains extremely large, from .075 – .515. Figure 13 shows the language input matrixes of the centroid exemplar of the Cluster A (English-dominant) and the individual with the furthest Euclidian distance from the centre.

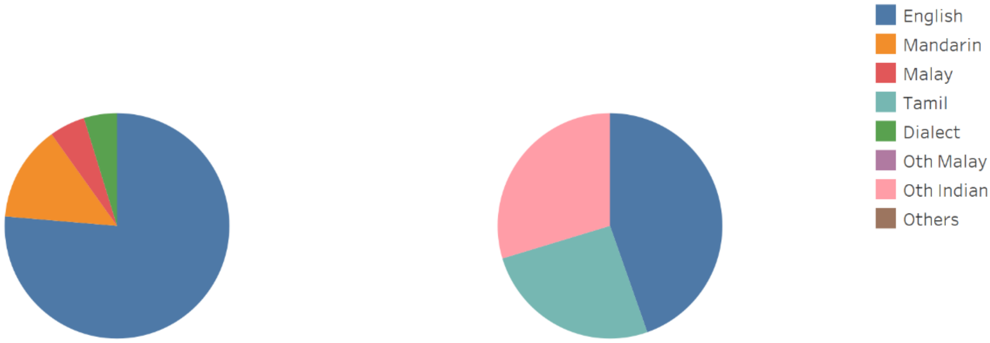


Figure 13. Variation of language input patterns within a single cluster (6-months): Centroid exemplar of Cluster A: English-dominant (left); individual furthest away from the cluster centre (right).

Figure 13 clearly illustrates that children with vastly different language input patterns are still grouped together in the same cluster when the number of clusters is set at 9.

Comparing the number of children in each cluster when the number of clusters is set at 4 and 9, we observed that the children within the Malay-dominant (Cluster C in 4-cluster) and D Tamil-dominant (Cluster D in 4-cluster) clusters have largely retained the same members. The increase in the number of clusters has resulted in finer distinction for Mandarin hearing children; separating those who hear predominantly Mandarin and those whose speech environments have most input from both Mandarin and Chinese Dialects.

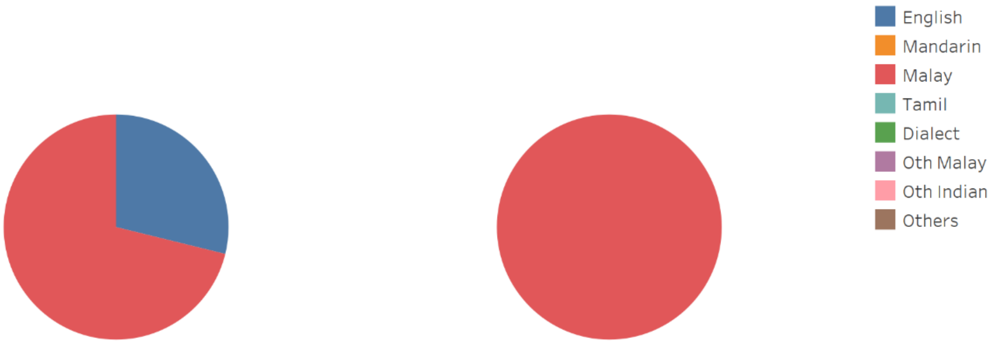


Figure 14. Variation of language input patterns within a single cluster: Centroid exemplar of Cluster C (Malay-dominant (left); individual furthest away from the cluster centre (right).

Figure 14 shows the visualization of language input matrixes of the centroid exemplar of the Cluster C (Malay-dominant) and the individual with the furthest Euclidian distance from the centre.

With the number of clusters set at 9, cluster centres are formed based on dominant languages and may not accurately represent children who received a balanced amount of input of their languages. For the purpose of this exploratory method of clustering, an increase in the number of clusters to 12 clusters was done to examine if it could provide a better representation of the variance among children with balanced language input.

12 clusters

When the designated number of clusters is set at 12, the three main state official Mother Tongues in Singapore: Mandarin, Malay and Tamil, appear to have two clusters each with one cluster representing a balanced English-Mother Tongue input while the other showing a Mother Tongue-Dominant pattern.

Figure 15 provides cluster centre means when the number of clusters was set at 12, the sizes of the clusters are indicated along with the visualisations of the language input matrix of the centroid exemplars. The distance from the cluster centre of all the members in each cluster is shown in the box-plot in the following Figure 16.

Cluster Centres

(12-cluster Lang Input matrix 6-months)

	English (A)	Mandarin (B)	Malay (C)	English & Mandarin (H)	English & Malay (I)	English & Tamil (J)	Tamil (D)	Other Indian (F)	Mandarin & Dialect (E)	English & Other Malay (L)	Dialect (N)	Others (G)
English	.82	.11	.17	.48	.46	.42	.05	.21	.13	.35	.02	.05
Mandarin	.10	.85	-	.48	-	-	-	-	.45	.03	.23	-
Dialect	.03	.03	-	.03	-	-	-	-	.42	-	.74	.01
Malay	.03	-	.83	-	.52	.01	-	-	-	.13	-	-
Other Bahasa	-	-	-	-	.01	-	-	-	-	.48	-	-
Tamil	.01	-	-	-	-	.55	.94	-	-	-	-	-
Other Indian	.01	-	-	-	-	.01	-	.77	-	-	-	-
Others	-	-	-	.01	.01	-	-	-	-	-	-	.94
N	99	86	64	60	53	21	21	16	11	7	5	1
Range of distances	.050 - .493	.025 - .260	.008 - .239	.045 - .381	.019 - .325	.028 - .407	.011 - .267	.045 - .439	.045 - .248	.061 - .505	.020 - .298	.000

Figure 15. Means of language input for each cluster (arranged according to cluster size) when the number of clusters is set at 12 and centroid exemplars for each cluster (Lang Input matrix at 6-months); the means for the dominant languages are highlighted. N represents the number of children in each cluster; The Range represents the Euclidean distance of the individuals in the cluster from the cluster centre. At the base of the table, a pie chart represents the centroid exemplar (the most typical child) from each cluster.

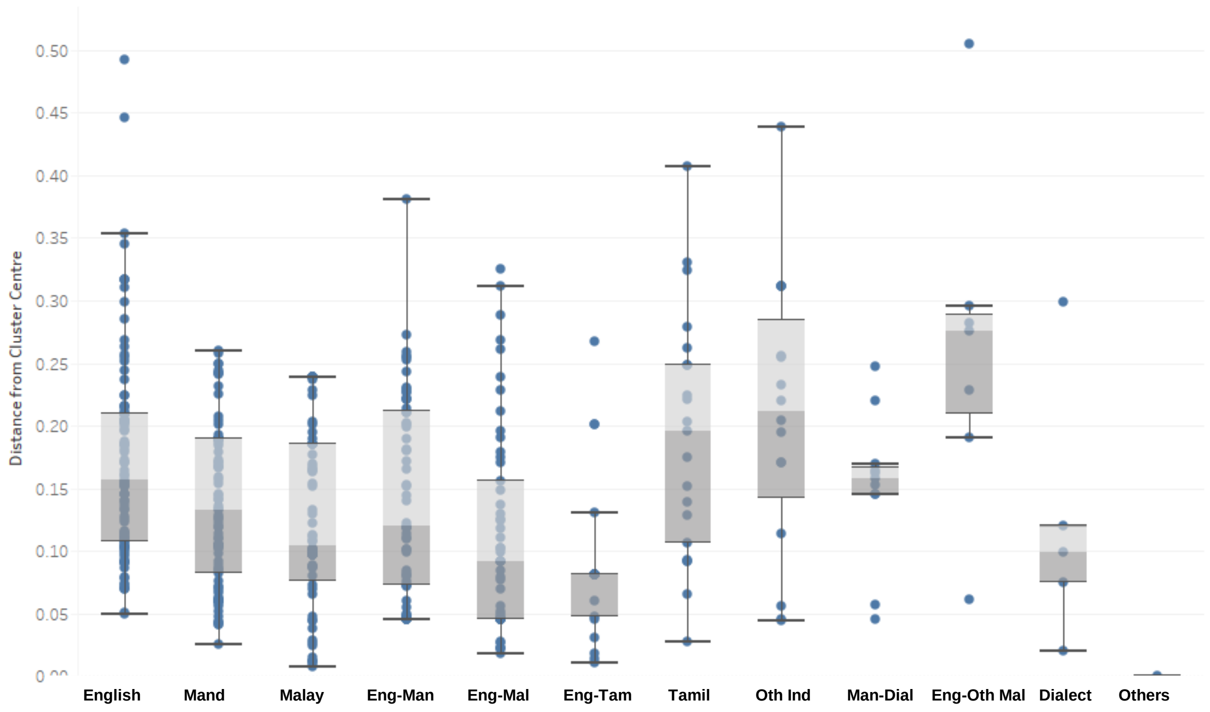


Figure 16. The distances of cluster membership from the cluster centre when there are 12 clusters, starting with the largest cluster (right to left). The boxes represent middle two quartiles of the distribution (25^{th} – 75^{th}). The whiskers extend 1.5 times from the interquartile range.

After 5 iterations, the cluster centres no longer changed and the cluster membership remained stable. This solution is the most stable; occurring 5 out of the 20 times the *K*-means analysis was carried out on 20 sets of randomly-ordered data. Each child remained in the same cluster across all 20 runs of the *K*-means analysis at a rate of 98.7%. With 12 clusters, the changes across the clusters were driven by language labels of the smaller clusters. For example, having a Chinese dialect-dominant cluster (N) in addition to the Mandarin-Dialect dominant cluster (E).

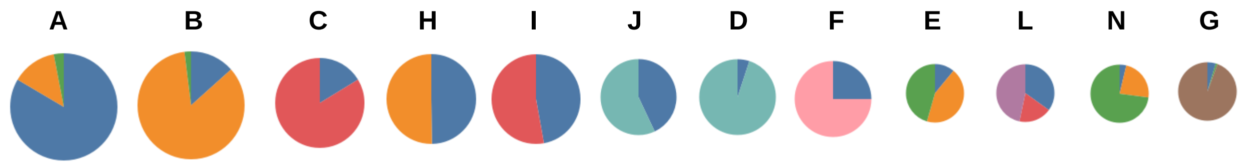


Figure 17. Language input variation of the centroid clusters when there are 12 clusters (6-months), scaled and arranged according to cluster size.

As highlighted by the distribution of language variation of the centroid clusters in Figure 17, there are similarities for the proportion of language input among all the language-dominant clusters and among the balanced language clusters. For the English, Mandarin and Malay dominant clusters (A, B, C), the proportion of the dominant languages is .82 – .85. The balanced bilingual groups (H, I, J) also have similar proportions.

From the distribution of language input of the centroid exemplars in Figure 17, it is clear that the differentiation among the clusters is largely driven by the labels of different languages and not by the amount of input per language. Hence, we decided to re-compute the Language Input matrix to group the Mother Tongues (Chinese, Malay, Tamil) under one label and 'variants' of these Mother Tongues (Dialect, Other Malay, Other Indian languages) under another label. This new Mother Tongue matrix quantified the Mother Tongue languages as contributing equal variances and removed the differences driven by individual labels. Separate *K*-means cluster analyses were conducted using the Mother Tongue matrix presented in the following section.

Analysis B: Mother Tongue Matrix clustering (6-months)

Methods

As seen from the earlier 12-cluster analysis, the differences in the patterns were driven by the labelling of individual language types. The data from the language input matrix was thus restructured to create a Mother Tongue matrix.

The children were first grouped according to their assigned Mother Tongues: Mandarin, Malay, and Tamil. This was determined by looking at parents' input of what languages other than English were used with their child. If Mandarin and/or Chinese Dialects were used, the child is classified as having Mandarin as his/her mother tongue. If Malay and/or Other Bahasa (i.e. Bahasa Indonesia) were used, the child is classified as having Malay as their Mother Tongue. Similarly, if Tamil or Other Indian languages were used, the child is classified as having Tamil as their Mother Tongue.

Secondly, four new variables were created and tabulated: Mother Tongue, Variation of (one's) Mother Tongue, Other Mother Tongues, and Variation of Other Mother Tongues. For example, for a child whose mother tongue is Mandarin, all Chinese dialects are considered as regional variants of one's mother tongue. Malay and Tamil were considered as Other MT and Other Bahasa and Other Indian languages were considered as regional variants of Other MT. Parents of one child used a Filipino language predominantly (presumably Tagalog) and some Hokkien thus the Filipino language was considered as the child's mother tongue and Hokkien was listed under Variations of Other MT. We acknowledge that from a narrow linguistic standpoint, Chinese dialects and Indian languages like Hindi may be considered as separate languages and not regional variations of Mandarin and Tamil. However, for the purposes of our analysis, we would list them as such.

Figure 18 provides an illustration of how the new variables were calculated for a child whose Mother Tongue is Mandarin. This approach allows us to capture shared variance between children with different mother tongues, that is, to group together Mother tongue dominant children from all mother tongues in a single cluster, separate from children who are English dominant (regardless of their mother tongues, and so on).

Example 1: Child classified as with Mandarin as MT

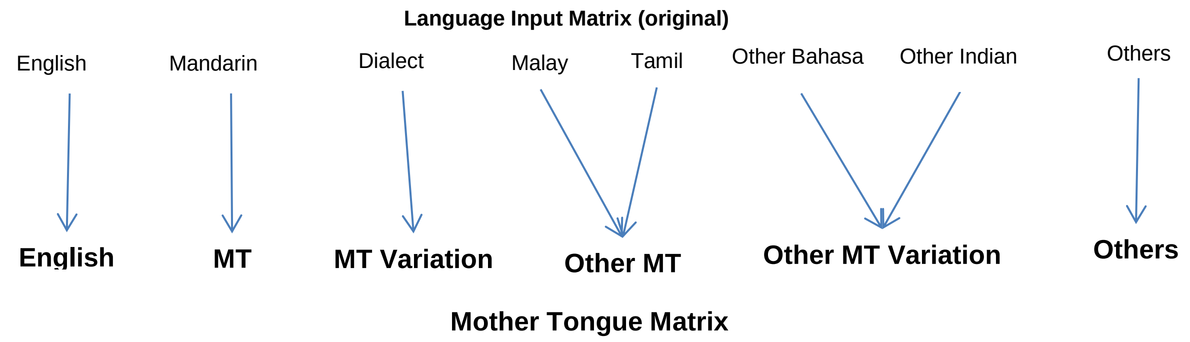


Figure 18. Reclassification of the components which make up the variables for Mother Tongue matrix.This is an example of a child with Mandarin as his mother tongue. The “Other MT” would be different for a Malay-hearing or Tamil-hearing child.

The newly computed Mother Tongue matrixes were subjected the same iterated procedure with *K*-means clustering as the Language Input matrix performed using SPSS (IBM Corp, 2016). We also calculated the optimal number of clusters with the elbow and the average silhouette method.

We began with three clusters to account for variance in the main significant language types in the Mother Tongue matrix – English, Mother Tongue and MT Variation respectively. The clustering process was iterated with a gradual increase of clusters by one for each round. For each *K*, we repeated the procedure of randomly sorting our data according to twenty sets of randomly generated numbers and running the *K*-means clustering analysis for each set to check the stability of our solution. In this section, we report the results for 3 and 4 clusters.

Results

The optimal number of clusters for MT matrixes of the 444 6-month olds in the GUSTO cohort as determined by the elbow method and average silhouette method are 4 and 3 respectively. Figure 19 shows the plots generated in R following the algorithm from Kassambara (2017).

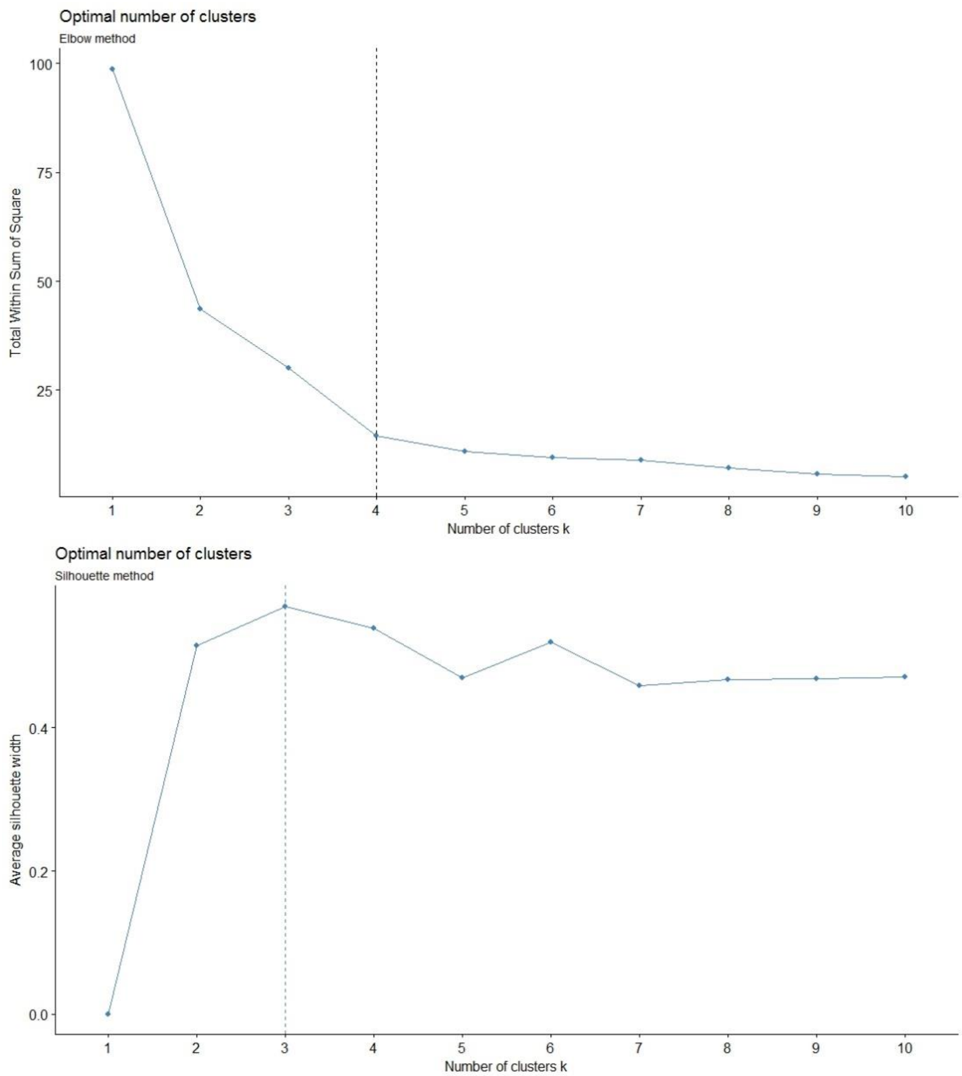


Figure 19. Plots for an optimal number of clusters calculated with elbow and average silhouette method. The algorithm for the calculation and plots were from Kassambara (2017).

3 clusters

When the number of clusters is set to three, the final cluster centres are set as follows: A: English-dominant, B: MT-dominant, C: MT Variation-dominant. Figure 18 shows the cluster centre means and the centroid exemplars of each cluster when *K* is set at 3 for MT matrix at 6-months. About half of the children (50.4%) fell into the Cluster B.

	Cluster Centres (3-cluster MT matrix 6-months)		
	Mother Tongue (A)	English (B)	MT Variation (C)
English	.18	.69	.17
Mother Tongue	.79	.27	.17
MT Variation	.02	.03	.66
Other MT	-	-	-
Other MT Variation	-	-	-
Others	-	-	-
N	228	183	33
Range of distances	.026 – .478	.036 – .533	.080 – .486

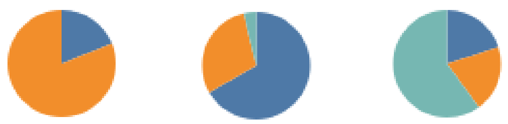


Figure 20. Means of language input for each cluster when the number of clusters is set at 3 (MT matrix at 6-months) with the centroid exemplar of each cluster. N represents the number of children in each cluster; The Range represents the Euclidean distance of the individuals in the cluster from the cluster centre. At the base of the table, a pie chart represents the centroid exemplar (the most typical child) from each cluster.

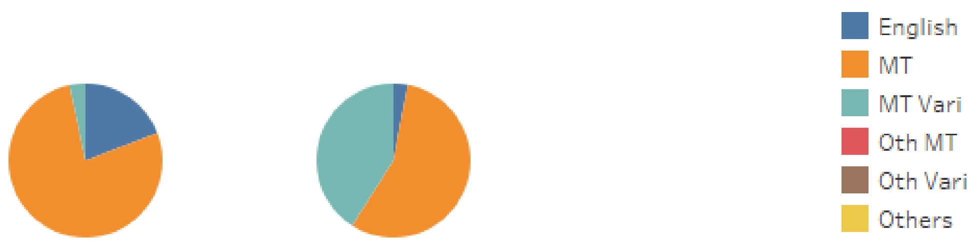


Figure 21. Variation of language input patterns within a cluster: Centroid exemplar of MT-dominant cluster (6-months) (left); individual furthest away from the cluster centre (right).

Figure 21 shows the visualisations of MT matrix of the centroid exemplar of Cluster B (MT-dominant) and the individual with the furthest Euclidian distance from the cluster centre. From the examples of children shown in Figure 21, the child with proportions of language input closest to the cluster means is vastly different from the child furthest away from the cluster centre; the latter receiving much more input from regional variants of their Mother Tongue and much less English input from their caregivers. As there is much variation in the language input among the children of the cluster, we increased the number of clusters by 1 to further differentiate the children.

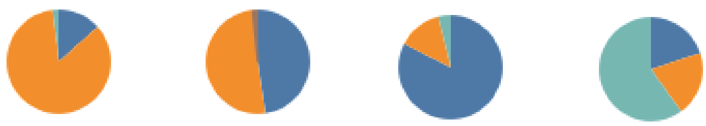
4 clusters

When the number of clusters is set to 4, a new cluster centre accounting for balanced language input emerged. The final cluster centres are as follows: A: MT-dominant, B: English-dominant, C: MT Variation- dominant, D: MT – English balanced. Figure 22 shows the means of each cluster centre and the visualisations of each centroid cluster when *K* is set at 4 for MT matrix at 6-months.

Cluster Centres

(4-cluster MT matrix 6-months)

	Mother Tongue (A)	English & MT (D)	English (B)	MT Variation (C)
English	.13	.46	.80	.16
Mother Tongue	.85	.51	.14	.16
MT Variation	.02	.03	.04	.68
Other MT	-	-	.01	-
Other MT Variation	-	-	-	-
Others	-	.01	-	-



	Mother Tongue (A)	English & MT (D)	English (B)	MT Variation (C)
N	179	124	109	32
Range of distances	.005 – .496	.034 – .386	.027 – .555	.092 – .493

Figure 22. Means of language input for each cluster when the number of clusters is set at 4 (MT Matrix at 6-months) with the centroid exemplars of each cluster. The clusters are arranged according to the sizes of clusters. N represents the number of children in each cluster; The Range represents the Euclidean distance of the individuals in the cluster from the cluster centre. At the base of the table, a pie chart represents the centroid exemplar (the most typical child) from each cluster.

This solution for 4 clusters is the most stable; occurring 16 times out of the 20 times the analysis was run on 20 sets of randomly-ordered data. 99.9% of the time the children remained the same cluster across all 20 runs. The emergence of a balanced MT and English input group is reflective of the previous clustering based on the earlier language input matrix. The total number of Mandarin/Malay/Tamil and English balanced input children was 124 (see Clusters H, I, J in Figure 15); the same number as in the current balanced MT-English input cluster (Cluster D).

Clustering beyond 5 groups

As the purpose of the cluster analysis is to identify groups of children with the most similarities in their language input patterns, further exploration with the number of clusters set at 5, 6, 7, 8 and 9 were run. The aim of our exploratory analysis to find clusters of children with the most within-group similarities, it is not our intention to have as many clusters as possible but rather a number of clusters that have members with high in-group similarity and outgroup differences. There is no fixed criterion for setting *K* and the process of defining *K* is subjective (Jain, 2010). As we increase the number of clusters by one for each round, we are careful not to overfit the data; where the clusters do not provide insights to help us understand the language input patterns of the children in our cohort.

As the number of clusters increases beyond 5, we observed that the larger clusters where the majority of the children are, remained the same while the smaller clusters are partitioned further to highlight the differences between them. Information regarding the clusters and their centroid exemplars are shown in Appendix 3. The further distinction did not serve to provide more noteworthy differences among the clusters but separated the children further according to slight differences in English input in their environments. 4 clusters were, therefore, the most informative groupings of children according to their patterns of language input.

Clutster difference on expressive vocabulary

In order to evaluate whether the K-means clustering is a useful way of characterising the language mixes of Singaporean children, we compare the Mother-tongue clusters at age 6m with the standardised test of English expressive vocabulary, the Communicative Developmental Inventories (CDI). The same children in the cohort are tracked from year to year and at 24 months, the Singapore English adaptation of the Communicative Development Inventories (Tan, 2009) was completed by the parents. The Communicative Developmental Inventories (CDI) is a well-studied and widely replicated tool used to gauge expressive vocabulary size of children (Reilly et al., 1993). Parents were given a catalogue of words in

30 different categories, ranging from animals to body parts and grammar prepositions, and asked to select the words their 24-month-old used regardless of the accuracy of pronunciation. Each word checked on the list is given one mark. A total score was tabulated by summing the marks across all 30 vocabulary categories. A higher CDI score is indicative of a larger expressive vocabulary size at age 24 months.

346 households from the GUSTO cohort completed the 24-month CDI, and scores ranged from 0 – 651, *M* = 161.1, *SD* = 137.3. The overlap between the 6-month Language Background Questionnaire and 24-month CDI is 275 households, their mean CDI score is 162.0, *SD* = 139.7. The CDI scores calculated followed a non-normal distribution, so non-parametric tests on SPSS (IBM Corp, 2016) were conducted. The amount of English language input at 6 months was significantly positively correlated to the CDI score at 24 months, $\rho(275) = .204$, $p < .001$, $d = .452$, $R^2 = .043$. Post-hoc power analysis was conducted with G*Power (Faul, Erdfelder, Lang, & Buchner, 2007). The power (1 – β error probability) to detect this moderate effect for the positive correlation was more than .999 with the sample size of 275.

The children in each cluster differed significantly in their CDI scores (MT-Variation group was omitted for the ANOVA analysis because their cluster was too small), $F(2,251) = 5.65$, $p = .004$, MT-dominant: 131.9 (*SD* = 115.1, N = 108), English-MT balanced: .180.5, (*SD* = 135.8, N = 73), English-dominant: 197.2 (*SD* = 164.4, N =73), MT-Variation: 129.4 (*SD* = 145.0, N = 21).

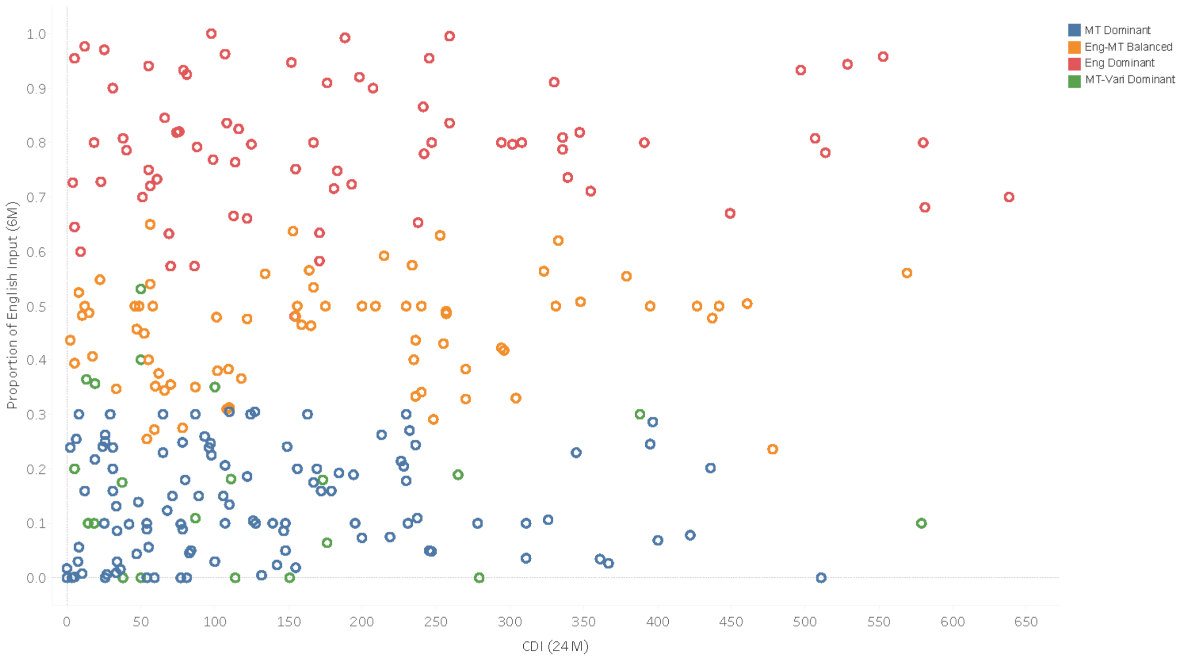


Figure 23. Scatterplot showing the relationship between the proportion of English input a child received at 6 months and their CDI score (expressive vocabulary) at 2 years. The children are characterised by their clusters accordingly.

Though children are not verbally producing any words at 6 months, it appears that the more English used by caregivers to speak to them at that age, the larger their expressive English vocabulary size would be at 2 years. The children who were hearing the most English in their environments (English-Dominant cluster) outperformed their peers and were shown to have more expressive vocabulary at 2 years.

By increasing the number of clusters without validation, there is a risk of overfitting the data; looking into the idiosyncrasies between the language use of caregivers with the children instead of focusing on grouping similar children together. Therefore, for language use by caregivers when speaking to the child at 6-months, we conclude that the children are best grouped into 4 clusters: MT-dominant, English- MT Balanced, English-Dominant, and MT-Variation Dominant. Table 15 shows the distribution of children in each cluster based on their Mother Tongues as determined by the use of Mother Tongues by their parents.

MT / Cluster	MT-Dominant (A)	English – MT Balanced (D)	English-Dominant (B)	MT Variation Dominant (C)	Total
Chinese	**88**	**60**	**77**	**12**	237
Malay	**63**	**51**	**15**	**4**	133
Tamil	**27**	**13**	**16**	**16**	72
Others	**1**	**0**	**1**	**0**	2

Table 14. The number of children in each cluster according to their designated MT (based on MT matrix at 6-months). The shades of blue represent the percentage band of the N in each subgroup corresponding to the total number in their group, starting at 10-20% band, with a 10% increment (lightest to darkest) till 40-50% band. The cells corresponding to the 'Others' group are left uncoloured.

For the Chinese-hearing children, the spread of members across MT-dominant, English-dominant and balanced input clusters is rather even. 5% of the children are receiving language inputs that are dominant in a Chinese dialect.

For the Malay-hearing children, most of them are receiving either a Malay-dominant or English-Malay balanced input from their caregivers. For the Tamil-hearing children, approximately 22% of them are receiving a language input dominant in an Indian language that is not Tamil.

Analysis C: Language Use of Caregivers with 18-month olds

When the child was 18 months old the same language background survey was sent to the parents and data collection from the second time-point was restructured in an identical fashion as the first time point. At the second time-point at 18 months, we are interested to see if the caregivers change their language input patterns as their child began to talk. Regarding linguistic development, these children may have experienced an exponential increase in vocabulary size and start to develop the use of syntactic structures at 18 months (Kuhl, 2011b). At 18 months, the children were most likely also more responsive to linguistic cues and interaction from their caregivers (Igualada, Bosch, & Prieto, 2015); this may result in the caregivers intentionally varying the patterns of their language(s) use. Furthermore, at this age, the children may have different care patterns so the combination of the caregiver(s), languages used, and the hours of exposure may differ from the 6-months time point.

Methods

Parents of 365 children from the GUSTO cohort returned completed questionnaires at the 18-month time-point, if which there was an overlap with 302 children from the first time-point. There was more missing data at the 18-month time-point. Identical procedures to account for missing data conducted for the data at 6-month time-point was applied to the data collected at the18-month time-point. The missing info on language use was especially prevalent for the language use of domestic helpers. 92 households reported that domestic helpers spend time caring for their 18-month-olds, but only 34 households reported the language use of their domestic helpers. For the remaining 54 households, we note that there will be a certain degree of error in the final language input matrix of the child. As the data are archival, we are not in a position to follow up with the families in question, however, future researchers should bear in mind that missing data is likely to be higher at this time point, and should endeavour to follow up with families around the time of data collection for completeness.

As with the data from 6-month time-point, we restructured the language input matrix into the Mother Tongue matrix. The same iterative process of *K*-means cluster analysis on the Mother Tongue matrixes of 365 18-month old children was conducted to meaningfully group the children together. We also determined the optimal number of clusters with the elbow method and the average silhouette method as we did with the data from the 6-month time-point. For the ease of comparison, the clusters for all *K* values are given consistent letter names in each cluster analysis, using the following code: A – MT-dominant, B – English-dominant, C – MT Variation-dominant, and D – English-MT balanced. All the figures reported about the cluster make-up are arranged according to the size of the cluster.

To check for the validity of the number of clusters, we also checked for possible cluster differences and the CDI scores at age 2. A comparison of the clusters at both time-points and the movement of cluster members are also presented in this following section.

Results

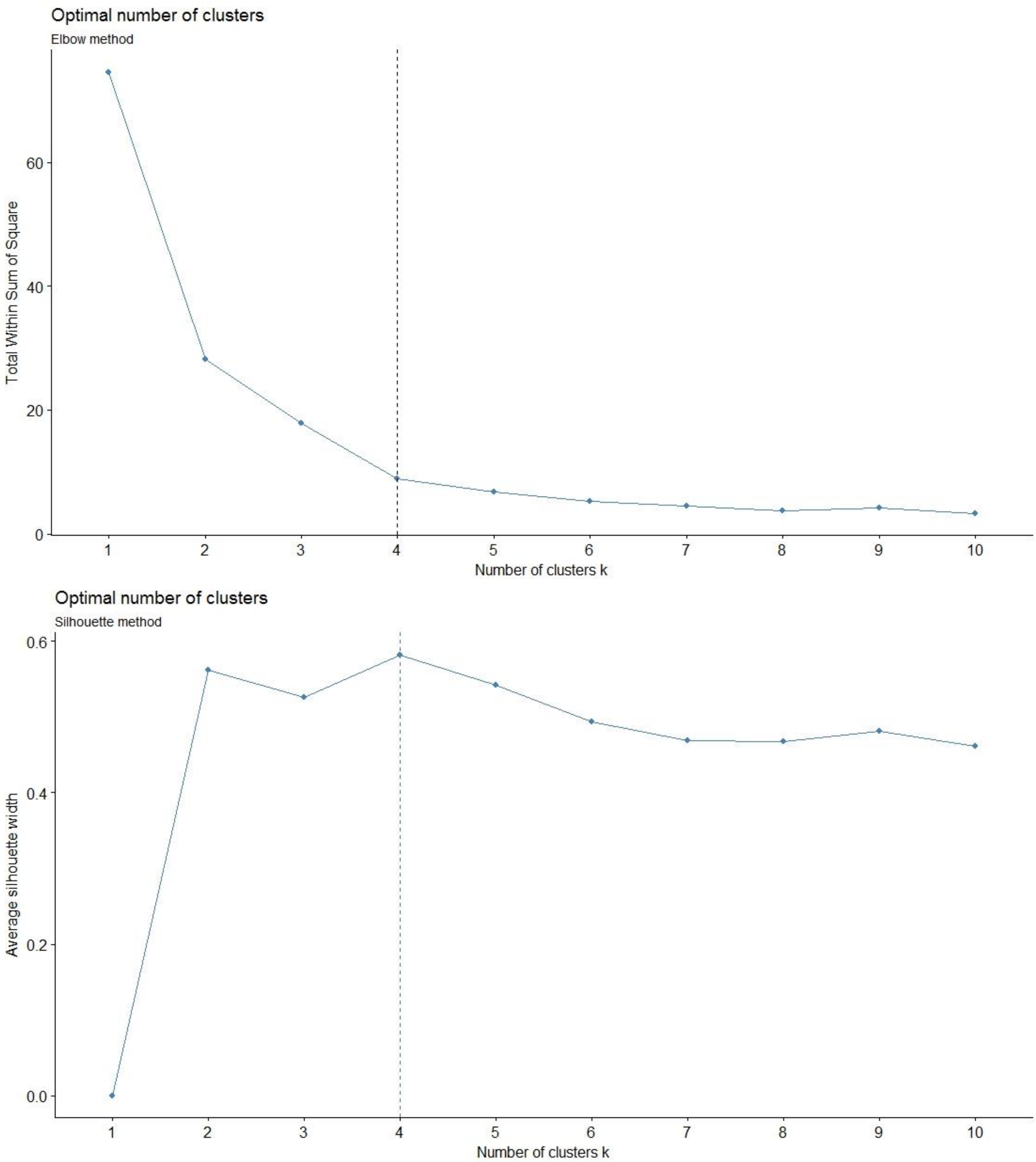


Figure 24. Plots for an optimal number of clusters calculated with elbow and average silhouette method. The algorithm for the calculation and plots were from Kassambara (2017).

Figure 24 shows that the optimal number of clusters for the Mother Tongue matrix at 18 months is 4 as calculated by the elbow method and the average silhouette method in R (Kassambara, 2017).

4 clusters

	Cluster Centres (4-cluster MT matrix at 18-months)			
	MT	English - MT	English	MT Variation
	(A)	(D)	(B)	(C)
English	.13	.45	.81	.28
Mother Tongue	.85	.53	.16	.04
MT Variation	.02	.02	.03	.68
Other MT	-	-	-	-
Other MT Variation	-	-	.01	-
Others	-	- (.01)	-	-
N	130	111	109	15
Range of distances	.021 – .333	.025 – .409	.026 – .375	.048 – .554

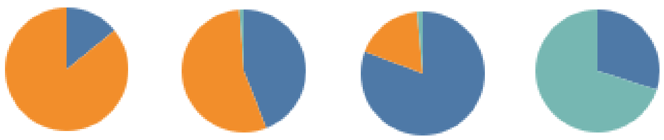


Figure 25. Means of language input for each cluster when the number of clusters is set at 4 (MT matrix at 18-months). Centroid exemplars of each cluster are shown. N represents the number of children in each cluster; The Range represents the Euclidean distance of the individuals in the cluster from the cluster centre. At the base of the table, a pie chart represents the centroid exemplar (the most typical child) from each cluster.

Figure 25 shows the means of each cluster centre and the centroid exemplars when *K* is set at 4 for the MT matrix data at 18-months. This solution was the most stable; occurring 18 times out of the 20 runs of analysis we ran on twenty different randomly-ordered sets of data.

For the children, the amount of English exposure from their caregivers at 18-months is strongly positively correlated to their expressive vocabulary size (CDI) at 24 months, $\rho(262) = .325$, $p < .001$, $d = .570$, $R^2 = .092$. Post-hoc power analysis showed that power to detect this moderate effect for the positive correlation was more than .999 with the sample size of 262. Scatterplot illustrating the amount of English input at 18-months and CDI scores at 24-months is shown in Figure 26.

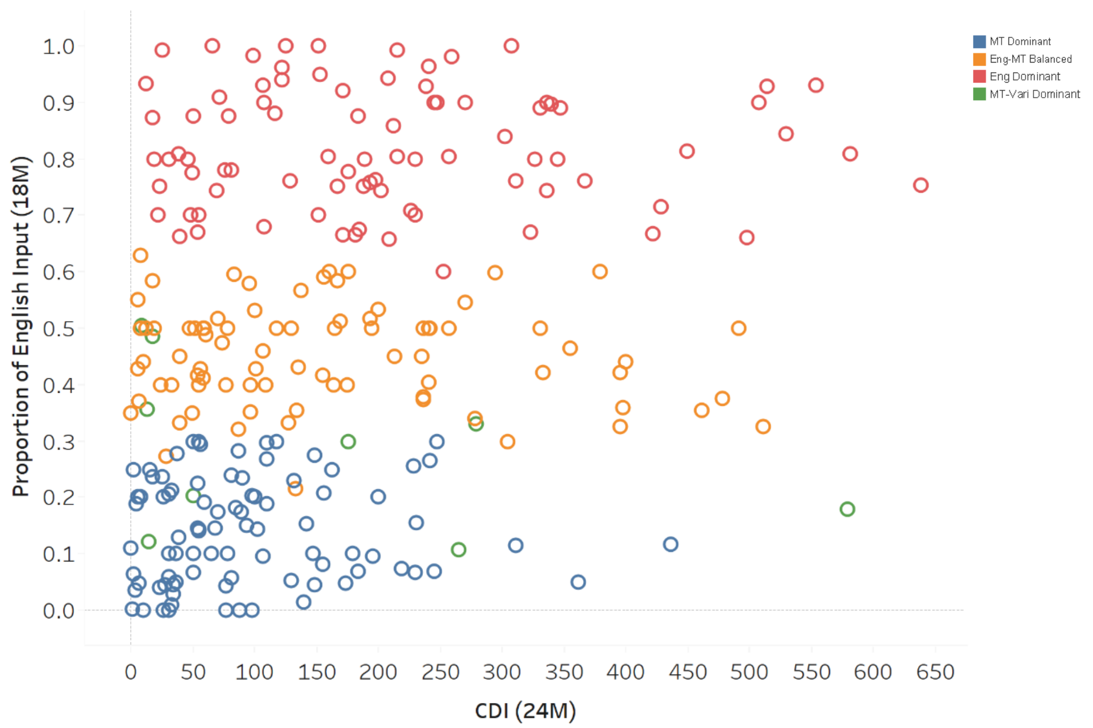


Figure 26. Scatterplot showing the relationship between the proportion of English input a child receives at 18 months and their CDI score (expressive vocabulary) at 2 years.

The children in each cluster differed significantly in their CDI scores (the MT-Variation cluster was not included in the ANOVA analysis as the cluster was too small), $F(2,252) = 18.2$, p <.005, MT-dominant: 96.2 (*SD* = 85.3, N = 87), English-MT balanced: 160.7, (*SD* = 131.9, N = 82), English-dominant: 210.3 (*SD* = 148.2, N = 84), MT-Variation: 155.9, (*SD* = 193.3, N = 9).

We concluded that for language use by caregivers when speaking to the child at 18 months, the children are best grouped into the same four clusters as at 6-months: MT-dominant, English- MT Balanced, English-Dominant, and MT-Variation Dominant. As with the 6-month time-point, the amount of English exposure at 18 months predicted the children's expressive vocabulary (CDI) score at 2 years. It is unsuprising that more variance of the CDI scores is accounted for by the amount of English expousre at the18-month time-point (9.2% compared to 4.3%) as the time between the two tests are closer.

The number of children classified according to their Mother Tongues and the clusters is presented in Table 15. We reduplicated Table 14 below for easier reference.

MT / Cluster	**MT-Dominant**	**English – MT Balanced**	**English-Dominant**	**MT Variation Dominant**	**Total**
Chinese	**64**	**62**	**58**	**12**	196
Malay	**44**	**38**	**32**	**1**	115
Indian	**21**	**10**	**19**	**2**	52
Others	**1**	**1**	**-**	**-**	2

Table 15. The number of children in each cluster for MT matrix (18-months) according to their Mother Tongues; determined by the parents' use of MT with their child. The shades of blue represent the percentage band of the N in each subgroup corresponding to the total number in their group, starting at 10-20% band, with a 10% increment (lightest to darkest) till 40-50% band. The cells corresponding to the 'Others' group are left uncoloured.

MT / Cluster	**MT-Dominant**	**English – MT Balanced**	**English-Dominant**	**MT Variation Dominant**	**Total**
Chinese	**88**	**60**	**77**	**12**	237
Malay	**63**	**51**	**15**	**4**	133
Tamil	**27**	**13**	**16**	**16**	72
Others	**1**	**0**	**1**	**0**	2

Table 14 (reduplicated). The number of children for each cluster according to their designated MT (based on MT matrix at **6-months**).

As shown in Tables 14 and 15, there is an increase in the number of members for the English-MT balanced group or the English-dominant clusters across all the Mother Tongues/ethnicity groups.

Language Backgrounds: Changes across two time-points

We tracked the changes of clusters if any of the 302 children whose language data was present in both time-points. Figure 27 illustrates the movement of cluster members across the time-points.

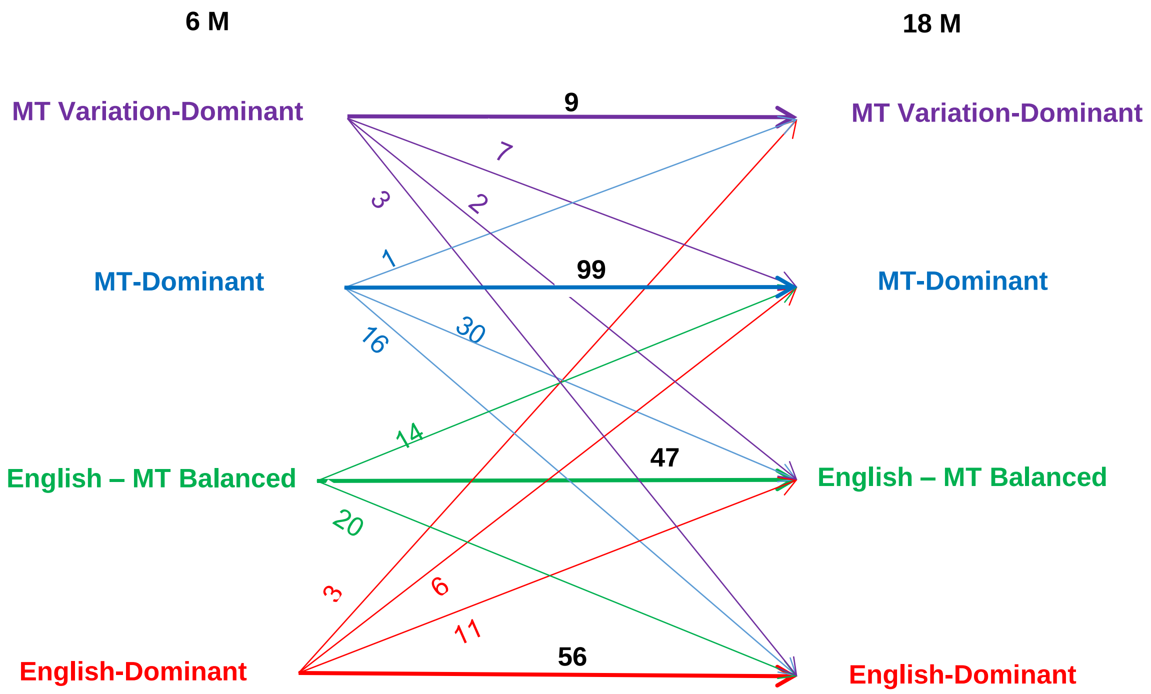


Figure 27. The movement of cluster members across both time-points (6 months → 18 months). The trend of moving towards balanced or English-dominant groups as the child gets older is present. In each group, the thickest line represents the most common pattern of movement.

The MT-dominant cluster remains the largest across both time-points. For each group, the most likely outcome is the maintenance of cluster identity. While the majority of the children remain in the same clusters (i.e. the proportion of language(s) their caregivers use with them did not change from 6-months to 18-months), 37.4% of the children did experience a change in the patterns of their language input from their parents and caregivers. When change does occur, the trend is towards the use of official or ‘standard’ languages with the children.

When considering the bilingual language policy exercised in Singapore, one expects that parents and caregivers may make a conscious effort to move towards providing the child with a balanced English – Mother Tongue input as their child grows and gradually becomes more verbal and more responsive to linguistic cues. While the movement of children into the various clusters from 6-months to 18-months seems multi-directional, in general, households shifted their language use towards more state-encouraged standards e.g. balanced bilingual or English-dominant and moved away from being MT-Variation dominant. There is also a shift towards the use of the officially recognised Mother Tongue e.g. Mandarin instead of Dialect with the children as they get older.

Discussion: Language Background of pre-lingual Children in Singapore

In this chapter, observable patterns and trends of languages that were used with Singaporean children at 6 months and at 18 months were presented. Archival data were gathered from the Language Background Questionnaires that were completed by parents. After accounting for missing data and restructuring the variables, we were able to present a descriptive analysis of the language backgrounds of Singaporean children at 6-months old. We also created a language input matrix for each child. The unique make-up of language input at 6-months for each child is presented as pie charts in Figure 8.

Following the descriptive analysis, we conducted K-means clustering analysis to allow our data to inform us of the different possible groups of language backgrounds of Singaporean children. We are aware that given the nature of the cluster analysis run in SPSS, we may encounter many possible solutions for each *K* selected. For each *K* presented, we selected the model that occurred the most frequently out of 20 runs with different initial clusters. Furthermore, we also checked for an optimal number of clusters using the elbow method and the average silhouette method (Kassambara, 2017).

K-means clustering ran on the language input matrixes of all children at 6-months showed that the children are well described at 9 clusters. By increasing the clusters to 12, we began to see finer distinctions in the smaller clusters while the larger clusters remain largely the same. Given the similarities which arose between the language-dominant and balanced language input groups from the clustering results of the language input matrixes, the language input matrixes were re-coded into the **Mother Tongue matrixes**. Categories for language use in the MT matrix are: English; Mother-Tongue; Mother Tongue Variation (regional variants of one's ascribed Mother Tongue e.g. Chinese dialects); Other Mother Tongues (other MTs not ascribed to your ethnicity); Other Variations (regional variants of other MTs); and Others.

For language backgrounds of the children at the 6-month and 18-month time-points, the children are best categorised into **four** clusters: Mother Tongue–dominant, English–Mother Tongue balanced, English–dominant, and Mother Tongue Variation–dominant. At both time-points, the Mother Tongue–dominant group had the most members. At 18-months, the percentage of English use by the caregivers for each cluster ranged from 13–81% and the amount of English present in a child's environment correlated with the child's expressive English vocabulary size (measured by parent-reported CDI) at 2 years. In this study, we were particularly interested in the proportion of English input, as almost all the available standardised tests for reading administered in Singapore is only in English and we will be exploring the relationship of cluster identity and predictors of reading in the following chapter.

Speech and language development researchers in Singapore agree that bilingualism in Singaporean children is difficult to define and quantify (En, Brebner, & McCormack, 2014). For the purposes of experimental work, some researchers have chosen to define Singaporean children with external norms, often separating the children into two or three groups (Singh, 2017). From our data, we suggest that Singaporean children fall into four main groups. We propose the cut-off for considering a child dominant in English is an input above 61 %. To be considered a balanced English-MT bilingual, a child should be receiving 32- 61% English input in his environment. For the small group of children whose dominant language is an MT-variation language (Chinese dialects, Bahasa Indonesia or an Indian language that is not Tamil), they pattern more closely with MT-dominant children, receiving up to 30% English in their language environments. This information helps us to characterise the children in our study, many of whom attempted our sound-symbol matching task at 6 years of age. It also helps to characterise children in Singapore in general; providing more accurate cut-offs for classification for future experiments.

Singapore census data shows that language use continues to shift towards the use of official languages, in particular, English, and away from vernaculars especially for the younger generations (Department of Statistics of Singapore, 2010). We observe a similar shift supporting this trend with the English-dominant and English-MT balanced groups increasing in size as children grew from 6-months to 18-months. This shift may be due to pragmatic reasons of change in care relationships e.g. placing the child into an English-speaking child-care class or it may be due to parents becoming more vigilant with the language(s) they choose to use with their child as they grow older and become more responsive. Future studies can take a more nuanced approach to how and why this shift occurs.

Commonly inculcated in Singaporeans via narratives surrounding bilingual language policies is the belief of ‘language as a tool’, and a good command of English is often linked to personal education and workplace gains and upward social mobility (Bolton & Ng, 2014). Furthermore, Pakir (1997) notes that the shift towards using more English is common among families trying to better prepare for their young children for primary school. The shift we observed in our data may be a reflection of attitudes towards languages by Singaporean parents and caregivers.

Chapter 3: Sound-Shape Matching Task & Predictors of Reading

Current literature regarding behavioural tasks to examine sound-symbolism have shown consistently shown that adults systematically prefer matching certain sounds to certain shapes. For example, adults match rounded vowels and voiced bilabial consonants e.g. [o, u, b, m] with curvy shapes and high front vowels and voiceless stop consonants and e.g. [i, e, t, k] to spiky shapes. The well-studied bouba-kiki paradigm is a two-alternative forced choice task used to elicit this sound-to-shape matching.

In accordance with our aim to examine the developmental trajectory of sound-symbolism and dyslexia, we adapted the bouba-kiki paradigm and designed a child-friendly task – the Alien Zoo task. For Study 1, we present the methods, results and findings of the Alien Zoo task conducted with pre-schoolers from the GUSTO cohort and adult controls. In the current literature, we do not have norms for children performing a task with the bouba-kiki paradigm. With our findings, we establish for the first time, norms for bilingual/multilingual pre-schoolers on a sound-to-shape matching task. We discuss the results of the adult controls in relation to studies in the current literature.

For Study 2, we present analyses exploring the correlations between the children's Alien Zoo task score with results from measures of predictors of early reading (phonological awareness, vocabulary size, and letter identification) taken at an earlier age. We preregistered these analyses (see Appendix 5) and predicted a positive correlation between the best-known risk factor for dyslexia (phonological awareness) and sound symbolic sensitivity (as measured in the Alien Zoo task). Then, we discuss these results in relation to dyslexia and reading development in children.

In Study 3, we present exploratory analyses conducted to examine the relationship of cluster identity based on language backgrounds (as determined in the previous chapter) with results from the Alien Zoo task. We also examined if these clusters correlated with known predictors of English reading ability. We are interested to see if the cluster analysis in the preceding chapter could provide a salient source of difference between individuals' performance on the Alien Zoo task and predictors of English reading ability.

Study 1: Alien Zoo Task

Participants

The children were recruited for the task as part of an ongoing longitudinal cohort study by GUSTO. 413 children were available for the study as part of the GUSTO programme of which a total of 388 attempted the task. The research assistants who conducted the experiment kept a log of each child's behaviour during the task. Prior to analysis, results from 10 children were removed because it was reported that the child did not perform the task according to instructions e.g. selecting the shape before the sound file was played. Additionally, another child was also removed from the analysis as the participant chose to drop out of the task after two trials. Of the 377 child participants whose results were analysed, 203 were male and174 were female. The children were aged between 5 years 10 months and 6 years 6 months (*M* age = 5 years 10 months, *SD* = 3.36). The GUSTO study was approved by the National Healthcare Group Domain Specific Review Board (NHG DSRB) and the Sing Health Centralized Institutional Review Board (CIRB). Written consent was obtained from mothers at the time of the test.

Adult participants were asked to participate in the study through online invitations. Of the total of 111 adult participants, 80 of them were undergraduates enrolled in an introductory psychology course and were given course credits for their participation. 2 participants did not indicate their age or gender. The rest of the participants were aged between 18 years and 40 years (Mean age bracket: 21-30 years), there were 39 male participants and 70 female participants. The study was approved by the NTU IRB. Digital consent was obtained before the participants proceeded with the task presented on the online platform, Qualtrics (Qualtrics, 2005).

Stimuli

Previous bouba-kiki studies involving children reviewed earlier used visual stimuli which were plain line drawings or coloured blobs (Maurer et al., 2006; Occelli et al., 2013). These visuals were neither interesting nor visually engaging for children. In order to ensure that the children enjoy the task and pay attention the stimuli, we used visually rich and coloured stimuli adapted from previous studies (Lai, 2015; Styles & Lai, 2017). The images were high definition photographic images of rounded and spiky shapes presented on a black background. These images were adapted from photographs of pollen and microscopic organisms and then coloured red, blue, green or purple. The images were either distinctly curvy or spiky. This judgement was made by a separate group of adults from the Cool Virus Set study where they were asked to rate the visual stimuli on dimensions of Shape (Round-Spiky), Texture (Rough – Smooth), and Plasticity (Hard-Soft) (Styles & Lai, 2017). To make the stimuli appealing to children, we added googly eyes on each image.

The images were used to create 16 pairs of colourful distinctive and contrastive visual stimuli presented on a black background. Each pair consisted of one round and one spiky shape in the same colour.

Some of the auditory stimuli used in previous studies were also questionable in its ability to elicit the desired bouba-kiki effect e.g. *Shhhh* and *Mmmm* (Oberman & Ramachandran, 2008). We avoided the use of phonemes like /s/, /z/, and /r/ as they have not been widely established in their ability to elicit the desired sound-symbol matching effect. Auditory stimuli in our study were also adapted from the same previous study by Styles &

Lai (2017). Auditory stimuli were pseudo-words created from phonemes widely recognised as ‘round’ (/m/, /o/, /u/, /b/,/l/) and ‘spiky’ (/k/,/t/, /p/, /i/, /e/) phonemes (D’Onofrio, 2014; Nielsen & Rendall, 2013; Styles & Gawne, 2017a). The pseudo-words were randomly generated and followed consonant-vowel structures commonly used in these sound-symbolic tasks e.g. CVCVCV (*maluma*). Each pair of visual stimuli was matched with one auditory token; creating 16 unique yoked trials. The full set of yoked trials is presented in Figure 28.

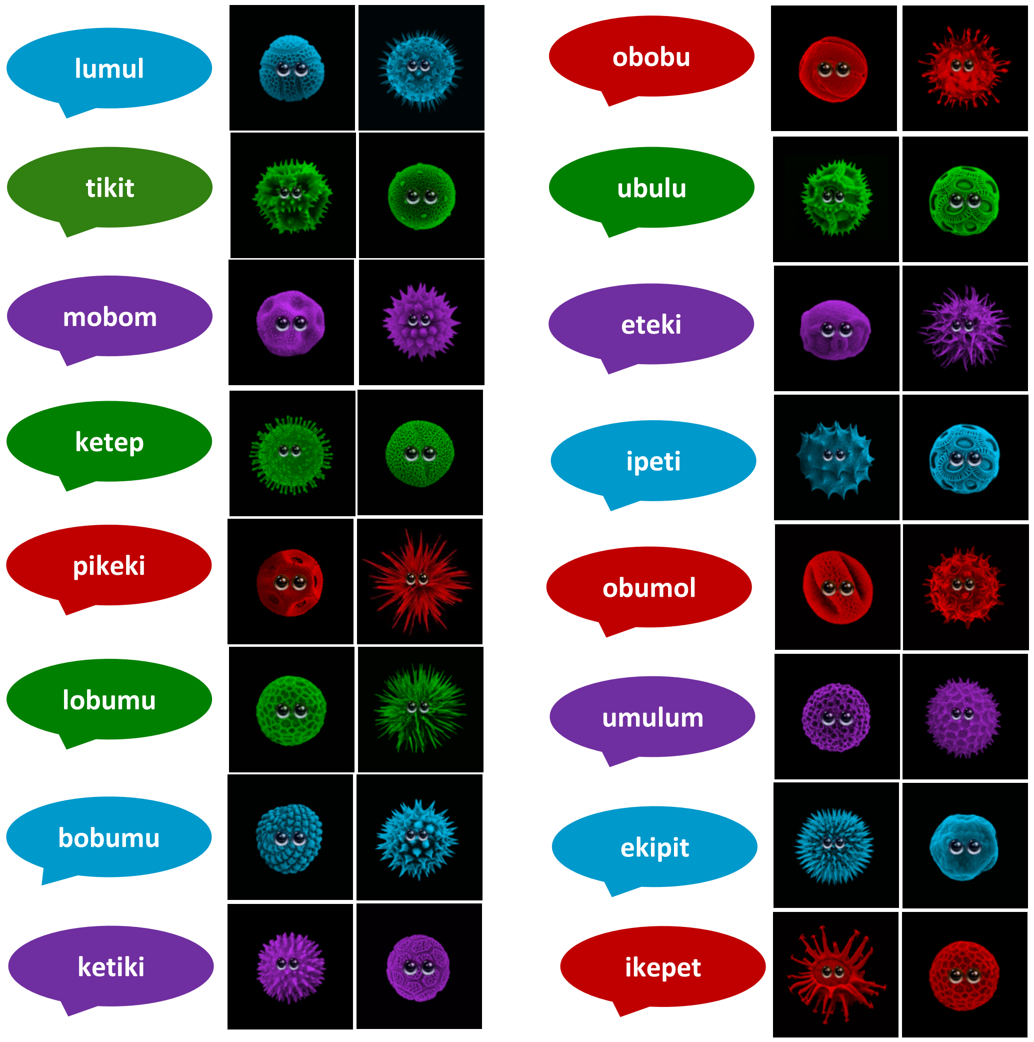


Figure 28. The yoked pairings of child-friendly visual stimuli and pseudo-words used for all 16 trials.

The order of the trials was randomised and the position of the spiky and rounded shapes was also randomised. None of the shapes was repeated. The study was configured and presented on a computer screen using the online programme for surveys, Qualtrics (Qualtrics, 2005).

Procedure

Children were tested in a quiet room individually with a research assistant who read aloud the instructions and questions to them. They sat beside the research assistant while facing the computer screen. Children were asked to choose from English, Mandarin or Malay versions of the task, depending on which language they were most familiar with. The yoked tokens remained unchanged for all versions of the task, only the onscreen written questions and storyline were translated. The translations were cross-checked by native speakers to ensure that the language used was child-friendly. Table 18 shows the number of children and adults who took the task in various languages offered.

Language (version of the task)	Children	Adults
English	353	111
Mandarin	21	-
Malay	3	-
Total	**377**	**111**

Table 16. Distribution of the children and adults according to the language they chose to complete the task in.

353 children completed the task in English, 21 completed the Mandarin version of the task and 3 completed the Malay language version of the task. A child who completed the English version of the task did so with verbal Tamil translation from the research assistant.

Adult participants completed the task with no guidance and while they were also given the option to choose between different language versions of the task, all participants completed the task in English.

In order to make the task interesting for the child participants, a storyline was created. All participants were told that they had arrived at the Alien Zoo where the aliens – represented by various shapes – were loose and jumbled up. The participants were told Zookeeper Clarice took care of two groups of aliens which looked different. They were asked if they could help Zookeeper Clarice to sort the aliens.

In each trial, the participants saw a pair of visual stimuli side by side. They were asked to guess, “Which alien do you think has this name [play auditory token]?” The auditory tokens were controlled manually and could be played repeatedly if necessary. As a participant should only make a choice after the auditory token was played, results from child participants who did not do so were removed before analysis. The children were told to point to their choice on the screen and the research assistant would click on the shape shown on the computer programme. When the shape was selected, it was highlighted with a red outline on the screen and the research assistant would confirm the child’s choice with them before moving to the next trial. Figure 33 provides an illustration of a trial.

Figure 29. A representation of a trial seen on the computer screen.

The child participants were given the option to drop out of the study after every trial should they want to. The research assistant responded with “good job” regardless of the answer selected by the child and assured them that there were no correct or incorrect answers. A detailed protocol for the experiment, including appropriate responses to say to the children, was provided to all the research assistants involved (See Appendix 4). Adult participants could quit the task at any time by exiting the Qualtrics programme.

After the 8th and 12th trials, messages appeared on the screen to encourage the participants; informing them that they were making progress with the task of helping Zookeeper Clarice. The messages were, “Good work! You have sorted half of Clarice’s aliens!” and “Almost there! You have almost sorted all of Clarice’s aliens!” respectively. After the last trial, a message appeared to thank the participants for their effort. The child participants concluded the task by answering three questions about their language use with their family, teachers, and friends respectively. They chose from a list of languages: English, Mandarin, Malay, Tamil, Hokkien, Teochew, Cantonese, Hakka, Others (specify). Adult participants moved on to another online task.

Results

A response for each trial is considered congruent when a rounded shape is chosen for the pseudo-word with ‘rounded’ sounds (/m/, /o/, /u/, /b/,/l/) and the angular shape for the pseudo-word with ‘spiky’ sounds (/k/,/t/, /p/, /i/, /e/). Figure 30 shows the congruent response for each of the auditory token.

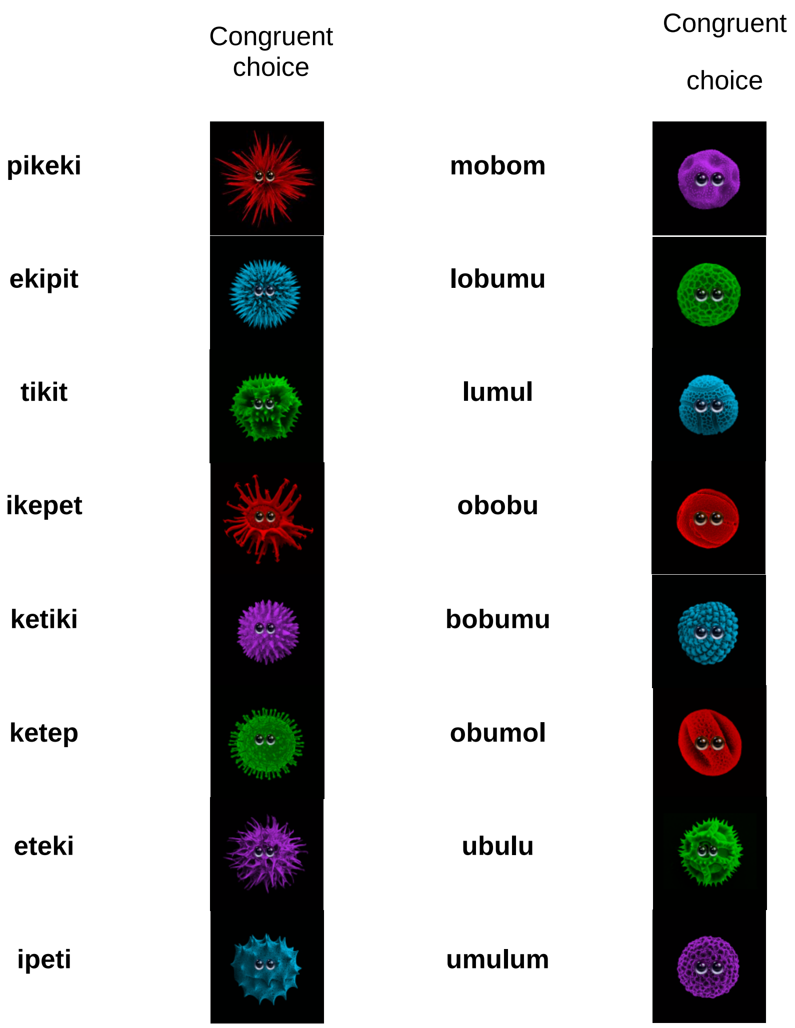


Figure 30. The congruent response for each of the auditory tokens.

The number of congruent responses for each participant was totalled. A total score of 16 demonstrates that the participant chose in the expected direction for every trial, and a total score of 8 indicates making unsystematic choices (half expected and half unexpected i.e., guessing at chance). Figure 31 shows the number of congruent responses made by each participant for the child and adult participants.

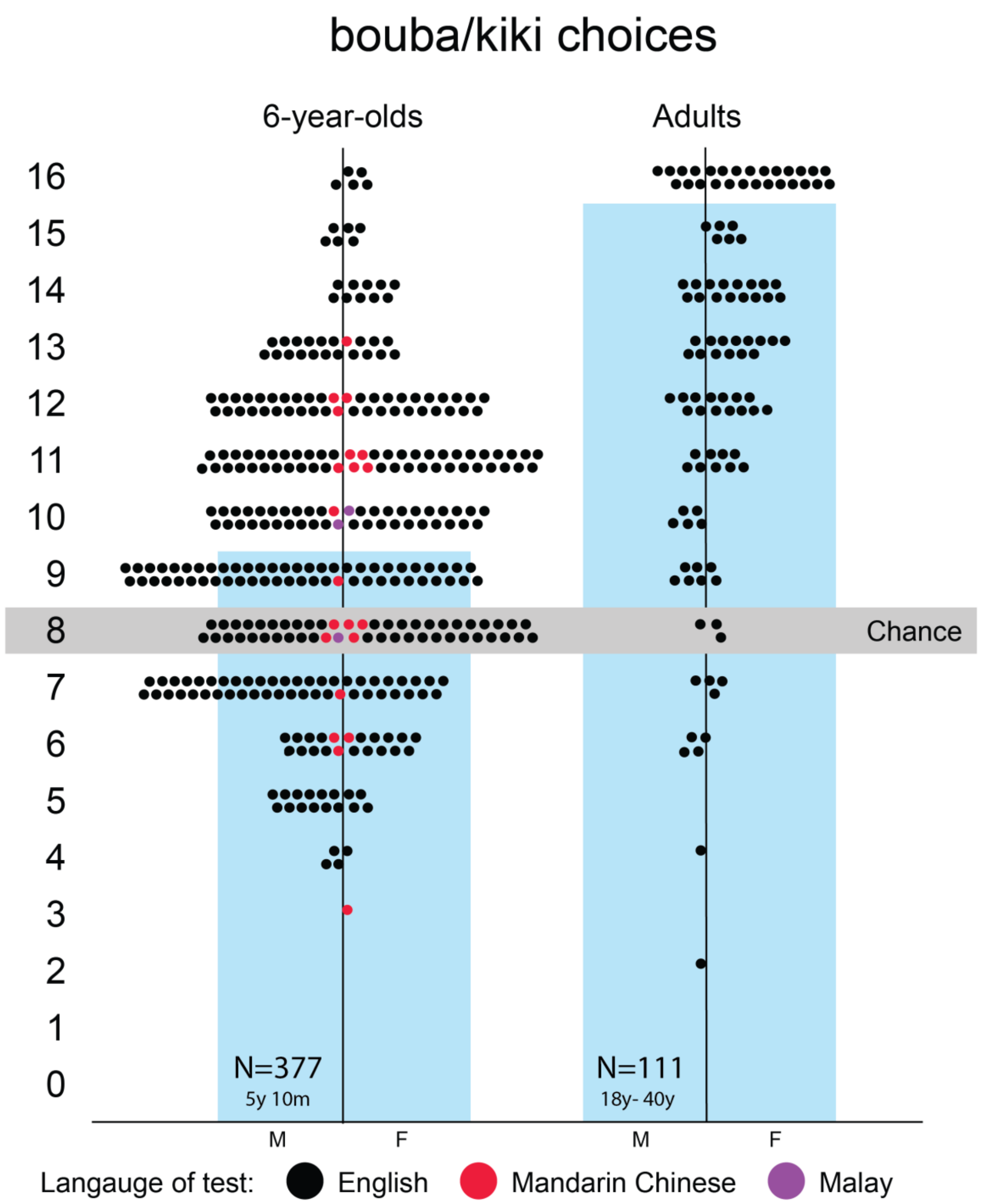


Figure 31. The number of congruent choices made by each participant in the Alien Zoo task, the blue bar represents median for congruent responses for each group. We report the median scores to account for the non-normal distribution of the adults' score and the granularity of our scoring.

Both children and adults chose the congruent responses significantly above chance. Children chose the congruent responses on a median of 9 and mean of 9.48 (SD = 2.54), $t(376)$ =11.32, $p < .0001$, d = .583 against the chance value of 8. The girls (M = 9.82, SD = 2.58) and boys (M = 9.19, SD = 2.47) differed significantly from each other, $t(375)$ = 2.40, p= .017, d = .249.

Adults also guessed significantly above chance so on a median of 13 and mean of 12.52 (*SD* = 3.14), $t(110) = 15.20$, $p < .0001$, $d = 1.44$, The adults made significantly more congruent responses compared to the children, by a difference in mean of 3.04, $t(486) = -10.5$, $p < .0001$.

A post-hoc power analysis on the number of congruent choices made by all the children was conducted with the programme, G*Power (Faul et al., 2007). The power to detect the moderate effect ($d = .583$) with our sample of children was high (.999), suggesting that the sample size of 377 children is well powered. In case the effect size is inflated in the current sample of children, we conservatively estimate that when power is set at .95 and $\alpha = .05$, two-tailed, a sample size of only 34 children would be needed for results of this task to be replicated.

Post-hoc power analysis was also conducted to look into the significant difference in the number of congruent choices made by boys and girls. The effect size was small ($d = .249$), and the observed power was .672. To detect this effect in future studies, a total sample of 838 children would be necessary.

While the adults were guessing significantly above chance, a ceiling effect was found. The large effect size in the adult population ($d = 1.44$) is expected given the current literature on similar 2-alternative forced choice sound-symbol matching tasks for adults. Similarly, post-hoc power analysis revealed that the power to detect this effect was high (.999). For future replications, with minimum power set at .95 and $\alpha = .05$, two-tailed, a sample size of 7 adults would suffice to detect the same effect.

Overall, for the adults, the rate of congruent responses was 78.3% (*SE* = .019). For the children, the rate of congruent responses was 59.2% (*SE* = .008).

Discussion

Children on the Alien Zoo task

Since the advent of the investigation into sound-symbol matching, we have been presented with a number of studies looking at various populations of adults with similar paradigm as the original work of Kohler (1929:1947) (Non-Western language speaking: Bremner et al., 2013; French-speaking: Fort, Martin, & Peperkamp, 2015; English-speaking:Nielsen & Rendall, 2011; Ramachandran & Hubbard, 2001). Evidence from these studies show that sound to symbol matching is not driven by knowledge of Western orthography or language and its effect is consistent across audio or visual presentations of the stimuli (Nielsen & Rendall, 2011). However, only a handful of studies has examined children using the same two-alternative forced choice paradigm (Maurer et al., 2006; Oberman & Ramachandran, 2008; Occelli et al., 2013; Spector & Maurer, 2013). These studies with children have small sample sizes and may have contained stimuli which were not optimal in eliciting a sound-symbol matching effect.

Our study is the first of its kind to examine a large cohort of 377 children performing the bouba-kiki task. With our stimulating child-friendly stimuli and experiment design, children's interest and attention were sustained throughout the task and we report an extremely low dropout rate (one child). Our sample size was large and the spread of the results was wide, so the task was sufficiently sensitive to identify children who were performing differently from the majority of their peers.

Here, we consolidate results of the bouba-kiki task with Singaporean pre-schoolers in the GUSTO study (the largest longitudinal developmental cohort in the country):

(i) Young children in Singapore with varying language backgrounds do make linguistic sound-symbol matches just like adults. The children showed systematic preferences for sounds and shapes; matching speech sounds /b, m, l, o, u/ with curvy shapes and /t, k, p, i, e/ with spiky shapes.

(ii) Children make fewer congruent sound-symbolic matches compared to adults.

The systematic preference is consistent with results from earlier studies on 2.5-year-old children (Maurer et al., 2006; Spector & Maurer, 2013). With their 4-item task and small sample size, Maurer et al (2006) found no overall significant difference in the rate of sound-shape matching between the children and adults. However, they noted that adults in their study made more sound-symbolic choices for two out of four of their test items. For their task, the visual stimuli were black line drawings presented on white paper and the auditory stimuli were verbally repeated by the experimenters. We suggest that there may have been confounding factors in their experimental procedure to account for the lack of difference in the number of congruent choices made between the children and adults. In our study, we used validated visual stimuli and pre-recorded auditory stimuli to minimise possible confounds. There were also more test items in our task which could have helped tease apart the differences between adults and children.

The girls in our study made significantly more congruent choices than the boys. However, the effect size was small and the effect was not well powered. An even larger sample of 838 children would be needed to replicate this finding. Thus, though the effect was significant, it is likely to be due to chance. It is commonly known that there are more boys diagnosed with reading impairments than girls, but the magnitude of gender differences is usually small and remains debatable (Quinn & Wagner, 2015). We think it is unwise to believe that there is a gender bias for sound-symbol matching ability with our results.

Adults on the Alien Zoo task

Styles and Gawne (2017a) conducted a meta-analysis examining the adults' rate of congruent responses when canonical speech sounds known to elicit sound-symbolic effects were used. The average reported rate is 84-94%. These canonical speech sounds were first examined by Köhler (1929:1947) and later adapted by Ramachandran and Hubbard (2001); /b, m, l, o, u, a/ matched with curvy shapes and /t, k, i, e/ matched with spiky shapes. The auditory stimuli used in our study were pseudo-words randomly generated with these sounds and the phoneme /p/. The phoneme /p/ was used previously to successfully elicit matches with spiky shapes (Fort et al., 2015; Nielsen & Rendall, 2011; Peiffer-Smadja & Cohen, 2010) Five tokens containing the phoneme /p/ was used in the word-initial '*pikeki*', word-medial '*ipeti*', and word-final '*ketep'* positions. The phoneme /p/ is a voiceless stop just like /t/ and /k/ known to match with spiky shapes but it is also bilabial (requires both lips as articulators) like /m/ and /b/ that are matched with curvy shapes. Ramachandran and Hubbard (2001) theorised that representations of certain lip and tongue movements in the motor areas are mapped in a non-arbitrary way to certain phonemic representations in the auditory cortex which in turn are mapped to visual representations of certain shapes. For example, the curved mouth-shape is linked to /u/ sound which in turn is linked to the visual percept of a round object.

For exploratory purposes, we reanalysed the number of congruent choices made after removing the five stimuli containing /p/. The rate of congruent choices made was 77.8% (*SE* = .020); still below the rate reported in the meta-analysis (Styles & Gawne, 2017a). Though Ramachandran and Hubbard's theory has yet to be supported with neurological evidence, it is perhaps that the /p/ phoneme is not an ideal speech sound to use for the bouba-kiki paradigm given its auditory properties as mentioned above.

Our auditory stimuli were pseudo-words generated following the phonotactics of English. Phonotactics are the rules governing the combinations of sounds which go together so for English, /skript/ is well-formed but /skrpit/ is not. While all of our adult participants are English speakers, 75.7% of them indicated Mandarin as one of their languages. In our study, there were five auditory tokens – *eteki*, *tikit*, *pikeki*, *ekipit*, *ketiki* – which contained the syllable /ki/. The syllable is not well-formed according to the phonotactics of Mandarin (although, we note that some dialects of Chinese do allow for this). When the five tokens are removed, the rate of congruent choices made was at 79.2% (*SE* = .018), still below the rate reported by Styles and Gawne (2017a).

In Figure 32, we present a forest plot listing results from our bouba-kiki task in comparison to all other adult bouba-kiki studies which used canonical speech sounds. The forest plot was created using the formulas developed by Styles and Gawne for the meta-analysis on canonical bouba-kiki sounds (2017a) and data tables for analysis can be found on the Open Science Framework repository for that article (https://osf.io/wt95v/).

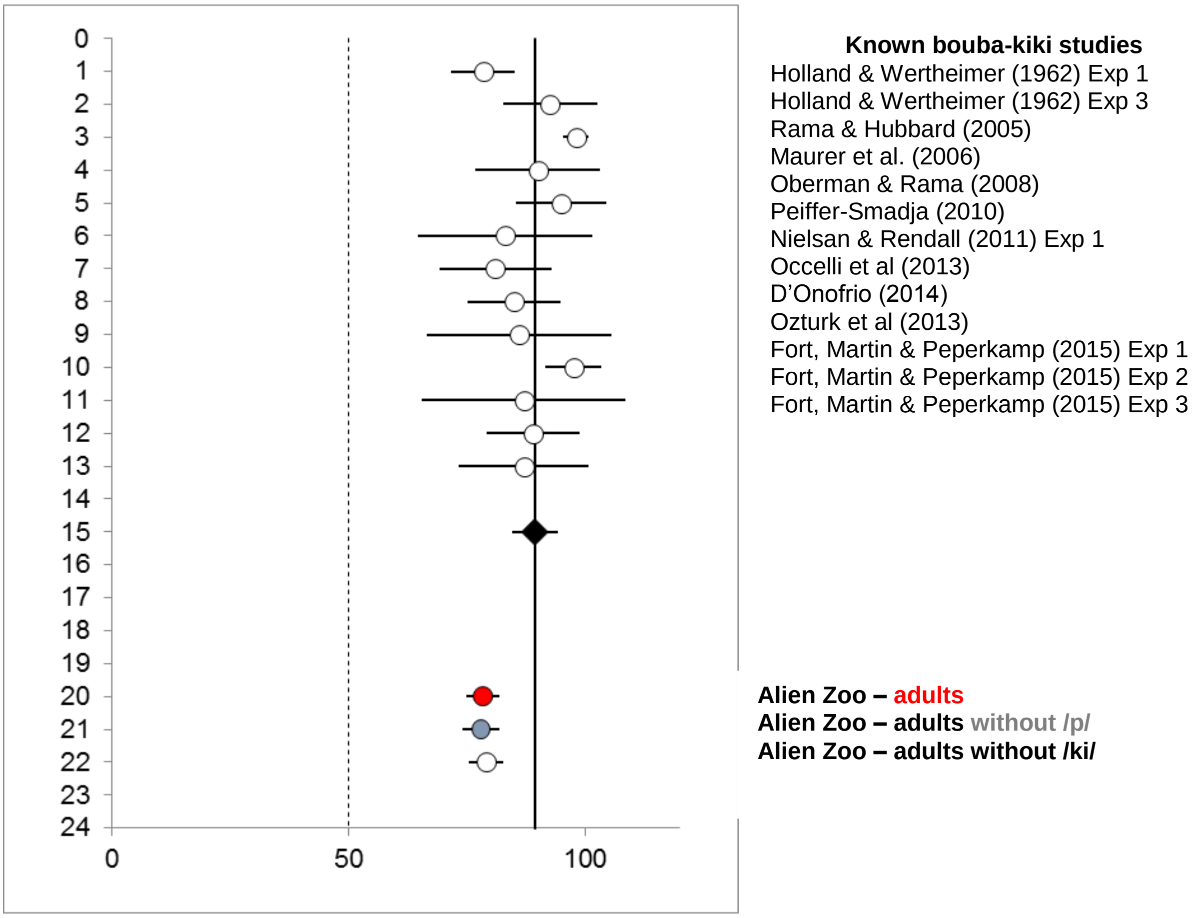


Figure 32. Forest plot of the rate of congruent choices made by adults in a bouba/kiki tasks. 1-13 are published studies containing stimuli using canonical speech sounds. The last three coloured points are from the Alien Zoo task.

As shown, our bouba-kiki task did not fall within the range of eliciting 84%-94% for congruent responses from adults.

While there is no concrete support for us to remove the auditory tokens from our analysis, we are aware that language background differences affect the way participants process these auditory tokens. When auditory tokens resembling human speech do not fall in line with the phonotactics of one's language, participants may process the tokens as non-speech sounds and the linguistic sound symbolic effect will fail to be elicited (Styles & Gawne, 2017b).

Also, our visual stimuli are very different from these other published studies, as they are colourful and have googly eyes on them. It is possible that our visual stimuli are less directional when it comes to eliciting sound-shape matching compared to the flat line drawings which are commonly used.

Study 2: Alien Zoo task in relation to predictors of early reading

A previous study on Dutch adults with dyslexia showed that these adults made fewer congruent sound-symbolic choices in a bouba-kiki task compared to normal adults (Drijvers et al., 2015). The authors suggested that dyslexic individuals may suffer from deficits in abstracting and/or integrating crossmodality (auditory and visual) cues which make them poor readers and causes them to make fewer congruent sound-shape matches. At present, theirs is the only study which has attempted to connect reading deficits with a sound-to-shape matching task. We are curious to know the developmental trajectory of dyslexia and its relation to sound-symbolism and if there is a direction of causality between reading deficits and deficits in crossmodal processing. As the children in our study are part of a longitudinal cohort study, we are able to explore this developmental trajectory by comparing results from our novel adaptation of the bouba-kiki task – the Alien Zoo task – with archival data on predictors of early reading.

In Study 1, we established norms for Singaporean children performing a sound-to-shape matching task. The newly normed data works as a scale for comparison with other developmental scales. As part of the GUSTO cohort study, the same children were tested for several measures on predictors of early reading prior to turning 6 years old, namely: (i) the Singapore adaptation of the Communicative Developmental Inventory (CDI) (Reilly et al., 1993; Tan, 2009) at 2 years; (ii) the Comprehensive Test of Phonological Processing (CTOPP-2) (Wagner, Torgesen, Rashotte, & Pearson, 2013) at 4 years; (iii) the Peabody Picture Vocabulary Test (PPVT-4) (Dunn & Dunn, 2007) at 4 years; (iv) and Lollipop Test (Chew, 1981) at 4 years.

Methods

Prior to linking the results of the Alien Zoo task to results of these measures, details of proposed analysis pathways for the measures and the Alien Zoo task were pre-registered on AsPredicted.org (see Appendix 5 for the preregistration). According to the previous literature suggesting links between dyslexia and sound symbolism (Drijvers et al., 2015), we predict that a possible causal link is multisensory processing deficit emerging before the diagnosis of dyslexia. Thus, we predict positive correlations between predictors of early reading (namely, phonological awareness) and the number of congruent responses made in the Alien Zoo task. Secondary analyses will check for correlations between other measures of early reading ability e.g. vocabulary size on the Alien Zoo task score. If positive correlations are found, we proposed using regression models to examine which measure had the most influence on the Alien Zoo task score.

Descriptions of the methods of the measures of predictors of reading are presented first, followed by results of the correlation analyses.

Phonological Awareness

Three subtests of the CTOPP-2 (R. Wagner et al., 2013) designed to test phonological awareness were administered to the children when they were 48 months of age. (1) Elision: removing speech segments from words, (2) Blending: combining speech segments into words, and (3) Sound Matching: identifying word-initial or word-final phones from a list of 3 words. Procedures and examples from each of the subtests are presented in Table 17. The total score from these three subtests was tabulated to form a composite score for phonological awareness.

	Procedure	Example
Elision	The child listened to a target word and is made to repeat it then asked to say the target word without a designated sound.	Say ‘toothbrush’ without saying ‘tooth’ Say ‘snail’ without saying ‘[n]’
Blending	The child listened to pre-recorded separate speech sounds then asked to blend the sounds together	What word do these sounds make? ‘[s]- [un]’ What word do these sounds make? ‘[m]-[u]-[n]’
Sound Matching	The tester read a word while pointing to its visual referent and then read three other words while pointing to their visual referents. The child is asked to point to the picture which referred to the word with target sound.	Which word starts with the same sound as ‘nap’? ‘Tape’, ‘Net’, or ‘Man’? Which word ends with the same sound as ‘Rob’? ‘Knot’, ‘Rain’, or ‘Tub’?

Table 17. Examples from the CTOPP-2 subtests administered to the children at age 48 months.

Vocabulary Size: CDI and PPVT

As mentioned in the earlier chapter, parents of the children in our study reported their child’s expressive vocabulary through the Singapore version of the Communicative Development Inventory (CDI) at 2 years of age (Tan, 2009). Parents indicated if their child used words listed in the inventory. Each word a child used was given one mark and a sum of all the marks makes up the child’s CDI score.

Another measure of the children’s vocabulary size was taken via the Peabody Picture Vocabulary Test IV (PPVT) at age 4 (Dunn & Dunn, 2007). The PPVT is a measure of English receptive vocabulary, containing 228 test items arranged in 19 sets of increasing difficulty. Each test item is accompanied by 4 visual referents (black line drawings on a white background). The tester would say a target word and ask the child to select the corresponding picture which best matches the target word. The PPVT does not require children to be able to verbalise the target word. The children responded by pointing to the picture which they thought best represented the target word; hence providing an indication of their receptive vocabulary. The total raw score is the number of correct responses and the raw scores were standardised to a mean of 100 with a standard deviation of 15 as recommended in the user manual.

Letter Identification

The Lollipop School Readiness Test (Chew, 1981) was administered when the children were at age 4. The test has been validated by previous research as successful in predicting future school achievements (Chew & Morris,1984; Monette, Bigras, & Guay, 2011). A subscale of the test examined the children's ability to identify letters and their ability to write them. As we are only interested in measures of reading ability, we omitted the last four test items in this subscale which required the children to write out letters and their name. Test items included asking a child to pick out a target letter e.g. "Show me the letter B" and asking a child to read out a target letter e.g. Point to the letter M and ask, "What letter is this?" A total score for the ten letter-identification test items was tabulated for each child.

Results

Alien Zoo and Phonological Awareness

There was an overlap of 304 children in both the CTOPP-2 task (M = 27.0, SD = 3.65) and the Alien Zoo task (M = 9.60, 2.51). A scatterplot of the number of congruent choices made and the children's phonological awareness score is shown in Figure 33.

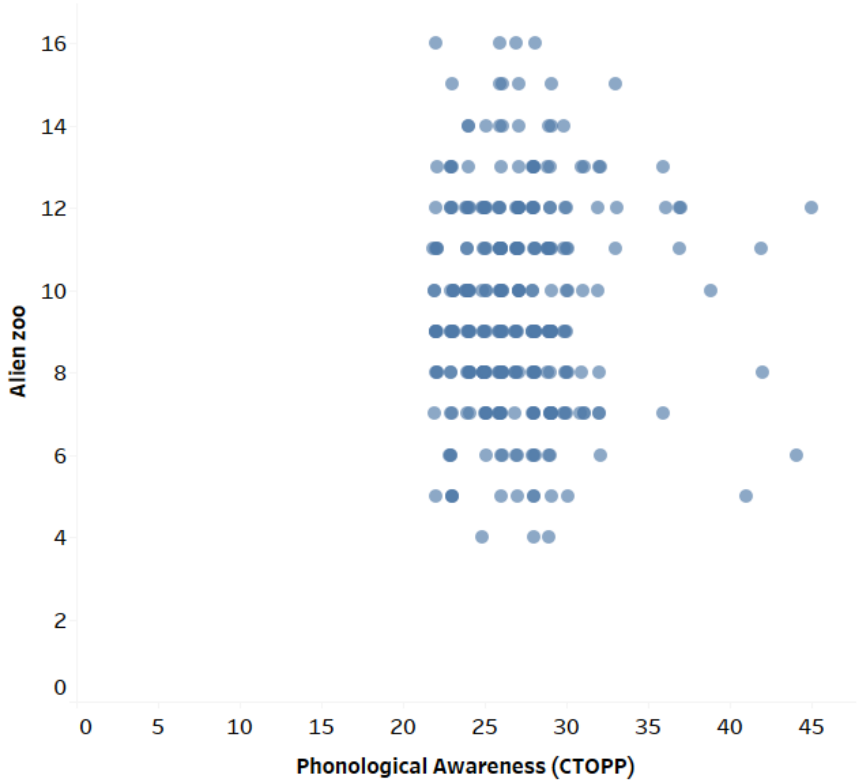


Figure 33. Scatter plot of Alien Zoo score (6y) and Phonological Awareness score (4y). The PA score has been jittered (± 0.25) according to a formula by Ashish (2017).

A Spearman's rank-order correlation was run to determine the relationship between 304 children's phonological awareness score and their performance on the Alien Zoo task. There was no significant correlation between phonological awareness score at age 4 and number of congruent responses selected during the Alien Zoo task at age 6, $\rho(304) = .007$, $p = .907$.

Alien Zoo and Vocabulary Size

187 children whose CDI scores were reported at age 2 (M = 158, SD = 140) participated in the Alien Zoo task at age 6 (M = 9.50, SD = 2.40). The scatterplot showing the number of congruent responses per child and their respective CDI score is shown in Figure 34.

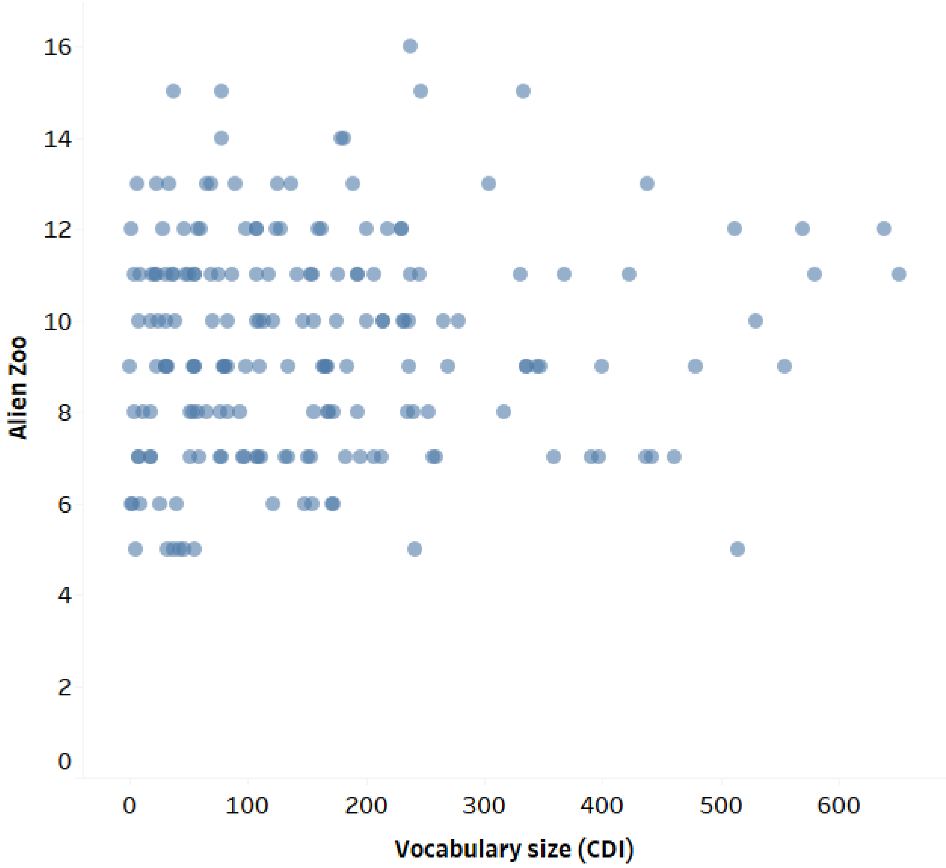


Figure 34. Scatterplot of Alien Zoo score (6y) and CDI (2y)

A Spearman's rank-order correlation was run to determine the relationship between 187 children's CDI score and their performance on the Alien Zoo task. There was no significant correlation between CDI score at age 2 and number of congruent responses selected during the Alien Zoo task at age 6, ρ(187) = .071, *p* = .333.

The other measure of vocabulary size was the PPVT task conducted when the children were 4-years old. 313 children whose PPVT scores were reported at age 4 (*M* = 87.0, *SD* = 16.4) participated in the Alien Zoo task at 6 years (*M* = 9.58, *SD* = 2.50). The scatterplot illustrating the number of congruent responses per child and their respective PPVT scores is shown in Figure 35.

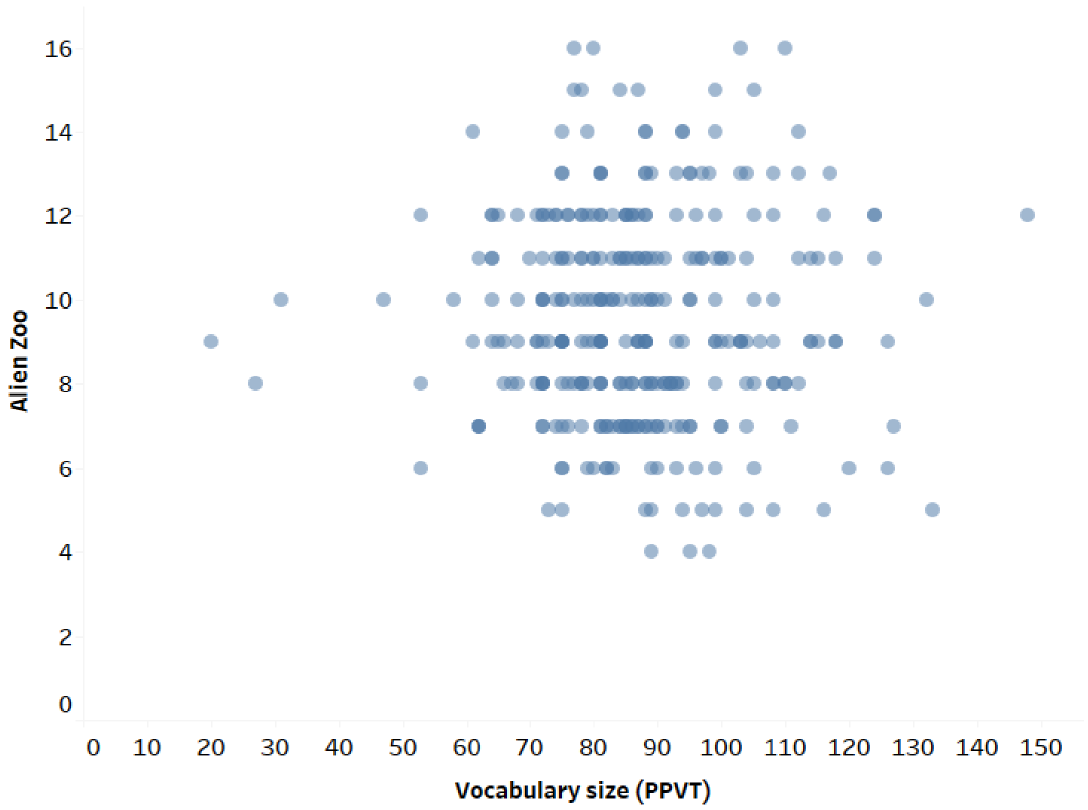


Figure 35. Scatter plot of Alien Zoo score (6y) and PPVT score (4y).

A Spearman's rank-order correlation was run to determine the relationship between children’s PPVT score and their performance on the Alien Zoo task. There was no significant correlation between PPVT score at age 4 and number of congruent responses selected during the Alien Zoo task at age 6, ρ(313) = .690, *p* = .333.

Alien Zoo and Identification of Letters

310 children whose Lollipop test scores (*M* = 6.56, *SD* = 3.71) were reported also participated in the Alien Zoo task (*M* = 9.59, *SD* = 2.51). Figure 36 is the scatterplot of the number of congruent choices made by the children and their corresponding Lollipop subscale test score.

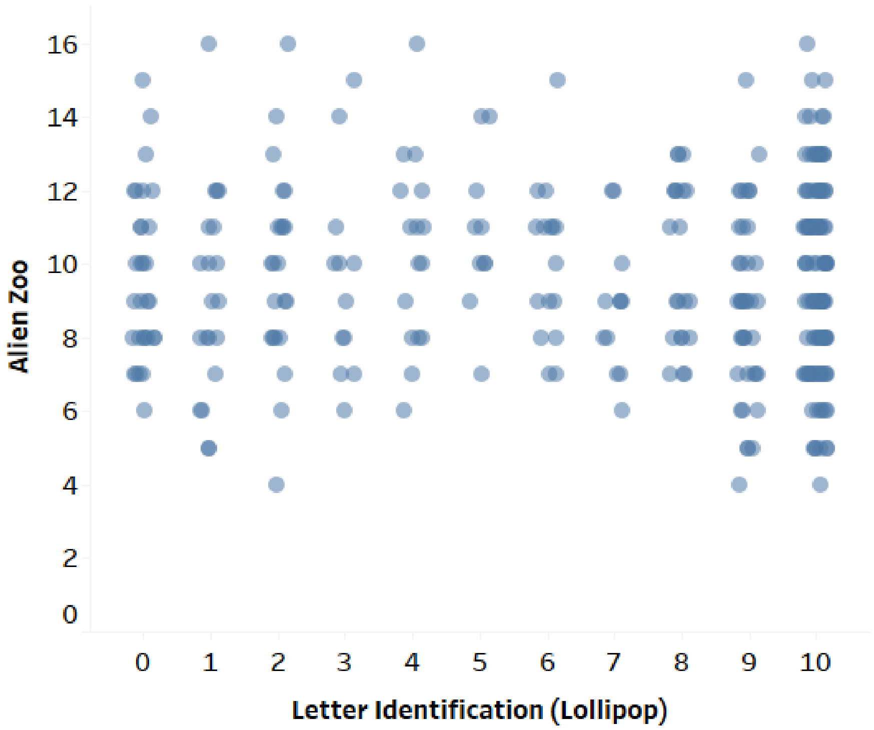


Figure 36. The scatterplot of Alien Zoo score (6y) and Lollipop score (4y). The Lollipop test score has been jittered (± 0.33) according to a formula for jittering on Tableau software (Ashish, 2017).

A Spearman's rank-order correlation was run to determine the relationship between children's Lollipop test score and their performance on the Alien Zoo task. There was no significant correlation between Lollipop score at age 4 and number of congruent responses selected during the Alien Zoo task at age 6, $\rho(310) = -0.16$, $p = .772$.

Although these predictors of early reading did not correlate with the Alien Zoo task, they are strongly correlated with one other. Table 18 presents the details of the Spearman's correlations across the four measures of predictors of early reading.

For the children in our study, a larger expressive vocabulary size at age 2 correlated positively with stronger phonological awareness, larger receptive vocabulary size and better letter identification ability at age 4.

Measures	Phonological Awareness (CTOPP)			Receptive Vocabulary (PPVT)			Letter Identification (Lollipop)		
	N	**ρ**	***p***	**N**	**ρ**	***p***	**N**	**ρ**	***p***
2 years									
Expressive Vocabulary (CDI)	285	.257	**>.0005**	291	.341	**>.0005**	291	.170	**.004**
4 years									
Phonological Awareness (CTOPP)				697	.459	**>.0005**	697	.352	**>.0005**
Receptive Vocabulary (PPVT)							716	.466	**>.0005**

Table 18. Positive correlations (Spearman's rho and significance) found among all the measures for predictors of early reading.

The scatterplots illustrating the relationship between these measures are shown in Figures 37 and 38.

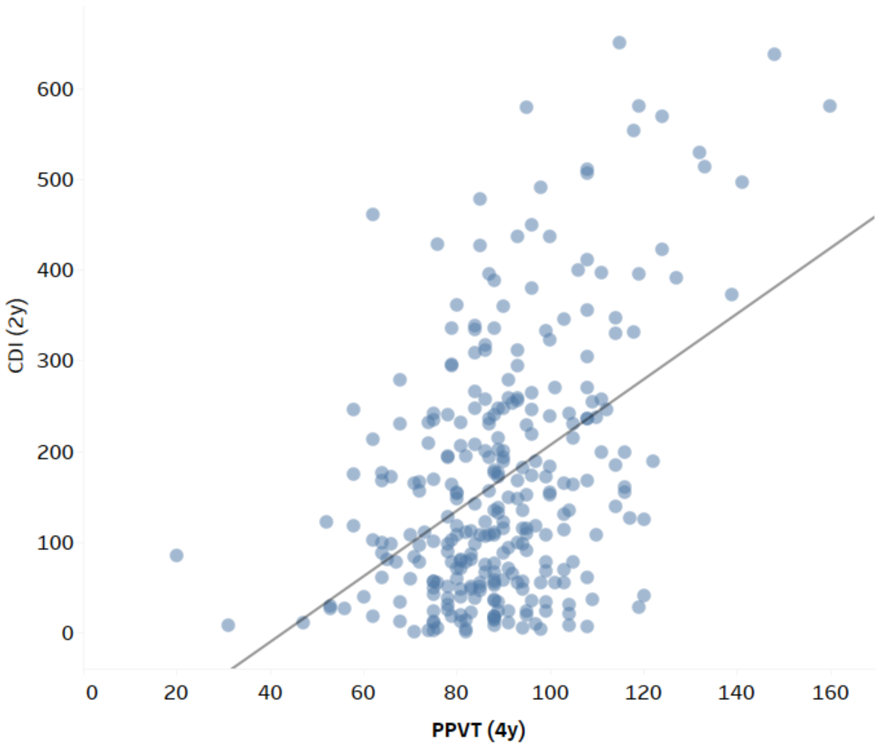


Figure 37. Scatter plot showing the relationship between both vocabulary measures, CDI (2y) and PPVT (4y).

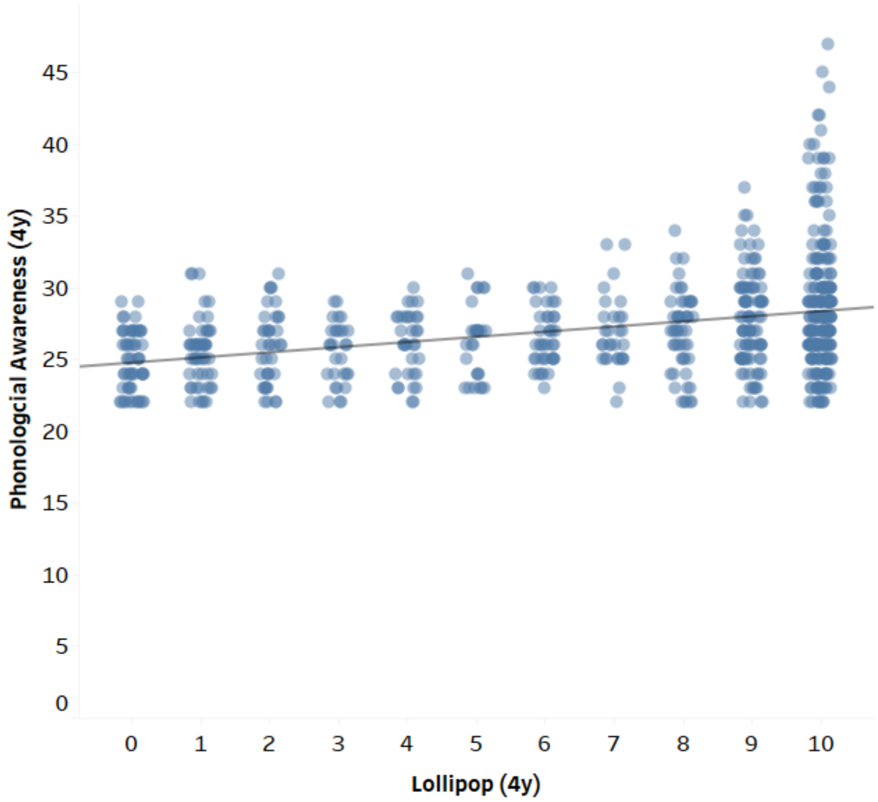


Figure 38. Scatter plot showing the relationship between phonological awareness and letter identification. Both measures are predictors for early reading. The Lollipop test score has been jittered (± 0.33) according to a formula for jittering on Tableau software (Ashish, 2017).

There was multicollinearity among the four predictors of early reading. We pre-registered the use of regression analysis to examine which of the four measures contributed most the variance in the results of the Alien Zoo task. However, as no correlations were found between these measures and the Alien Zoo task, stepwise multiple regression models investigating the independent contributions of these measures on the Alien Zoo task score were not warranted.

Discussion

Beyond establishing norms in a linguistic sound-symbol matching task with Singaporean children, we wanted to know if the number of congruent choices made by the children correlated with measures of known predictors of early reading skills. This was largely motivated by Drijvers et al (2015) who tested dyslexic adults on a two-alternative forced choice bouba-kiki task and found that they made significantly fewer sound-symbol matches (60%) compared to adult controls (73%). The authors suggested that the dyslexic individuals make fewer sound-symbol matches due to deficits in audio-visual abstraction.

It is well-studied that dyslexic individuals have auditory processing deficits (Goswami, Wang, et al., 2011; McBride-Chang et al., 2011; Serniclaes & Sprenger-Charolles, 2003; Zhang et al., 2012). They also exhibit deficits in auditory-visual temporal coordination (Meyler & Breznitz, 2005) and in letter-speech matching and integration (Blau et al., 2009; Froyen, Willems, & Blomert, 2011). However, many audio-visual tasks are reading-related or require reading skills, making it difficult to test pre-reading children. As such, there is no evidence about the developmental trajectory of sound-symbolism and dyslexia.

While we do not have results of current measures for reading ability on the pre-schoolers in our study, we predicted that results from predictors of early reading ability e.g. phonological awareness and vocabulary size would correlate with results from our Alien Zoo task. Our predictions were not borne out though those measures were strongly positively correlated with one another.

Learning to read requires the association of two low-level sensory processing – auditory (speech sounds) and visual (letters/orthographic representations). These measures of predictors of early reading ability may be looking into aspects of reading untapped by or unrelated to our Alien Zoo task. Hence, correlations were not found in our study.

The CTOPP task, examining phonological awareness, investigates the auditory processing of speech sounds. Vocabulary measures, CDI and PPVT, surveys word comprehension and expression (CDI only) abilities of the children, both of which involves higher level processing beyond low-level cross-modality associations. Vocabulary growth is dependent on the development of associations between semantics, phonological, and orthographic representations (Nation, 2009). The subscale of the Lollipop test for letter identification looks only into visual object recognition. It tested the children’s ability to recognise supposedly familiar letters to provide an indication of early literacy training (i.e. children who could not identify letters were most likely ones who have yet to receive any instruction on reading). While these measures have individually been known to be linked with reading comprehension (Clarke, Snowling, Truelove, & Hulme, 2010; Nation, 2009; Shankweiler & Foweler, 2004), the children in the study have yet to begin formal education and have therefore not been evaluated for their reading ability.

On the other hand, with our Alien Zoo task, we observed the audio-visual associations made by children. Neurological processes of this task involve abstracting and integrating auditory and visual cues; processes that are also involved in reading. That is to say, children who made few congruent sound-symbol matches in our study may have deficits in abstracting and/or integrating auditory and visual cues, and later turn out to be poorer readers. To test this concretely, we would require direct reading measures (word, non-word, comprehension) for the children once they begin primary school education and receive formal instruction for reading.

Causal link: Sound-symbolism and reading

We chose to investigate the developmental trajectory of sound-symbolism and dyslexia with a prediction that sound-symbolism deficits may show up before reading deficits since sound-symbol matching involves low-level sensory processing systems which are fully developed before reading occurs, and also evidence suggests that even infants as young as four months prefer congruent sound-shape matches (Pena, Mehler, & Nespor, 2011). However, we do not rule out the possibility that learning to read actually strengthens sound-to-symbol matching thus a better reader also makes more sound-symbolic choices in a bouba-kiki task.

Adults in our study made significantly more congruent choices than the children. Compared to the children, they have received some years of formal reading instruction in English and more than a decade of reading practice, hence, undergoing countless rehearsals for sound-symbol cross-modal sensory associations. Synaptic plasticity allows for strengthening and increasing connections of synapses whose neurons co-activate regularly, and the countless rehearsals of speech sound-letter would require co-activations of neurons processing audio and visual representations, thus potentially altering and strengthening these connections in brains of literate adults (Owens & Tanner, 2017). Hence, the adults may have made more congruent sound-symbolic choices because they are fluent readers. Furthermore, a recent study looking at goodness of fit of shapes to speech sounds shows that English-speaking adults may be influenced by orthography (curves and lines on letters) and authors suggested that continued exposure to the English language and its orthography weighs heavily in influencing literate subjects in a bouba-kiki task (Cuskley et al., 2017). Therefore, it is no surprise that the literate adults in our study make more congruent sound-symbolic choices that pre-reading children.

A way to further study this causal link between reading and sound-symbolism is to retest the same children in our cohort and evaluate their reading skills after they have begun primary school and have received formal instruction on reading. As high-profile as the bouba-kiki effect has been, there has been no validation of test-retest reliability. Given the ages of the children in our study, we expect a possibility of high variability in a behavioural task like ours, thus we expect future retesting (as part of the follow-up) to be highly informative.

Study 3: Linking language backgrounds to the Alien Zoo and predictors of reading

In the previous chapter, we characterised the children in the GUSTO cohort into four main groups as determined by clustering analyses on their language backgrounds. The children are identified as being Mother Tongue-dominant; English-MT balanced; English-dominant; and Mother Tongue Variation-dominant. The main difference between these groups is the amount of English language input the children were receiving at 18 months.

Literature on the sound-symbolic bouba-kiki paradigm suggests that language backgrounds and (English) language knowledge do not drive the selection of congruent choices in non-English speaking populations (Bremner et al., 2013) and children (Maurer et al., 2006) , so long as the sounds are sufficiently 'wordy' for the people taking the test (Styles & Gawne, 2017a). Therefore, we are interested to see if there are group differences in the Alien Zoo scores when the children are grouped according to language backgrounds.

Additionally, as presented in Study 2, the children were tested on predictors of (English) reading with tests on phonological awareness (CTOPP), vocabulary size (CDI and PPVT), and letter identification (Lollipop test). We are interested to know if there are group differences between the various clusters and these predictors of reading ability.

Methods

We ran univariate ANOVA between the three largest clusters (MT-dominant, English-MT balanced and English-dominant) and the Alien Zoo score. The analyses did not include the fourth cluster (MT- Variation dominant) because the number of children in the cluster is too small (N = 7).

Univariate ANOVA between the same clusters and the measures of reading (CDI at age 2, CTOPP at age 4, PPVT at age 4, and Lollipop test at age 4) were run as well.

Results

Cluster differences on Alien Zoo

Children in each cluster did not differ significantly in the Alien Zoo task, F(2,188) = 1.71, p = .184. MT-dominant children (N = 67) made an average of 9.81 congruent choices (*SD* = 2.44), English-MT balanced children (N = 67) made an average of 9.03 congruent choices (*SD* = 2.50), the English-dominant children (N = 57) made an average 9.46 congruent choices (*SD* = 2.33), the MT-dominant children (N = 7) made an average of 9.43 congruent choices (*SD* = 1.81). Figure 39 shows the scatterplot of the children's English input proportion and their Alien Zoo scores with the four clusters shown separately.

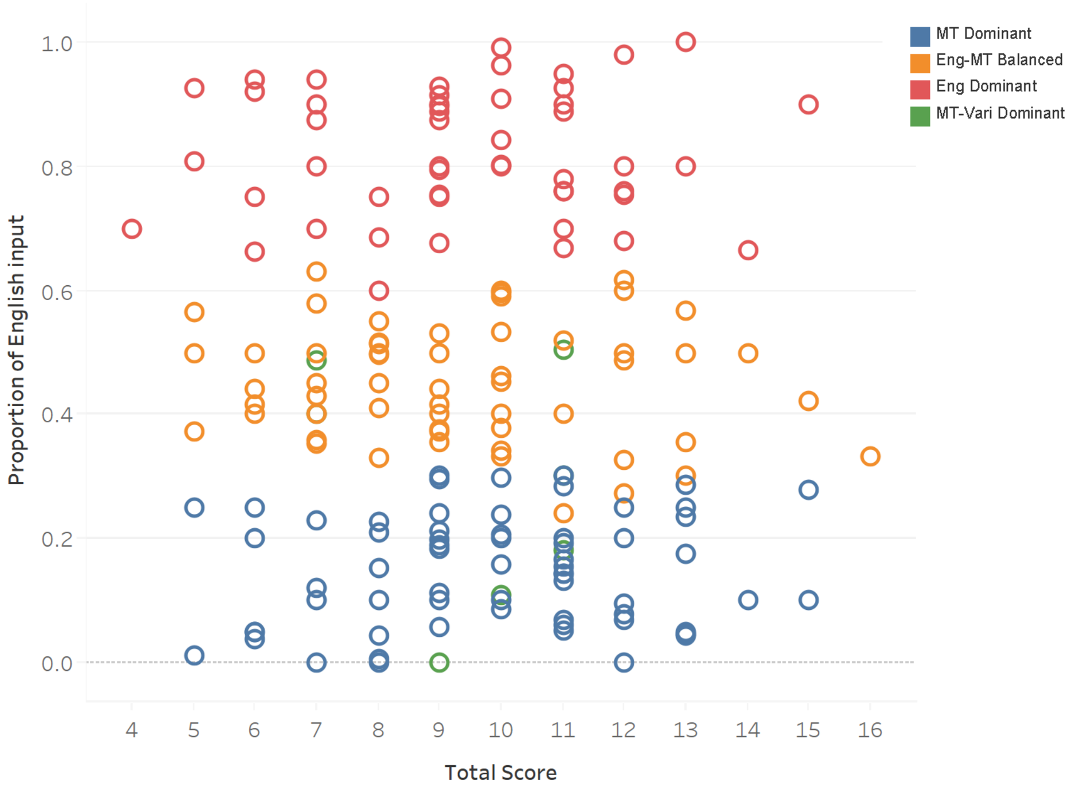


Figure 39. Scatter plot showing the relationship between proportion of English input and Alien zoo score. The children are grouped according to their cluster identity based on language backgrounds.

Cluster differences on predictors of reading

The children in each cluster differed significantly when comparing their scores on vocabulary size measured by Communicative Developmental Inventories (CDI) and Peabody Picture Vocabulary (PPVT) and letter identification ability measured by the Lollipop test. There was no significant difference between the members of different clusters on the measure of phonological awareness, CTOPP.

The CDI is parent-reported scores on productive English vocabulary of the children at 2 years. The children in each cluster differed significantly, $F(2,252) = 18.2$, $p < .005$, MT-dominant: 96.2 (*SD* = 85.3, N = 87), English-MT balanced: 160.7, (*SD* = 131.9, N = 82), English-dominant: 210.3 (*SD* = 148.2, N = 84). Figure 40 (repeat of Figure 26) shows the relationship between English input at 18 months and CDI scores at 2 years.

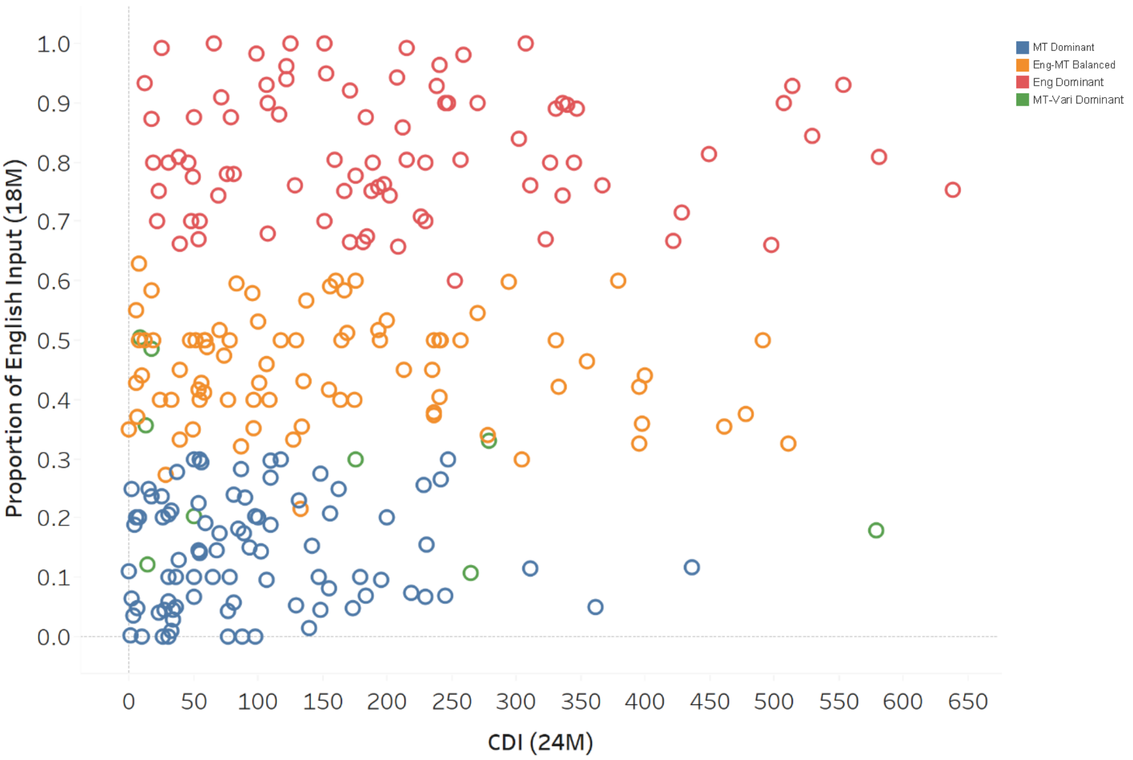


Figure 40 (Repeat of Figure 26). Scatterplot showing the relationship between the proportion of English input (18m) and CDI scores (2y). The children are grouped according to their cluster identity.

The children in different clusters also differed significantly in the other measure of vocabulary size, the Peabody Picture Vocabulary Test, $F(2, 148) = 10.5$, p <.005, MT-dominant: 81.5, (SD = 14.9, N = 14.9), English-MT balanced: 88.61, (SD = 14.1, N = 51), English-dominant: 95.1, (SD = 15.0, N = 38). Figure 41 shows the relationship between the proportion of English input at 18 months and PPVT at 4 years.

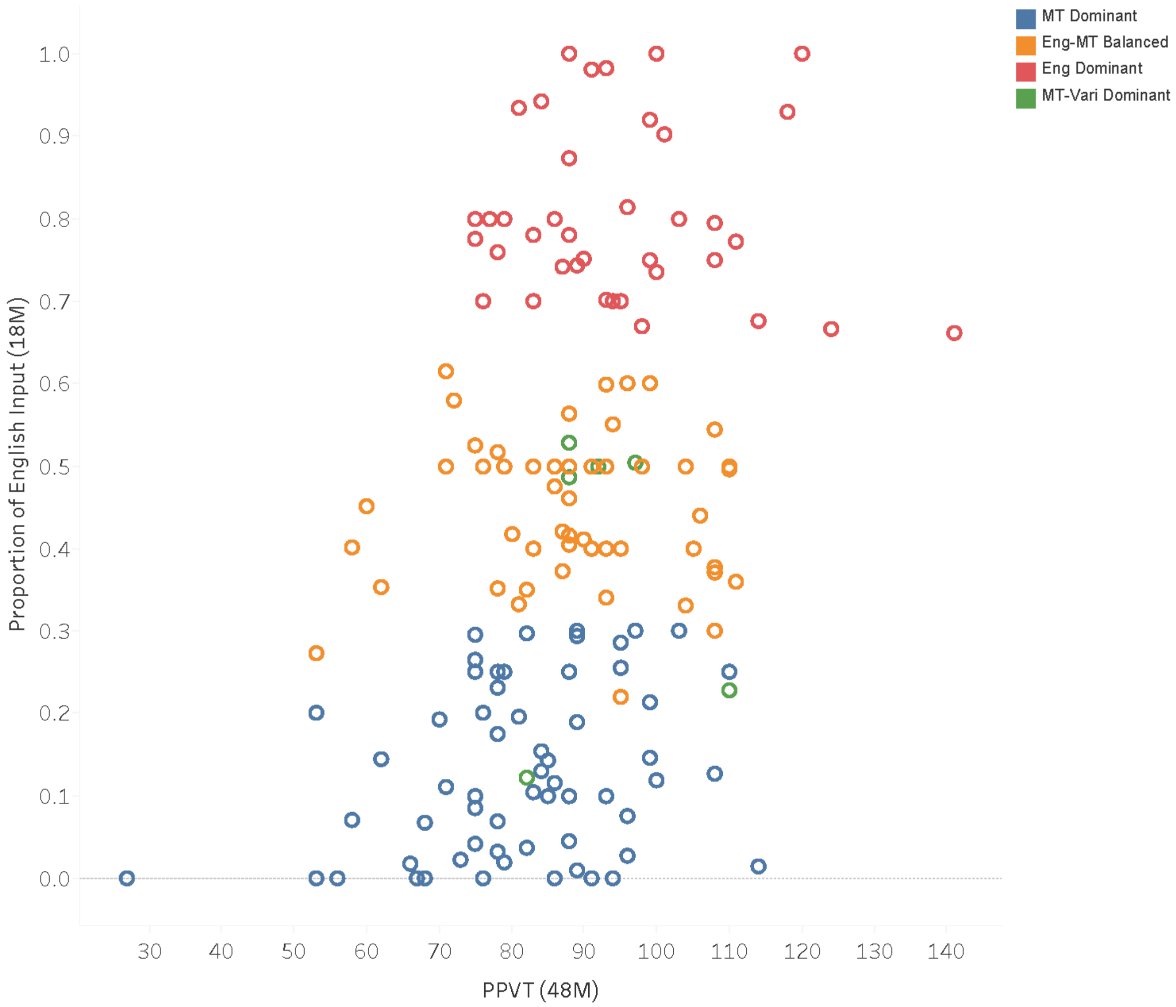


Figure 41. Scatterplot showing the relationship between the proportion of English input (18m) and the PPVT (4y). The children are grouped according to their cluster identity.

The children in different clusters also differed significantly in the Lollipop test, $F(2, 142) = 3.92$, $p = .022$, MT-dominant: 5.74, (SD = 3.60, N = 61), English-MT balanced:

6.38, (*SD* = 3.79, N = 48), English-dominant: 7.78, (*SD* = 2.74, N = 36). Figure 42 shows the proportion of English input at 18 months and the Lollipop test at 4 years.

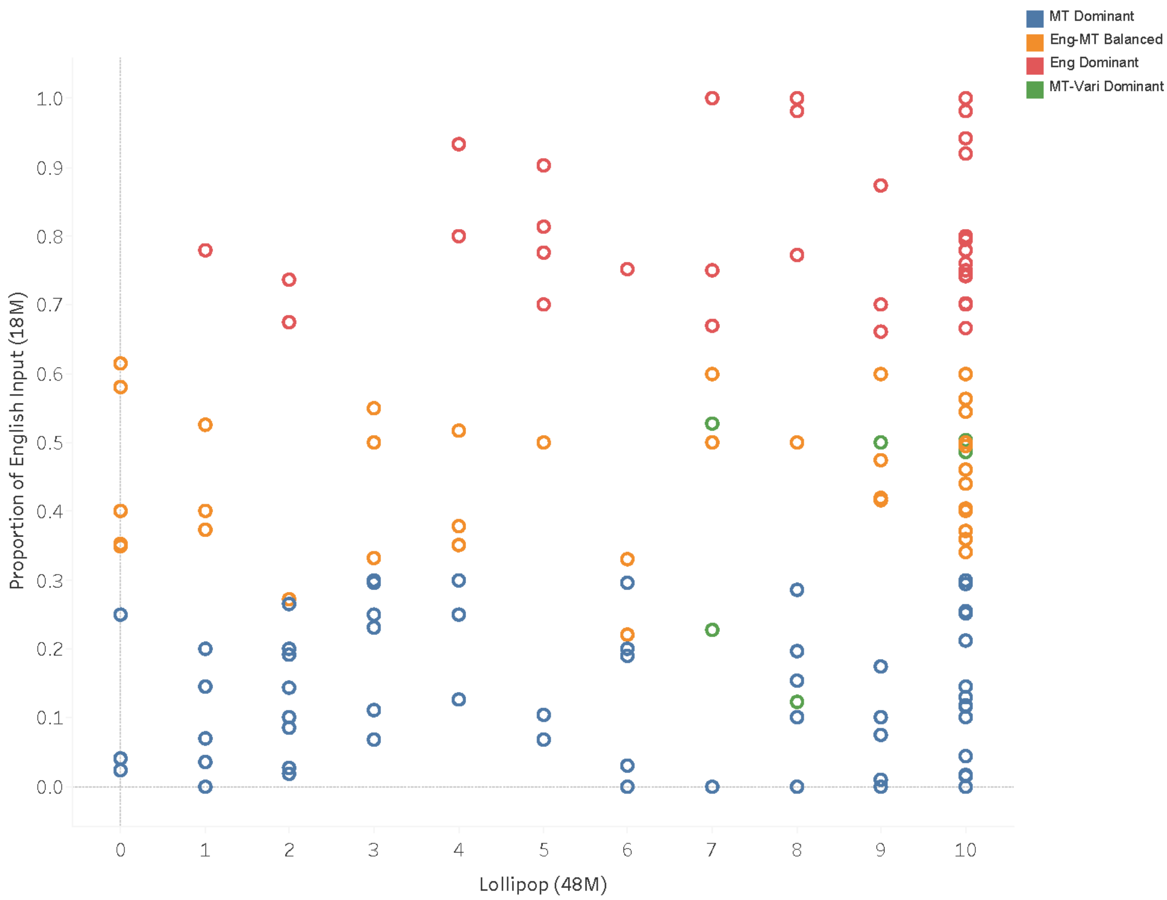


Figure 42. Scatterplot showing the relationship between the proportion of English input (18m) and the Lollipop test (4y). The children are grouped according to their cluster identity.

However, the children did not differ significantly on the measure of phonological awareness, measured by the CTOPP test, $F(2, 140) = .223$, $p = .800$. MT-dominant: 26.8, (*SD* = 3.10, N = 58), English-MT balanced: 27.2, (*SD* = 3.03, N = 49), English-dominant: 27.2, (*SD* = 3.65, N = 36). Taken together, these results show that the clusters (based on the mix of languages in children's input) are reliable predictors of their vocabulary size, but *not* their phonological awareness and *not* their performance on the Alien Zoo Task.

Discussion

Earlier work to establish that the sound-shape matching is systematic and largely universal showed that people groups with little or no exposure to English orthography also made congruent sound-symbolic choices as English speakers (Bremner et al., 2013) and even 2.5-year-olds with little experience with English and a small vocabulary size make identical congruent matches as adults (Maurer et al., 2006). Our Alien Zoo task is shown to be unbiased by language backgrounds of the children in the current sample. As the key differentiator of the different clusters is the amount of English language input, our analysis showed that differing amounts of English language input at a younger age (18m) did not influence the number of congruent sound-symbolic matches made by the children at age 6.

From the current literature, we see that multisensory processing deficits can be shown with the lack of congruent choices made in a bouba-kiki task as seen in adults with dyslexia and individuals with autism (Drijvers et al., 2015; Oberman & Ramachandran, 2008; Occelli et al., 2013). While we did not see a correlation between our Alien Zoo task scores and the children's performance on known predictors of reading in Study 2, we posit that follow-up studies with the children will help us establish a relationship between multisensory processing skills and reading acquisition. Thus, by establishing that our Alien Zoo task is independent of language exposure, we are hopeful that it will be a helpful tool to test preliterate children in future studies.

The cluster analyses presented in Chapter 1 helped to characterise children in our study into four main groups differing largely by the amount of English language input present in their language environments: MT-dominant (<30%); English-MT balanced (<62%); English-dominant (>62%). It is unsurprising that children who received more English language input performed better in measures of English vocabulary (CDI and PPVT) and of English letter identification (Lollipop test).

The children of different clusters did not differ on the measure of phonological awareness because unlike the other three tests which tests on solely on the English language, the phonological awareness measure (CTOPP) tested the children's ability to identify and manipulate speech sounds that not only exist in English but also mostly also exist in other languages spoken by the children like Mandarin, Malay and Tamil. In addition, by age 4, most, if not all of the children, would have some exposure to English so gaining ability to manipulate phonemes need not be linked with varying amounts of English language input at a younger age.

Chapter 4: Overall Discussions & Conclusions

By tracking Singaporean children across four years in a longitudinal cohort study, we present a unique investigation exploring the research gap to link sound-symbolism and reading deficits. This thesis began with descriptive analyses of the language backgrounds of Singaporean children in the largest longitudinal cohort study in the country. By running clustering analyses on parent-reported data of language use, we then reported different models of bilingualism in these pre-verbal children tailored to the Singapore landscape. These analyses characterised the children in our study to allow for analyses of group differences. They also help to characterise Singaporean children in general. The mean language inputs of each of these language background models serve as a better guideline to characterise a bilingual child in Singapore.

The same children performed a novel sound-shape matching task that we designed for children, at around age 6. We reported, for the first time, norms of young pre-reading children on the bouba-kiki paradigm using our novel Alien Zoo task. The results of the sound-shape matching task were then analysed for correlations with known predictors of early reading conducted when the children were at age 4. We also examined possible language background group differences on the Alien Zoo task and the same predictors of reading. We discuss our findings in light of current research gaps in reading development.

Language backgrounds of Singaporean children: What have we learnt?

Presently, in Singapore, there are no definitive criteria for determining if a pre-literate child is bilingual or not nor is there good descriptive data on how language inputs contribute to the kind of bilingualism a child may develop. There is a strongly biased assumption that children in Singapore grow up in a bilingual environment, specifically with English and an ethnically indexed Mother Tongue. However, in reality, children in Singapore may be exposed to languages beyond what is expected according to their official race. Furthermore, the definition of 'Mother Tongue' itself can be problematic as each numerically significant racial group in Singapore (Chinese, Malay, Indian) has only one corresponding Mother Tongue (Mandarin, Bahasa Melayu, Tamil) despite the fact that a myriad of Chinese dialects and Indian languages are spoken here (Silver & Bokhorst-Heng, 2016). Also, no ethnic group can claim English as their Mother Tongue.

To systematically describe the language backgrounds of children in our study, we ran *K*-means clustering analysis on the language input matrixes of the children. The language input matrixes were computed from parent-reported data on language use of all caregivers with the child at 6-months and 18-months in the GUSTO cohort. We recomputed the language input matrixes into Mother Tongue matrixes, grouping Mother Tongues (Mandarin, Malay, Tamil) and Mother Tongue 'variant' languages (Chinese dialects, Bahasa Indonesia, other Indian languages) together. This ensured that the partitions of new clusters were not driven by language labels and allowed for similar languages to be treated equally regardless of the number of children in each language group.

We presented pioneering work regarding classification of language backgrounds of Singaporean children; where clusters are data-driven and not pre-determined according to state-sanctioned race-based language use or by pre-existing expectations of the researchers. Across the 6-month and 18-month time-points, the classification of the children into four clusters accounted best for the within-group language input similarities and maximised between-group language input differences.

The four clusters were best described as follows: MT–dominant; English–MT balanced; English–dominant; or MT Variation–dominant. Based on our data, a typical child growing up in an English-dominant or Mother Tongue-dominant environment receives at around 80% input in that dominant language. A typical child with balanced bilingual input receives approximately 45% English and at least 53% Mother Tongue input. The corresponding percentages of languages in the language mix for each cluster provides better guidelines for determining the cut-offs when classifying a bilingual child in Singapore compared to theorised cut-offs.

By having four clusters which differed in the amount of English input from caregivers, we found strong positive correlations with English input at 6-months and 18-months and the children' English expressive vocabulary size at two years as reported by parents via the Singapore adaptation of Communicative Developmental Inventory (Tan, 2009). We also saw that the children in different clusters differed in their performance on measures of predictors of reading. Children with increased English language input (>62%) at 18 months outperformed their peers in English productive vocabulary (CDI at 2 years and PPVT at 4 years) and in English letter identification (Lollipop test at 4 years).

For both time-points, the MT–dominant group was the largest. While most children remained in the same clusters across both time-points, approximately 40% of the children moved to different clusters at 18-months. Factors like changes in caregivers and changes in care time by each caregiver could have accounted for the movement.

Notably, we see a trend of children moving into clusters that are aligned with language policies in Singapore at 18-months. The number of children in the English –MT balanced and English–dominant groups increased from 6-month to 18-month time-points. This may have been a conscious effort on the part of parents and caregivers to shift towards using more English with their child; driven by attitudes regarding the utility of the English language. The importance of a good command of English is actively promoted by the

Singapore government and is presented as synonymous with success in school and at the workplace (Bolton & Ng, 2014; Pakir, 1993).

As Bolton & Ng (2014) predict, for the next population census in 2020, more Singaporeans across all ethnic groups are expected to report English as their dominant home language. The use of Mandarin, Malay and Tamil as well as dialects will be further reduced. While the MT-dominant group was the largest cluster across both time-points, we predict that most children would use English as their dominant language as they get older. Indeed, the majority of children in the GUSTO cohort chose to complete the English version of our task at age 6 instead of the Mother Tongue versions.

Limitations of Parent-reported data

The data used to understand language backgrounds of children were parent-reported information from the Language Background Questionnaire completed from 2010-2012. We note that the GUSTO team typically administers the test and collects responses from about 800 mother-child dyads yearly, comparatively the response rates for the questionnaires were low (444 at 6-months and 365 at 18-months). The high attrition rate may be because the questionnaires were sent home to be completed and were not closely followed up at the time of data collection.

We recognise the limitations of analysis of archival data including our inability to account for possible transference errors from paper to digital archive and emergence of confounds due to missing data. We accounted for missing data from the questionnaires in a systematic fashion (see Appendix 2 for a summary of data handling) and recognise that for households with missing information on language use (typically of grandparents and domestic helpers) or time spent by caregivers, some minor misrepresentations of languages used with the child may have arisen.

The design of the questionnaire was sub-optimal; it was not immediately clear for parents how they should go about filling in the information of language use and caregiving time in percentages (see Appendix 1 for the original form). The options provided for questions might not have allowed for information on language use to be captured fully as some assumptions about caregivers we made. For example, language use of domestic helpers was restricted to English and one other language.

The parent-reported measure of expressive vocabulary, the Communicative Developmental Inventory (CDI), was used for comparative analyses in our study. Researchers agree that parent-reported measures on their children's language abilities have good validity and reliability and results correlate with laboratory observations of children's language abilities (Law & Roy, 2008; Paradis et al., 2010). The reliability of the reports is also unencumbered by mothers' occupation and level of education. While limitations regarding archival data still apply to the reported CDI scores, we analysed the scores with confidence in their validity.

Moving forward, surveys to elicit language use would be better carried out on a digital platform like Qualtrics (2005) where respondents can be prompted to complete null responses and data verification (like ensuring percentages sum to 100) is also possible. One additional possibility would be to allow options provided for language use to be unbiased and to elicit accurate representations. A respondent-friendly design would increase in response rates. We note that introspective surveys from parents provide only particular information about language use. For example, parents can report accurately about their child's knowledge of words but may find it difficult to describe the frequency of talk time and type. Hart and Risley (1995, 2003) presented massively influential work known as the '30 million word gap'; children whose parents spoke less to them at age 1-2 had smaller vocabularies by age 3 and lower academic achievement at age 9-10. Their findings prompted research into the measurements of talk type and time (Golinkoff, Can, Soderstrom, & Hirsh-Pasek, 2015). Thus, for more informative details regarding word quantity and utterance types, we propose full-day recordings with technology like the LENA system which allows for quantification and classification of utterances by child and caregivers (Cristia, 2013; Greenwood, Thiemeann-Bourque, Walker, Buzhardt, & Gilkerson, 2011).

Sound-symbolism in pre-reading Singaporean children: What have we learnt?

Linguistic iconicity, first explored by Kohler (1929:1947) and Sapir (1929) regarding the size of objects and vowels, provides a radically contrastive view to language arbitrariness (de Saussure, 1959). Arbitrariness – where there is no inherent relation between word and its referent – has been long been recognised as a cornerstone of language. Hockett (1960) listed arbitrariness as one of the key design-features of language; stating that “whale is a small word for a large object”. But as research regarding the symbolic links between language and meaning by examining speech and signed languages emerge, we might argue that ‘whale’ /wail/ is better suited to describe the gargantuan sea mammal than ‘wil’ /wil/.

The bouba-kiki paradigm, coined from Ramachandran and Hubbard’s version of the sound-symbol matching task by Kohler (1929:1947), is a well-studied experimental paradigm showing people’s systematic preference for matching certain sounds with certain shapes; providing evidence against complete arbitrariness in language (Ramachandran & Hubbard, 2001). In particular, speakers match rounded vowels and voiced bilabial consonants e.g. [o, u, b, m] with curvy shapes and high front vowels and voiceless stop consonants and e.g. [i, e, t, k] to spiky shapes.

Maurer et al (2006) were the first to demonstrate that pre-reading children make sound-symbol matches just like adults; ruling out effects of experience with orthography as the cause for the preference. With their 4-items task and small sample size, they found no overall significant difference in the rate of sound-shape matching between the children and adults. Though not significantly different, adults in their study made more sound-symbolic choices for two out of four of their test items.

We adapted the bouba-kiki paradigm into the 16-item Alien Zoo task and conducted the experiment with 377 children (mean age: 5 y 10 m) growing up in Singapore. We created 16 auditory tokens which consisted of canonically 'curvy' or 'spiky' speech sounds e.g. *bobumu* and *tikit.* Our shapes were pairs of distinctively curvy or spiky 'aliens' and the pre-schoolers were asked, "Which alien has this name [auditory token plays]?" Just like the children in the study by Maurer et al, the children guessed significantly above chance; making systematic sound-shape matching choices.

With our large sample size and the normal spread of number the congruent sound-shape made by the children, we established norms for young Singaporean children performing a bouba-kiki task. Firstly, young Singaporean children do make sound-symbolic matches. Secondly, the rate of congruent sound-shape choices made by children (59.2%) was significantly lower compared to the adults (78.3%). The girls in our study made significantly more congruent choices than boys. However, we posit that this effect is not meaningful and likely due to chance, because of the small effect size (.25) and low achieved power (.67).

Limitations of the Alien Zoo task

Styles and Gawne (2017a) performed a meta-analysis on adults' responses to a sound-symbol matching task using the bouba-kiki paradigm. They found that the average rate for choosing congruent sound-shape choices is between 84%-94% if the auditory tokens were made up of canonical speech sounds known to elicit the effect. These speech sounds were used in the original study by Kohler (1929:1947) and later adapted by Ramachandran and Hubbard (2001) and summarised in Table 19.

Shape	**Speech sounds**
	b, m, l, o, u, a
	t, k, i, e,

Table 19. Canonical speech sounds and corresponding congruent shapes for the bouba-kiki paradigm. Images were retrieved from Garcia,(2016).

Deviations from this list of sounds may result in a lower rate of congruent choices. An example of this was present in the study by Drijvers et al. (2015) where the Dutch adults guessed at a rate of 73%; albeit still significantly higher than Dutch adults with dyslexia (60%). The auditory tokens they used were pseudo-words consisting fricatives /z, v/ and sonorants /r, n/ and these speech sounds have not been established as sounds that evoke matches with curvy or pointy shapes. However, when presented in a forced-choice task, normal adults would still make more congruent sound-shape matches compared to dyslexic adults.

The adults in our study made congruent choices at a rate of 78%, also lower than the rates reported by Styles and Gawne (2017a). Our auditory stimuli were made up of the canonical speech sounds as well as the phoneme /p/. To explore possible reasons for this lowered rate, we removed auditory tokens containing the speech sound /p/ (commonly used by not canonical) and in a separate analysis, we removed tokens containing the syllable /ki/ (not well-formed according to phonotactics of Mandarin). However, both analyses did not significantly improve the rate of congruent choices made overall. We, therefore, reasoned that the phonology was probably not implicated in the lower rate of responding.

Our visual stimuli, as mentioned, are very different from these other published studies, as they are colourful and have animated eyes on them. It is possible that our visual stimuli are less directional when it comes to eliciting sound-shape matching compared to the flat line drawings which are commonly used. Also, as pointed out in the meta-analysis by Styles and Gawne (2017b), there may be a publication bias for the bouba-kiki paradigm. The currently published studies have highly successful rates and underreporting of completed studies with low rates of congruent matching may have inflated the rates. We, therefore, conclude that the lower response rates for adults may have been due to a) less canonical visual stimuli or b) bias in the publication of previous literature

Sound-symbolism and reading development: What have we learnt?

Learning to read is a complex process; involving the association of two low-level sensory processing – auditory (speech sounds) and visual (letters). While we may be wired to acquire speech, researchers on language evolution theorise that reading acquisition builds upon a system anchored for speech (Lieberman, 2006; Lieberman & McCarthy, 2014). Therefore, early research has focused on difficulties in speech processing (especially at phoneme level) to have causative effects on reading deficits. However, beyond widely studied phonological deficits of dyslexic individuals, current research into deficits of auditory-visual integration and processing highlights that reading difficulties could be caused by disturbances in the connections of neural networks which subserves cross-modal processing of speech and visual object recognition (Blomert, 2011). Figure 43 illustrates some of the processes and interactions involved in reading and disruptions to any of these processes would result in reading deficits.

A sound-shape matching task, the bouba-kiki paradigm, serves as an easy-to-implement behavioural task to examine cross-modal interaction. From a previous study, we have some indication that adults with dyslexia do not make as many sound-symbolic choices as do the typical population; suggesting deficits in cross-modal processing (Drijvers et al., 2015). At present, we do not understand the developmental trajectory of dyslexia and we do not know if difficulty in making sound-symbolic choices is necessarily a feature of dyslexia, or if learning to read will allow the strong readers to enhance their multisensory processing and thus, outperforming the dyslexic children in making sound-symbolic choices and in reading.

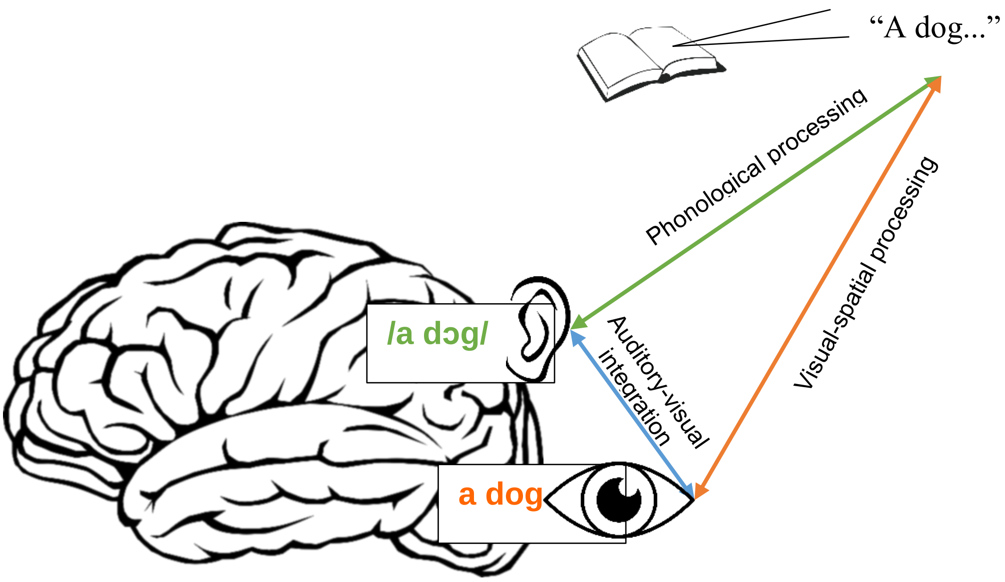


Figure 43. A simplistic model of some processes involved in reading. All the images used in this figure are licensed under Creative Commons (free to use without attribution).

In our attempt to understand the developmental trajectory of dyslexia and sound-symbolism, we examined the connections between the performance of the children on the cross-modal bouba-kiki task and their scores on measures of three predictors of early reading: phonological awareness; vocabulary size; letter identification and naming. The data collected on these measures was completed when the children were at age 2 and 4. We predicted that number of congruent responses made by the children in the Alien Zoo task (adapted from the bouba-kiki paradigm) task prior to turning 6 years old would correlate with these measures individually; that is, a child who has better phonological awareness, or larger vocabulary size, or better ability in letter identification and naming would make more congruent sound-shape choices. However, our predictions were not borne out.

The number of congruent sound-shape choices did not correlate with any of the measures of predictors of early reading even though the measures correlated among one another strongly. When we consider the various processes involved in successful reading (see Figure 43), the lack of correlations may be borne out of different processes being tested in these various measures. The Alien Zoo task is a behavioural task probing into cross-modal processing whereas phonological awareness relies on auditory processing, letter identification and naming rely on visual processing and measure of vocabulary size goes beyond low-level cross-modality associations and is tied to the processing of semantics, memory systems and attention. The children who made few congruent sound-symbol matches in the Alien Zoo task may have difficulties in abstracting and/or integrating auditory and visual information and later, become poor readers. Later stages of the ongoing project will provide valuable information about whether this is the case.

Limitations of comparisons with predictors of reading and what can we do next?

Firstly, we recognise the limitations of working with archival data regarding these reading measures. There could have been transference errors from paper to digital archive which we would not be about to account for. Secondly, the CDI was recorded when the children were at age 2 and the other predictors of early reading were measured when the children were at age 4. We are unable to account for the developmental changes in these children over the years which could have affected their ability to make sound-symbolic choices. Thirdly, there may be an underlying factor that is less obvious or harder to detect e.g.

visual attention that affects cross-modal processing and reading separately and we were not able to examine it in this study.

Finally, we explored the connections between sound-symbol matching task and reading deficits following the results reported by Drijvers et al (2015). As far as we know, no other studies have tested dyslexic individuals with the bouba-kiki paradigm. Replication is necessary to rule out that it was a perhaps a chance finding by those researchers. We will only be able to fully examine the developmental trajectory of cross-modal perception and reading deficits when we retest the children with the bouba-kiki paradigm after they have received formal instruction in reading and collect actual measures of reading.

Conclusions

Growing up in Singapore, children are exposed to various languages and it is difficult to define or quantify bilingualism though being raised bilingual is an expected norm for Singaporean children. For the first time, structured analyses of parent reports of caregivers' language use are conducted to reveal common patterns of caregiver language use.

Between the ages of 6-18 months, children either received dominant English/Mother Tongue/Variation of Mother Tongues input from their caregivers or receive a balanced bilingual (English – MT) input. We provided the means of the variables in the language mix for each of these clusters and that information provides better guidelines for classifying Singaporean children according to the mix of their languages.

At 18-months, more households shifted towards using English mainly with their child or started providing an English-Mother Tongue bilingual language environment for their child. This shift provides an indication of Singaporean parents' attitudes towards language utility and their assumptions of links between language ability and future success.

Consistent findings regarding sound-symbolism across age groups, language backgrounds and cultures, atypical/typical populations have shown that the phenomenon is not a chance finding and contributes to our understanding of language processing and language evolution.

Here, for the first time, we have shown that young children in Singapore make sound-symbolic associations just like adult controls though at a lower rate. Our large-scale study provided norms for the sound-symbol matching task with pre-reading children. The task is not reading-based and can be used as a validated tool for future studies with pre-literate children. We suggest that learning to read may strengthen cross-modal processing and in turn, makes adults better at making sound-symbolic associations. We hope to further this research to shed light on the developmental trajectory of reading deficits and its relation to crossmodal processing.

References

Arriaga, R. I., Fenson, L., Cronan, T., & Pethick, S. J. (1998). Scores on the MacArthur Communicative Development Inventory of children from low and middle-income families. *Applied Psycholinguistics*, *19*(2), 209. https://doi.org/10.1017/S0142716400010043

Ashish. (2017). How to Create Jitter Plot (Strip Plot) in Tableau. Retrieved July 7, 2017, from https://www.doingdata.org/blog/how-to-create-jitter-plot-strip-plot-in-tableau

Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., … Gillam, R. B. (2012). The measure matters: Language dominance profiles across measures in Spanish-English bilingual children. *Bilingualism (Cambridge, England)*, *15*(3), 616–629. https://doi.org/10.1017/S1366728912000090

Blau, V., van Atteveldt, N., Ekkebus, M., Goebel, R., & Blomert, L. (2009). Reduced Neural Integration of Letters and Speech Sounds Links Phonological and Reading Deficits in Adult Dyslexia. *Current Biology*, *19*(6), 503–508. https://doi.org/10.1016/j.cub.2009.01.065

Blomert, L. (2011). The neural signature of orthographic-phonological binding in successful and failing reading development. *NeuroImage*, *57*(3), 695–703. https://doi.org/10.1016/j.neuroimage.2010.11.003

Bolton, K., & Ng, B. C. (2014). The dynamics of multilingualism in contemporary Singapore. *World Englishes*, *33*(3), 307–318. https://doi.org/10.1111/weng.12092

Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., & Spence, C. (2013). " Bouba" and " Kiki" in Namibia? A remote culture makes similar shape-sound matches, but different shape-taste matches to Westerners. *Cognition*, *126*(2). https://doi.org/10.1016/j.cognition.2012.09.007

Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's New, Pussycat? On Talking to Babies and Animals. *Science*, *296*(5572), 1435. https://doi.org/10.1126/science.1069587

Chew, A. (1981). *The Lollipop Test - A Diagnostic Test of School Readiness*. Atlanta: Humanics Pub. Group.

Chew, A., & Morris, J. D. (1984). Validation of the Lollipop Test: A Diagnostic Screening Test of School Readiness. *Educational and Psychological Measurement*, *44*, 987–991.

Chung, K. K. H., McBride-Chang, C., Wong, S. W. L., Cheung, H., Penney, T. B., & Ho, C. S. H. (2008). The role of visual and auditory temporal processing for Chinese children with developmental dyslexia. *Annals of Dyslexia*, *58*(1), 15–35. https://doi.org/10.1007/s11881-008-0015-4

Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating Children's Reading-Comprehension Difficulties. *Psychological Science*, *21*(8), 1106–1116. https://doi.org/10.1177/0956797610375449

Cristia, A. (2013). Input to Language: The Phonetics and Perception of Infant-Directed Speech. *Linguistics and Language Compass*, *7*(3). https://doi.org/10.1111/lnc3.12015

Cunningham, E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934–945. https://doi.org/10.1037/0012-1649.33.6.934

Cuskley, C., Simner, J., & Kirby, S. (2017). Phonological and orthographic influences in the bouba-kiki effect. *Psychological Research*, *81*(1), 119–130. https://doi.org/10.1007/s00426-015-0709-2

D'Onofrio, a. (2014). Phonetic Detail and Dimensionality in Sound-shape Correspondences: Refining the Bouba-Kiki Paradigm. *Language and Speech*, *57*(3), 367–393. https://doi.org/10.1177/0023830913507694

Davis, R. (1961). The fitness of names to drawings. A cross-cultural study in Tanganyika. *British Journal of Psychology*, *52*(3), 259–268. https://doi.org/10.1111/j.2044-8295.1961.tb00788.x

de Saussure, F. (1959). *Course in general linguistics (W. Baskin Trans.).* New York: The Philosophical Library Inc.

Department of Statistics. (2015). *Key Findings: General Household Survey 2015*. Singapore. Retrieved from https://www.singstat.gov.sg/docs/default-source/default-document-library/publications/publications_and_papers/GHS/ghs2015/findings.pdf

Department of Statistics of Singapore. (2010). *Census of Population 2010 Statistical Release 1: Demographic Characteristics, Education, Language and Religion*. Singapore. Retrieved from http://www.singstat.gov.sg/docs/default-source/default-document-library/publications/publications_and_papers/cop2010/census_2010_release1/findings.pdf

Deroy, O., & Spence, C. (2016). Crossmodal Correspondences: Four Challenges. *Multisensory Research*, *29*(1–3), 29–48. https://doi.org/10.1163/22134808-00002488

Dixon, Q. (2005). The Bilingual Education Policy in Singapore: Implications for Second Language Acquisition. In J. Cohen, K. McAlister, K. Rolstad, & J. MacSwan (Eds.), *Proceedings of the 4th International Symposium on Bilingualism*. Somerville: Cascadilla Press. Retrieved from http://www.lingref.com/isb/4/047ISB4.PDF

Drijvers, L., Zaadnoordijk, L., & Dingemanse, M. (2015). Sound-Symbolism is Disrupted in Dyslexia : Implications for the Role of Cross-Modal Abstraction Processes, 602–607.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (Fourth). MN: Pearson Assessments.

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. *International Journal of Language and Communication Disorders*, *49*(3), 317–332. https://doi.org/10.1111/1460-6984.12075

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3:A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Forget-Dubois, N., Lemelin, J.-P., Boivin, M., Dionne, G., Séguin, J. R., Vitaro, F., & Tremblay, R. E. (2007). Predicting Early School Achievement With the EDI: A Longitudinal Population-Based Study. *Early Education & Development*, *18*(3), 405–426. https://doi.org/10.1080/10409280701610796

Fort, M., Martin, A., & Peperkamp, S. (2015). Consonants are More Important than Vowels in the Bouba-kiki Effect. *Language and Speech*, *58*(2). https://doi.org/10.1177/0023830914534951

Fort, M., Weiß, A., Martin, A., & Peperkamp, S. (2013). Looking for the bouba-kiki effect in prelexical infants. In S. Ouni, F. Berthommier, & A. Jesse (Eds.), *International Conference on Auditory-Visual Speech Processing* (pp. 71–76). Annecy.

Froyen, D., Willems, G., & Blomert, L. (2011). Evidence for a specific cross-modal association deficit in dyslexia: an electrophysiological study of letter-speech sound processing. *Developmental Science*, *14*(4), 635–648. https://doi.org/10.1111/j.1467-7687.2010.01007.x

Garcia, X. (2016). Media Guide: The Bouba-Kiki Effect - Science Friday. Retrieved June 5, 2017, from https://www.sciencefriday.com/educational-resources/media-guide-the-bouba-kiki-effect/

Golinkoff, R. M., Can, D. D., Soderstrom, M., & Hirsh-Pasek, K. (2015). (Baby)Talk to Me. *Current Directions in Psychological Science*, *24*(5), 339–344. https://doi.org/10.1177/0963721415595345

Goswami, U., Fosker, T., Huss, M., & Mead, N. (2011). Rise time and formant transition duration in the discrimination of speech sounds : the Ba – Wa distinction in developmental dyslexia and De, *1*, 34–43. https://doi.org/10.1111/j.1467-7687.2010.00955.x

Goswami, U., Wang, H.-L. S., Cruz, A., Fosker, T., Mead, N., & Huss, M. (2011). Language-universal sensory deficits in developmental dyslexia: English, Spanish, and Chinese. *Journal of Cognitive Neuroscience*, *23*(2), 325–337. https://doi.org/10.1162/jocn.2010.21453

Greenwood, C. R., Thiemeann-Bourque, K., Walker, D., Buzhardt, J., & Gilkerson, J. (2011). Assessing children's home language environments using automatic speech recognition technology. *Communication Disorders Quarterly*, *32*, 83–92.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore: MD: Brookes.

Hart, B., & Risley, T. R. (2003). The Early Catastrophe. *American Educator*, *27*(1), 4–9. Retrieved from https://www.aft.org/sites/default/files/periodicals/TheEarlyCatastrophe.pdf

Hockett, C. (1960). The origin of speech. *Scientific American*, *203*, 89–97. https://doi.org/10.1038/scientificamerican0960-88

Hsieh, P.-J., Hung, S.-M., & Styles, S. (2016). Can a word sound sharp before you have seen it? Sound-shape mapping prior to conscious awareness. *Journal of Vision*, *16*(12), 468. https://doi.org/10.1167/16.12.468

IBM Corp. (2016). IBM SPSS Statistics for Windows. Armonk, NY: IBM Corp.

Igualada, A., Bosch, L., & Prieto, P. (2015). Language development at 18 months is related to multimodal communicative strategies at 12 months. *Infant Behavior and Development*, *39*, 42–52. https://doi.org/10.1016/j.infbeh.2015.02.004

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *369*(1651).

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, *109*(1), 54–65. https://doi.org/10.1016/j.cognition.2008.07.015

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, *31*(8), 651–666. https://doi.org/10.1016/j.patrec.2009.09.011

Kalashnikova, M., Goswami, U., & Burnham, D. (2016). Mothers speak differently to infants at-risk for dyslexia. *Developmental Science*, 1–15. https://doi.org/10.1111/desc.12487

Kassambara, A. (2017). Determining The Optimal Number Of Clusters: 3 Must Know Methods - Articles - STHDA. Retrieved February 9, 2018, from http://www.sthda.com/english/articles/29-cluster-validation-essentials/96-determining-the-optimal-number-of-clusters-3-must-know-methods/

Köhler, W. (1929). *Gestalt psychology, an introduction to new concepts in modern psychology* (2nd ed.). NY: Liveright.

Kuhl, P. K. (2011a). Early Language Learning and Literacy: Neuroscience Implications for Education. *Mind, Brain and Education : The Official Journal of the International Mind, Brain, and Education Society*, *5*(3), 128–142. https://doi.org/10.1111/j.1751-228X.2011.01121.x

Kuhl, P. K. (2011b). Early Language Learning and Literacy: Neuroscience Implications for Education. *Mind, Brain and Education : The Official Journal of the International Mind, Brain, and Education Society*, *5*(3), 128–142. https://doi.org/10.1111/j.1751-228X.2011.01121.x

Lai, S. (2015). *Word learning and Crossmodal Perception (unpublished undergraduate dissertation)*. Nanyang Technological University.

Law, J., & Roy, P. (2008). Parental report of infant language skills: A review of the development and application of the communicative development inventories. *Child and Adolescent Mental Health*, *13*(4), 198–206. https://doi.org/10.1111/j.1475-3588.2008.00503.x

Lieberman, P. (2006). *Toward an evolutionary biology of language*. Cambridge, MA: Havard University Press.

Lieberman, P., & McCarthy, R. C. (2014). The Evolution of Speech and Language. In W. Henke & I. Tattlesall (Eds.), *Handbook of Paleoanthropology* (pp. 873–920). Springer Berlin Heidelberg.

Lockwood, G., & Dingemanse, M. (2015). Iconicity in the lab: a review of behavioral, developmental, and neuroimaging research into sound-symbolism. *Frontiers in*

Psychology, *6*, 1246. https://doi.org/10.3389/fpsyg.2015.01246

Lockwood, G., Dingemanse, M., Hagoort, P., Lockwood, G., Dingemanse, M., & Hagoort, P. (2016). Journal of Experimental Psychology : Learning, Memory, and Cognition Sound-Symbolism Boosts Novel Word Learning Sound-Symbolism Boosts Novel Word Learning. https://doi.org/10.1037/a0028642

Low, E. L., & Brown, A. (2005). *English in Singapore: An Introduction*. Singapore: McGraw Hill.

Lyon, G. R., Sl, B. A., Catts, H., Dickman, E., Eden, G., Fletcher, J., … Viall, T. (1995). Defining Dyslexia, Comorbidity, Teachers â€™ Knowledge of Language and Reading A Definition of Dyslexia, *53*.

Mattock, K., & Burnham, D. (2006). Chinese and English Infants' Tone Perception: Evidence for Perceptual Reorganization. *Infancy*, *10*(3), 241–265. https://doi.org/10.1207/s15327078in1003_3

Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: Sound-shape correspondences in toddlers and adults. *Developmental Science*, *9*(3), 316–322. https://doi.org/10.1111/j.1467-7687.2006.00495.x

McBride-Chang, C., Lam, F., Lam, C., Chan, B., Fong, C. Y. C., Wong, T. T. Y., & Wong, S. W. L. (2011). Early predictors of dyslexia in Chinese children: Familial history of dyslexia, language delay, and cognitive profiles. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *52*(2), 204–211. https://doi.org/10.1111/j.1469-7610.2010.02299.x

McMurray, B., Kovack-Lesh, K. A., Goodwin, D., & McEchron, W. (2013). Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence? *Cognition*, *129*(2). https://doi.org/10.1016/j.cognition.2013.07.015

Meyler, A., & Breznitz, Z. (2005). Visual, auditory and cross-modal processing of linguistic and nonlinguistic temporal patterns among adult dyslexic readers. *Dyslexia*, *11*(2), 93–115. https://doi.org/10.1002/dys.294

Nation, K. (2009). Reading comprehension and vocabulary: what's the connection? In R. K. Wagner, C. Schatschneider, & C. Phythian-Sence (Eds.), *Beyond decoding : the behavioral and biological foundations of reading comprehension*. The Guilford Press. Retrieved from https://www.guilford.com/books/Beyond-Decoding/Wagner-Schatschneider-Phythian-Sence/9781606233108/summary

Nielsen, A. K. S., & Rendall, D. (2011). The sound of round: evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology = Revue Canadienne de Psychologie Experimentale*, *65*(2), 115–124. https://doi.org/10.1037/a0022268

Nielsen, A. K. S., & Rendall, D. (2013). Parsing the role of consonants versus vowels in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *67*(2). https://doi.org/10.1037/a0030553

Nygaard, L. C., Herold, D. S., & Namy, L. L. (2009). The semantics of prosody: Acoustic and perceptual evidence of prosodic correlates to word meaning. *Cognitive Science, 33*(1), 127–146. https://doi.org/10.1111/j.1551-6709.2008.01007.x

Oberman, L. M., & Ramachandran, V. S. (2008). Preliminary evidence for deficits in multisensory integration in autism spectrum disorders: the mirror neuron hypothesis. *Social Neuroscience*, *3*(3–4), 348–355. https://doi.org/10.1080/17470910701563681

Occelli, V., Esposito, G., Venuti, P., Arduino, G. M., & Zampini, M. (2013). The takete-maluma phenomenon in autism spectrum disorders. *Perception*, *42*(2), 233–241. https://doi.org/10.1068/p7357

Overview of GUSTO. (2009). Retrieved August 7, 2017, from http://devos.sg/about/GUSTO.html

Owens, M. T., & Tanner, K. D. (2017). Teaching as Brain Changing: Exploring Connections between Neuroscience and Innovative Teaching. *CBE Life Sciences Education*, *16*(2), fe2. https://doi.org/10.1187/cbe.17-01-0005

Ozturk, O., Krehm, M., & Vouloumanos, A. (2012). Sound symbolism in infancy: Evidence for sound-shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology*. https://doi.org/10.1016/j.jecp.2012.05.004

Pakir, A. (1993). Two tongue tied: Bilingualism in Singapore. *Journal of Multilingual and Multicultural Development*, *14*(1&2), 73–90. https://doi.org/10.1080/01434632.1993.9994521

Pakir, A. (1997). Education and invisible language planning: The case of the English language in Singapore. In J. Tan, S. Gopinathan, & W. K. Ho (Eds.), *Education in Singapore: A book of readings* (pp. 57–74). Singapore: Prentice Hall.

Paradis, J., Emmerzael, K., & Sorenson Duncan, T. (2010). Assessment of English language learners: Using parent report on first language development. *Journal of Communication Disorders*, *43*, 474–497. Retrieved from http://www.elsevier.com/copyright

Peiffer-Smadja, N., & Cohen, P. L. (2010). Exploring the bouba / kiki effect : a behavioral and fMRI study. *Small*.

Pena, M., Mehler, J., & Nespor, M. (2011). The role of audiovisual processing in early conceptual development. *Psychological Science, 22*(11), 1419–1421. https://doi.org/10.1177/0956797611421791

Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology*, *1*(DEC), 1–15. https://doi.org/10.3389/fpsyg.2010.00227

Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651). https://doi.org/10.1098/rstb.2013.0300

Phua, K. H., Hui, R., Nodzenski, M., & Bacolod, N. (2012). *Health of Migrants in Singapore*. Singapore. Retrieved from https://www.asef.org/images/docs/Session 3_2_Kai Hong Phua_Preliminary results of studies of Singapore and Hon Kong SAR_1.pdf

Pires, N. (2017). VennDiagrams.net. Retrieved from http://www.venndiagrams.net/

Pua, E. P. K., Lee, M. L. C., Rickard Liow, S. J., W., L. M., M., W., M., O., … J., R. (2017). Screening Bilingual Preschoolers for Language Difficulties: Utility of Teacher and Parent Reports. *Journal of Speech Language and Hearing Research*, *60*(4), 950. https://doi.org/10.1044/2016_JSLHR-L-16-0122

Qualtrics. (2005). Qualtrics. Provo, Utah.

Quinn, J. M., & Wagner, R. K. (2015). Gender Differences in Reading Impairment and in the Identification of Impaired Readers: Results From a Large-Scale Study of At-Risk Readers. *Journal of Learning Disabilities*, *48*(4), 433–445. https://doi.org/10.1177/0022219413508323

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia - A Window Into Perception, Thought and Language. *Journal of Consciousness Studies*, *8*(12), 3–34.

Reilly, J. S., Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., … Pethick, S. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual* (2nd ed.). San Diego: Singular Publishing Group.

Revill, K. P., Namy, L. L., DeFife, L. C., & Nygaard, L. C. (2014). Cross-linguistic sound symbolism and crossmodal correspondence: Evidence from fMRI and DTI. *Brain and Language*, *128*(1), 18–24. https://doi.org/10.1016/j.bandl.2013.11.002

Sapir, E. (1929). The study in phonetic symbolism. *Journal of Experimental Psychology*, *12*(3), 225. https://doi.org/10.1515/9783110198867

Serniclaes, W., & Sprenger-Charolles, L. (2003). Categorical perception of speech sounds and dyslexia. *Http://cpl.revues.org*, (10, NaN, 2003).

Shang, N. (2017). *Crossmodal correspondences between linguistic tones and complex 3-D shapes and their low-level sensory properties (pitch and spatial frequency)*. Nanyang Technological University.

Shankweiler, D. ., & Foweler, A. E. (2004). Questions people ask about the role of phonological process in learning to read. *Reading and Writing: An Interdisciplinary Journal*, *17*, 483–515.

Silver, R., & Bokhorst-Heng, W. D. (Eds.). (2016). *Quadrilingual education in Singapore : pedagogical innovation in language education*. Singapore: Springer.

Sim, C. (2016). Bilingual policy. Retrieved May 20, 2017, from http://eresources.nlb.gov.sg/infopedia/articles/SIP_2016-09-01_093402.html

Singh, L. (2017). He said, she said: effects of bilingualism on cross-talker word recognition in infancy. *Journal of Child Language*, 1–13. https://doi.org/10.1017/S0305000917000186

Siok, W. T., Spinks, J. A., Jin, Z., & Tan, L. H. (2009). Developmental dyslexia is characterized by the co-existence of visuospatial and phonological disorders in Chinese children. *Current Biology*, *19*(19), 890–892. https://doi.org/10.1016/j.cub.2009.08.014

Soh, S.-E., Tint, M. T., Gluckman, P. D., Godfrey, K. M., Rifkin-Graboi, A., Chan, Y. H., …

Venkatesh, S. K. (2014). Cohort Profile: Growing Up in Singapore Towards healthy Outcomes (GUSTO) birth cohort study. *International Journal of Epidemiology*, *43*(5), 1401–1409. https://doi.org/10.1093/ije/dyt125

Spector, F., & Maurer, D. (2009). Synesthesia: a new approach to understanding the development of perception. *Developmental Psychology*, *45*(1), 175–189file:///Users/Abdul/Downloads/art%253A10.10. https://doi.org/10.1037/a0014171

Spector, F., & Maurer, D. (2013). Early sound symbolism for vowel sounds. *I-Perception*, *4*(4). https://doi.org/10.1068/i0535

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*(4), 971–995. https://doi.org/10.3758/s13414-010-0073-7

Spence, C., & Parise, C. (2010). Prior-entry: A review. *Consciousness and Cognition*, *19*(1), 364–379. https://doi.org/10.1016/j.concog.2009.12.001

Styles, S. J., & Gawne, L. (2017a). Systematic Review & Meta-Analysis of canonical maluma/takete effects: Supplementary Material for "Two key failures and a meta-analysis suggest that phonology and phonotactics matter in maluma/takete tasks" (in press). *I-Perception*.

Styles, S. J., & Gawne, L. (2017b). When does maluma/takete fail? Two key failures and a meta-analysis suggest that phonology and phonotactics matter. *I-Perception*.

Styles, S. J., & Lai, S. Z. (2017). *Does it look like Ebola or Zika? Sound-symbolism enhances virus learning*. Singapore.

Styles, S. J., & Plunkett, K. (2009). What is "word understanding" for the parent of a one-year-old? Matching the difficulty of a lexical comprehension task to parental CDI report. *Journal of Child Language*, *36*(4), 895–908. https://doi.org/10.1017/S0305000908009264

Tableau Software Inc. (2017). Tableau Software: Business Intelligence and Analytics. Seattle, WA: Tableau Software Inc.

Tan, S. H. (2009). *Singapore Communicative Developmental Inventories*. Singapore.

Wagner, R., Torgesen, J., Rashotte, C., & Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing, Second Edition (CTOPP-2)* (2nd ed.). Austin, TX: PRO-ED.

Wang, J. J., Bi, H. Y., Gao, L. Q., & Wydell, T. N. (2010). The visual magnocellular pathway in Chinese-speaking children with developmental dyslexia. *Neuropsychologia*, *48*(12), 3627–3633. https://doi.org/10.1016/j.neuropsychologia.2010.08.015

Wewalaarachchi, T. D., Wong, L. H., & Singh, L. (2017). Vowels, consonants, and lexical tones: Sensitivity to phonological variation in monolingual Mandarin and bilingual English–Mandarin toddlers. *Journal of Experimental Child Psychology*, *159*, 16–33. https://doi.org/10.1016/j.jecp.2017.01.009

Xu Rattanasone, N., Burnham, D., & Reilly, R. G. (2013). Tone and vowel enhancement in Cantonese infant-directed speech at 3, 6, 9, and 12 months of age. *Journal of Phonetics*, *41*(5), 332–343. https://doi.org/10.1016/j.wocn.2013.06.001

Zhang, Y., Zhang, L., Shu, H., Xi, J., Wu, H., Zhang, Y., & Li, P. (2012). Universality of categorical perception deficit in developmental dyslexia: An investigation of Mandarin Chinese tones. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *53*(8), 874–882. https://doi.org/10.1111/j.1469-7610.2012.02528.x

Ziegler, J. C., & Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

Appendix 1: GUSTO Language Background Questionnaire (Original)

Language Background Questionnaire

Please fill in the blanks or tick the appropriate answer. All information is kept confidential. Thank you.

Subject id ____01-14054-02_____

Child's date of birth _______________ Child's gender ☐ Male ☐ Female

Your relationship to the child ☐ Mother ☐ Father ☐ Other ______

Language background

1. When **you** *talk to your child*,

a. how often do you use English? ____ % of the time
Mandarin? ____ % of the time
Dialect? ____ % of the time
Malay? ____ % of the time
Tamil? ____ % of the time
Other languages e.g., Peranakan? _____________ about ____ % of the time

b. If you use dialect, which dialect do you use? ____________________

2. When **your husband/wife** *talks to your child*, how often does he/she use
English? ____ % of the time
Mandarin? ____ % of the time
Dialect? ____ % of the time
Malay? ____ % of the time
Tamil? ____ % of the time
Other languages e.g., Peranakan? _____________ about ____ % of the time

3. How often do **your parents** see your child?

a. at least once a day
b. at least once a week
c. at least once a month
d. less often than that

4. How often do **your husband's/wife's parents** see your child?

a. at least once a day
b. at least once a week
c. at least once a month
d. less often than that

5. When **your parents** talk to your child,

does your mother use English? ____ % of the time
Mandarin? ____ % of the time
Dialect? ____ % of the time
Malay? ____ % of the time
Tamil? ____ % of the time
Other languages? _____________ about ____ % of the time

does your father use English? ____ % of the time

Mandarin? _____ % of the time

Dialect? _____ % of the time

Malay? _____ % of the time

Tamil? _____ % of the time

Other languages? ________________ about _____ % of the time

6. When **your husband/wife's parents** talk to your child,

does your husband/wife's mother use English? _____ % of the time

Mandarin? _____ % of the time

Dialect? _____ % of the time

Malay? _____ % of the time

Tamil? _____ % of the time

Other languages? ________________ about _____ % of the time

does your husband/wife's father use English? _____ % of the time

Mandarin? _____ % of the time

Dialect? _____ % of the time

Malay? _____ % of the time

Tamil? _____ % of the time

Other languages? ________________ about _____ % of the time

7. If you have a **maid**, when she *talks to your child*, does she use

English? _____ % of the time

Another language? _____________________ about _____ % of the time

8. If someone else also looks after your child, does this person use

English? _____ % of the time

Mandarin? _____ % of the time

Dialect? _____ % of the time (Which dialect? _________________)

Malay? _____ % of the time

Tamil? _____ % of the time

Other languages? __________________________ about _____ % of the time

9. At home, when you and your husband/wife talk to each other, do you use

English? _____ % of the time

Mandarin? _____ % of the time

Dialect? _____ % of the time (Which dialect? _________________)

Malay? _____ % of the time

Tamil? _____ % of the time

Other languages? __________________________ about _____ % of the time

Caregiving arrangements

15. How much time (on average) *per week* do _____ spend looking after your child?

Caregiver		hours *per week*	% of the week
	you	_____	_____

your husband/wife	_____	_____
your mother	_____	_____
your father	_____	_____
your husband/wife's mother	_____	_____
your husband/wife's father	_____	_____
maid	_____	_____
childcare centre	_____	_____
Other ________________	_____	_____
		Total 100

16. Do you have other children? How old are they? ______ years , _____ years , _____ years, _____ mths

17. If you did not grow up in Singapore, what variety of languages do you speak:

- ☐ Malaysian English ☐ PRC Mandarin ☐ Bahasa Malaysia.
- ☐ UK English ☐ Taiwanese Mandarin ☐ Bahasa Indonesia
- ☐ US English ☐ Malaysian Mandarin ☐ Other_____
- ☐ Australian English ☐ Other_____
- ☐ Other_____

18. If your husband/wife did not grow up in Singapore, what variety of languages does he/she speak:

- ☐ Malaysian English ☐ PRC Mandarin ☐ Bahasa Malaysia.
- ☐ UK English ☐ Taiwanese Mandarin ☐ Bahasa Indonesia
- ☐ US English ☐ Malaysian Mandarin ☐ Other_____
- ☐ Australian English ☐ Other_____
- ☐ Other_____

Appendix 2: Summary of Data restructuring for Language Background Questionnaire

For numerical calculations, all errors from “not_applicable” or “not_answered” were replaced with “0”

Language spoken to child per parent/caregiver (lang use)

1. Sum total percentage of languages spoken
2. Weighted language/per parent or caregiver = $\frac{Per\ lang\ \%}{Total\ lang\ \%}$
3. Sum of all languages for $\frac{Per\ lang\ \%}{Total\ lang\ \%}$ = 1 (check)
4. Dominant language: $\frac{proportion\ of\ each\ language}{sum\ of\ proportion\ all\ other\ languages}$

Steps 1-4 were completed for each parent/caregiver:

a) Parent 1
b) Parent 2
c) Grandparent 1 (related to P1)
d) Grandparent 2 (related to P1)
e) Grandparent 3 (related to P2)
f) Grandparent 4 (related to P2)
g) Maid
h) Someone else

Family types: created 16 family types based on presence of data for language use; if data is a caregiver not provided, we assume that particular caregiver is not involved (NB: Grandparents’ frequency of interaction considered instead of actual time spent data)

Every household is assigned to one family type.

Time spent looking after the child per parent or caregiver (care time)

1. Sum total % of time spent across all
2. Sum total hours of time spent across all
3. Weighted time (proportion) per parent or caregiver = $\frac{per\ caregiver\ \%}{Total\ time\ \%}$ (if % not provided, consider hours instead)
4. Sum of time for all caregivers = 1 (check)

Family types: assigned each household to each of the 16 family types based on time spent (calculated via the actual time spent data)

Calculated the ***mean*** time spent by each caregiver for each family type.

IFF information on time is missing, mean time for based on family type from language use is used instead.

Lang*Time per parent or caregiver (Talk time)

1. Language spoken x time spent was calculated for each parent or caregiver

Language Input Matrix per child

1. Sum of Lang*Time for each language by each parent or caregiver= language input for that particular language
2. Sum of **all language(s) input = 1 (final check)**
3. If Step 2 ≠ 1, check the care time and language input data (sometimes one is missing for a particular caregiver)
4. Care time recalculated; removing the time spent by caregiver with missing data
5. Final Care time tabulated
6. Repeat steps 1 & 2

Final: Each child has a language input matrix with a maximum of 6 languages.

Appendix 3: 5-cluster and 7, 8, 9-cluster on Mother Tongue Matrix at 6-months

Cluster Centres (5-cluster MT matrix 6-months)

	Mother Tongue	**English & MT**	**English**	**MT Vari**	**MT Vari & MT**
English	.13	.46	.80	.19	.15
Mother Tongue	.85	.51	.14	.01	.38
MT Variation	.02	.02	.04	.80	.47
Other MT	-	-	.01	-	-
Other MT Variation	-	-	.01	-	-
Others	-	.01	-	-	-
N	177	121	109	19	18
Range of distances	.007 – .261	.028 – .388	.027 – .555	.070 – .359	.019 – .476

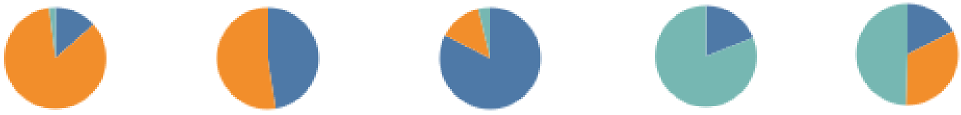


Figure 1. Means of language input for each cluster when the number of clusters is set at 5 (MT matrix at 6-months) with the centroid exemplars for each cluster. The clusters are arranged according to size.

Cluster Centres

(7-cluster MT matrix at 6-months)

	Mother Tongue	English	English & MT	MT Vari & MT	English & MT Vari	MT Vari	English & Oth Vari
English	.11	.81	.43	.13	.48	.10	.48
Mother Tongue	.87	.15	.54	.40	.08	.02	.06
MT Variation	.02	.03	.02	.48	.43	.88	.01
Other MT	-	.01	-	-	.01	-	-
Other MT Vari	-	-	-	-	-	-	.45
Others	-	-	.01	-	-	-	-
N	158	136	105	17	14	13	1
Range of distances	.020 – .271	.026 – .398	.025 – .326	.047 – .299	.098 – .299	.030 – .268	.000

Figure 2. Means of language input for each cluster when the number of clusters is set at 7 (MT matrix at 6-months) with the centroid exemplars for each cluster.

Cluster Centres

(8-cluster MT matrix at 6-months)

	Mother Tongue (A1)	**Mother Tongue (A2)**	**English & MT**	**English**	**MT Vari & MT**	**English & MT Vari**	**MT Vari**	**English & Oth Vari**
English	**.06**	**.28**	.53	.84	.13	.48	.10	.48
MT	.92	.70	.45	.12	.40	.06	.02	.06
MT Vari	.02	.01	.02	.03	.48	.45	.88	.01
Other MT	-	-	-	.01	-	.01	-	-
Oth MT Vari	-	-	-	-	-	-	-	.45
Others	-	-	.01	-	-	-	-	-
N	108	103	101	86	17	13	13	1
Range of distances	.014 – .259	.021– .278	.016 – .389	.020 – .336	.047 – .299	.079 -.252	.030 –.268	.000

Figure 3. Means of language input for each cluster when the number of clusters is set at 8 (MT matrix at 6-months). The MT dominant group splits into two clusters with one cluster (A1) receiving much less English input than the other (A2).

	Cluster Centres (9-cluster MT matrix at 6-months)							
	Mother Tongue	**Mother Tongue**	**English & MT**	**English**	**MT Vari & MT**	**English & MT Vari**	**MT Vari**	**English & Oth Vari**
English	**.06**	**.28**	.53	.84	.13	.48	.10	.48
Mother Tongue	.92	.70	.45	.12	.40	.06	.02	.06
MT Variation	.02	.01	.02	.03	.48	.45	.88	.01
Other MT	-	-	-	.01	-	.01	-	-
Other MT Variation	-	-	-	-	-	-	-	.45
Others	-	-	.01	-	-	-	-	-
N	108	103	101	86	17	13	13	1
Range of distances	.014 – .259	.021 – .278	.016 – .389	.020 – .336	.047 – .299	.079 – .252	.030 – .268	.000

Figure 4. Means of language input for each cluster when number of clusters is set at 9 (MT matrix at 6-months).

Appendix 4: Protocol for Alien Zoo task

BIG Mixup at Alien Zoo
Protocol for BLIP Lab/GUSTO study

Design by Suzy Styles, Shang Nan & Woon Fei Ting

1. Before the start of the test, please make sure that each participant is coded. All identifying information should be removed.
2. CHECK for the language that he/she is more comfortable with and more responsive in. e.g. for a Chinese child, you can address them in both English and Mandarin and ask them which language they want to do the test in. Check for their responsiveness in each language.
    a. Our test has 3 versions: English, Mandarin, and Malay. Please allow the participant to be tested in their MOST proficient language. If the test is not in child's dominant language e.g. Tamil, a researcher proficient in the child's dominant language may translate the English test verbally with appropriate child-directed language.
3. Start ONLY when all the pictures and the audio bars appear for each question.
4. READ out all the instructions and storyline on the screen for the participant.
5. Make sure you CLICK to play the audio for each trial AFTER reading the question.
6. NEVER point to the pictures one at a time while asking (e.g., do you think it is this one?)
7. ASK the child to POINT to the screen to make their choice.
8. CLICK responses for the participant (selected response turns red).
9. If the child is having difficulty making their decision, you can
    a. repeat the question
    b. rephrase the question ("e.g., one of these aliens makes this noise - which one?")
    c. play the sound file again
    d. remind children of the 'story' and how they their guess will help the Zookeeper
    e. suggest the children look carefully at the aliens, to decide which one 'looks like' it would have that name or make that sound
10. AFTER the first question in one Zookeeper's section of the study, it is OK to...
    a. substitute/simplify the question text (e.g., What about these aliens?)
    b. omit reading aloud the question text if the child understands and is keen to answer quickly
11. NEVER imitate the sound from the audio files, always play it again if participant requires repetition

12. NEVER give evaluations if a child articulates an answering strategy
    a. e.g. They may say “all the sharp ones are kiki”. Try to stick to responses like “Really?” “Do you think so?” or “Let’s see…”
13. If the child is worried that their guess might not be correct, you can...
    a. reassure them that the zookeeper will check all the aliens are in the right place at the end
    b. remind them that they are helping the Zookeepers do the sorting
    c. "Let's wait and see" (there is positive feedback at the end no matter what choices the children make)
14. If the child makes a selection before hearing the sound file, please play the sound file and ask the child to verify his/her answer.
15. For the survey questions about languages they speak, the children/research staff can select more than one. If the child is not cooperative in providing answers, the research staff can do the selection based on language background of the child.
16. Please make sure you click >> button until the page which states the results are recorded is loaded (after the THANK YOU page).

BLIP Lab Contact person: WOON Fei Ting: woon0039@e.ntu.edu.sg

Appendix 5: Pre-registered report

AsPredicted
Pre-Registration made easy

Links between Developmental Trajectories of Sound Symbolism and Dyslexia (#4599)

Author(s)
Suzy J Styles (Nanyang Technological University (Singapore)) - suzy.styles@ntu.edu.sg
Fei Ting Woon (Nanyang Technological University (Singapore)) - woon0039@e.ntu.edu.sg

Created: 06/22/2017 01:37 AM (PT)
Public: 08/23/2017 10:50 PM (PT)

1) What's the main question being asked or hypothesis being tested in this study?

Previous research has shown that adults (Kolher, 1929:1947, Ramachandran and Hubbard, 2001) and toddlers (Maurer, Pathman & Mondloch, 2006) systematically match certain kinds of words to certain kinds of shapes according to the sounds of their phonemes (e.g., 'kiki'-spiky 'bouba'-curvy). Studies examining crossmodal matching among dyslexic adults (Drijvers & Dingemanse, 2003) have shown fewer sound-symbolic choices compared to normal adults, suggesting that crossmodal deficits may be a cause of the developmental difference. However, the developmental trajectory of dyslexia and sound symbolism remain largely unknown. Phonological awareness and vocabulary size at an earlier age are known predictors of later reading ability (Bradley & Bryant, 1983; Lyytinen & Lyytinen, 2004), yet no studies compare these skills with sound-symbolic matching in pre-reading children.

We will ask whether standardised measures of phonological awareness and vocabulary size of children at 4 years are related to later ability in sound-symbolic matching, by assessing whether there is a correlation between performance on phonological awareness score and vocabulary size at 4 years, and a sound symbolic matching task at 6 years.

A study for sound-shape matching in children between the ages of 5'10 and 6'6 years has been conducted, using the Alien Zoo task. The study is being run in collaboration with Growing Up in Singapore Towards Healthy Outcomes (GUSTO), a longitudinal birth cohort study in Singapore (www.gusto.org). As part of the GUSTO study, parents were asked to give an indication of their child's vocabulary size via Communicative Developmental Inventories (S-CDI) at 18 months and 24 months. At 48 months, children completed three subtests from CTOPP-2 (Elision, Blending, Sound Matching), the Peabody Picture Vocabulary task (Dunn & Dunn, 2007), and the Lollipop Test (Chew, 1981).

2) Describe the key dependent variable(s) specifying how they will be measured.

The 16-item Alien Zoo task provides each child with a Sound Symbolism score from 0-16. Testing for this part of the project has been completed, with data going directly to researchers in the BLIP Lab at NTU (n.b., Chinese Wall policy). In addition to this primary DV, the following tests have been conducted by the GUSTO Project:

1.S-CDI at 18 months and 24 months – total score for vocabulary size.
2.CTOPP-2 at 48 months – score for elision, blending and sound matching, composite score of phonological awareness.
3.PPVT at 48 months – total score for vocabulary size
4.Lollipop test at 48 months – Subscale for Identification of letters and writing

Testing for the CDI at 18 months and 24 months was completed between 2011 and 2012. Testing for CTOPP-2, PPVT and Lollipop task were completed in 2014. All testing was conducted by researchers at NUH, A*Star, and NUS, working directly on the GUSTO project. According to the Chinese Wall policy in the BLIP Lab, no data will be linked at the level of the individual until pre-registration has been completed.

3) How many and which conditions will participants be assigned to?

All participants who completed the Alien Zoo task and were included in the analysable cohort for each of the GUSTO measures (according to their in-house exclusion criteria) will be analysed, in a single group.

4) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

Primary Analysis: To test whether phonological awareness at 4y is related to sound symbolic matching at 6y, we will run a correlation analysis between the Alien Zoo score and the composite Phonological Awareness score in the CTOPP. According to the previous literature suggesting links between dyslexia and sound symbolism, we predict a positive correlation between pre-reading skills and scores on the Alien Zoo task. If these variables are correlated, we will follow up with an ANCOVA to investigate whether the continuous variable Phonological Awareness interacts with gender (male, female) to influence scores on the Alien Zoo task.

5) Any secondary analyses?

1. Inter-correlations between all variables will be checked before proceeding to the next analysis:
2. If data assumptions are met, a stepwise regression analysis will investigate the independent contributions of the following predictors on the Alien Zoo task:
a.CTOPP score.
b.CDI score (24m).
c.PPVT score (48m).
d.Lollipop test: Identification of letters and writing (48m)
In the event of multicollinearity between variables (e.g., CDI and PPVT), models excluding correlated variables be compared against each other, to establish which is the greater predictor of variance.
e. Gender (IFF significant in primary analysis)
3. If massive multicollinearity precludes rational use of regression models, Multiple Mixed Models will be considered as an alternative.
4. Any additional analyses will be considered strictly exploratory.

6) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.

The full GUSTO Cohort is comprised of around 800 children. Each neurocognitive testing wave invites all cohort members to be involved in testing, but not all children in the cohort take part in each round, due to drop-out. In addition, of those children who take part, not all children complete all tasks, due to drop-out/fussiness etc. Drop-outs are removed without replacement. The Alien Zoo task began after the 6y testing wave had already started. After this point, all children involved in neurocognitive testing were invited to perform the test. Hence, in this study, we will include all children who took part in all the data collection in the 4y and the 6y testing waves, and also…

1.Completed the Alien Zoo Task (N = 388)
2.Completed the CTOPP-2 at 48 months
3.Completed the PPVT at 48 months
4.Completed the Lollipop test at 48 months

The final N will be confirmed after data linkage.

7) Anything else you would like to pre-register? (e.g., data exclusions, variables collected for exploratory purposes, unusual analyses planned?)

The Alien Zoo task includes a small number of demographic questions about language use in different contexts. These data were collected for descriptive purposes, but may be used in further exploratory data analysis. The GUSTO cohort has also been involved in detailed language use surveys at 6m and 18m. These details may also be included in exploratory analyses. Two additional tests were conducted when the children were 48 months old, Visually Cued Recall and Random Object Span Test. We will not investigate the relationship between these variables and others included in the study in this pre-registered design.

Exclusions: If the child was not able to complete the Alien Zoo Task or was not compliant with the instructions e.g. picking shapes without paying attention to sounds, their results will be removed from the analysis according to the GUSTO team's notes on exclusions. Since this study involves archival results of the CDI, CTOPP-2, PPVT and Lollipop test previously conducted by the GUSTO team (under a Chinese Wall policy), any exclusions or data deemed unusable by that team will be taken as final and will not be included in the analysis registered here.

Data 'collection,' access and linkage: The Chinese Wall Policy

Testing was concluded for the previous waves of the GUSTO cohort, however, no identifiable details have been accessed by the BLIP lab research team, and data have been shared in a zipped, de-identified file. The data have not been unzipped, and no data linkage has taken place. According to the Chinese Wall policy:

• Data for the Alien Zoo Task is accessible to researchers at NTU. This data has been checked for distributional properties only to check task validity. No feedback from the data analysis has yet been shared with collaborators in the other research stream, who were conducting the test on a pre-defined cohort of participants.

• Data collection for the CDI at 18 months and 24 months were completed between 2011 and 2012. Data collection for CTOPP-2, PPVT and Lollipop task were completed in 2014. All testing was conducted by researchers at NUH, A*Star and NUS. Data have been transferred in a zipped de-identified format. Data have not been unzipped and data linkage has not been performed.

• Data linkage (i.e., "data collection") between individuals' different tests will not occur until preregistration is complete.

8) Have any data been collected for this study already?

No, no data have been collected for this study yet