From Alien Zoo to Spy School: A Preregistered Study of Linguistic Sound Symbolism and its Links to Reading in 8-year-olds

Fei Ting Woon (feitingwoon@ntu.edu.sg)[1],
Defu Yap (defu.yap@ntu.edu.sg)[1],
Shirong Cai(obgcais@nus.edu.sg)[2],
Evelyn C. Law (evelyn_law@nuhs.edu.sg)[3],
Lourdes Mary Daniel (mary.daniel@kkh.com.sg)[4],
Suzy J. Styles (suzy.styles@ntu.edu.sg)[1,2]

[1]Psychology, School of Social Sciences, Nanyang Technological University, 14 Nanyang Drive, S(637332), Singapore,
[2]Singapore Institute for Clinical Sciences, Agency for Science Technology and Research (A*STAR), Singapore,
[3]National University Health System, Singapore,
[4]Department of Child Development, KK Women's and Children's Hospital, Singapore

Abstract

Adults and children systematically match certain kinds of words to certain kinds of shapes according to the sounds of their phonemes (e.g., 'kiki'-spiky 'bouba'-curvy). These sound-shape mappings rely on multisensory processing of perceived goodness of fit between vision and audition. Dyslexic individuals have shown deficits in general multisensory processing and sound-symbolic matching suggesting that multisensory processing deficits may be developmentally implicated in early reading difficulties. A longitudinal cohort study tracking bilingual children in Singapore showed that early predictors of English reading at 4 years (e.g., phonological awareness, vocabulary size and letter knowledge) did not correlate with a novel child-friendly task eliciting the bouba-kiki effect at 6 years. However, since the children had not yet started formal reading instruction, it is difficult to interpret the lack of relationship. In the current study, we followed the same cohort of children into early reading years and tested their English word and pseudoword reading abilities at 8.5 years. In our preregistered analysis, no significant relationship was observed between earlier multisensory sound-shape matching reading outcomes but known predictors of reading showed strong relationships in this cohort of bilingual children.

Keywords: reading; bouba-kiki; early predictors of reading; developmental trajectory; dyslexia; multisensory processing

Introduction

Learning to read represents the acquisition of stable links between stimulus types processed in vision and audition; fluent readers integrate visual symbols (graphemes) with abstract categories of speech sounds (phonemes) effortlessly when they engage with text. The intersensory linkages are therefore a kind of cross-modal association. Research suggests that individuals with difficulty in reading also show difficulties in making multisensory mappings between graphemes and phonemes (Blau et al., 2010; Froyen et al., 2011), and adults diagnosed with developmental dyslexia even show impaired multisensory mapping in nonlinguistic domains including white noise and Gabor patches (Harrar et al., 2014). These findings suggest that impairments in multisensory processing/mapping may be related to impairments in reading.

In an audio-visual matching task, popularly known as the 'bouba-kiki' task, participants are asked to match auditory speech tokens with either rounded or spiky shapes (Köhler, 1929; Ramachandran & Hubbard, 2001). Adults and 2-year-olds systematically match certain speech sounds to certain kinds of shapes (e.g., 'bouba'-curvy', kiki'-spiky) (Maurer et al., 2006). This preferential matching task has been replicated across various populations. One recent meta-analysis showed that around 89% of normal adults make 'congruent' matches when the nonsense words contain canonically 'round' or canonically 'spiky' phonemes (/b, m, l, o, u/ and /k, t, i, e/, respectively), but this task fails when the tested speech tokens do not comply with phonotactic rules of the participants' language (Styles & Gawne, 2017).

What about individuals who struggle to read? Drijvers, Zaadnoordijk, & Dingemanse (2015) tested Dutch dyslexic adults with the bouba-kiki paradigm and found that they made fewer congruent sound-shape matches compared to normal adults, and suggested that their matching task may be a marker of a broader pattern of multisensory deficits that impair normal reading. Since systematic preferences for sound-shape matches occur in pre-reading children (Maurer et al., 2006), tests of multisensory matching may be perfect for detecting children at risk of reading difficulties, using a task that does not rely on higher-order metalinguistic knowledge (e.g., letter knowledge, phonological awareness).

Recently, 377 bilingual English-speaking children (average age: 5 years 10 months) in Singapore completed a novel, child-friendly task eliciting this bouba-kiki effect, the Alien Zoo task (Woon & Styles, 2017). The children were presented with sixteen trials of yoked auditory-visual stimuli and were asked to choose "which Alien (visual token) do you think has this name (recorded auditory token e.g, 'pikeki')?". The children systematically made congruent matches albeit at a lower rate than the adults in the control group. In a preregistered analysis, their results in the Alien Zoo task were examined with their predictors of English reading at 4 years

(i.e., phonological awareness, vocabulary size and letter knowledge). Strong inter-correlations were found amongst the predictors of reading. However, no significant relationship was found between predictors of reading at 4 years and sound-symbolism at 6 years (Woon et al., 2018). Hence, although sound-symbolic matching was not related to known predictors of dyslexia in this group of pre-schoolers, it may be the case that differences only emerge after the onset of formal reading instruction. We planned to follow up the same group of children with a reading assessment conducted after the first grade of primary school (P1) to see if those children who made fewer congruent sound-symbolic choices would go on to show weaker reading skills.

Methods

We conducted a preregistered test of the relationships among linguistic sound-symbolic matching, early predictors of reading, and later reading outcomes in a large cohort of 8.5-year-olds in Singapore. The children are enrolled in a longitudinal birth cohort and have been tracked since birth. We compared the children's previously recorded scores on known predictors of reading at age 2 and 4, with their tendency for sound-symbolic matching at 6 years in the Alien Zoo task, and their reading outcomes at 8.5 years.

Participants

The children were recruited from Singapore's largest longitudinal cohort study, Growing Up in Singapore Toward healthy Outcomes (GUSTO)(Soh et al., 2014). They are bilingual in English and one of the following languages: Mandarin Chinese, Malay, or Tamil. 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. Following the GUSTO cohort study design, the children were measured on a variety of cognitive and developmental tests including known predictors of reading ability. At 2 years of age, children were administered a vocabulary inventory, a Singapore adaptation of the Macarthur-Bates Communicative Developmental Inventory (CDI) (Tan, 2009); at 4 years of age the Comprehensive Test of Phonological Processing (CTOPP-2) (Wagner et al., 2013); the Peabody Picture Vocabulary Test (PPVT-4) (Dunn & Dunn, 2007), and the Lollipop Test (Chew, 1981).

In our previous study, at 6 years of age, 377 children from the cohort completed a sound-shape matching task, Alien Zoo, designed to elicit their preference for congruent sound-shape matches (Woon et al., 2018). Given the relationship between multisensory processing and reading, we predicted that children who achieved low scores on the multisensory matching task would go on to have difficulties learning to read. At 8.5 years, 285 of the GUSTO children (Range: 8.5y – 9.9y, Ave = 8.8 years old, *SD* = 0.11) completed the TOWRE-2 reading task, which comprised of a Word-reading and a Non-word reading task (Torgesen et al., 2012). Data collection for the 8.5-year-old age group was interrupted due to the COVID-19 lockdown measures in Singapore and was concluded in August 2020. Preliminary results for the 6-year wave were reported previously. 118 children from the GUSTO cohort completed all tasks at ages 2-8.5y and are included in the preregistered analysis reported here.

GUSTO Recruitment	Known Reading Predictors				Multisensory Processing	Reading Skills	
	CDI Score	PPVT (receptive)	CTOPP (Phon Awareness)	Lollipop (Letter Knowledge)	Alien Zoo	TOWRE 2 Words	TOWRE 2 Nonwords
Pre-natal	2y	4y			6-y	8.5-y	
2009-2010	Previously reported relationships (Woon et al., 2018)					Current Report	

Figure 1. Timeline of study elements

Stimuli & Procedure

As part of the cohort study, the children are invited for a day of activities and tests as they turn 8.5 years. The TOWRE-2 reading and non-word reading tasks are administered as the last task of the day for the children. The TOWRE-2 test of reading requires children to read aloud from a list of words and non-words as quickly as possible within a time limit. There are 4 sets of test stimuli (A, B, C, D), each containing one 'Word' list and one 'Non-word' list. According to the TOWRE-2 manual, all four sets are equivalent in difficulty (Torgesen et al., 2012). There are separate word and non-word lists, and children always complete the 'Word' task before the 'Non-word' task. Before the start of each task, the researchers conducted a practice task. For each task, the children are given 45 seconds and were asked to read as many words as possible. Each list starts with relatively easy words (e.g., 'is') and gradually increases in difficultly (e.g., 'limousine'). Children are allowed to skip words.

The words and non-words increase in the number of syllables, complexity, syllable length. The non-words are designed to reflect grapheme-phoneme correspondences in English. Participants were sequentially assigned one set of the test lists, starting with Participant 1 receiving set A, Participant 2 receiving set B and so on. During testing, research assistants awarded one mark for each accurately read word. For this analysis, we followed standard TOWRE-2 scoring guidelines.

Our previous experience has shown that research assistants are sometimes uncertain of whether to mark a word as correct or incorrect for the local variety of English. To facilitate future offline scoring, children's verbal responses were recorded on a Zoom H4n Pro recorder. They wore a Rode Lav headset with a SmartLav+ lapel microphone to ensure clear near-field recordings. To encourage children to wear the headset and make the task more engaging, we created a novel

storyline for the task, a 'wrapper' for the TOWRE-2 task called Spy School. Children were told that they had been recruited into a Spy School and their first task was to read "code words" as quickly as possible to help their fellow spies escape from a maze. Each "code word" they read aloud opened a door for their fellow spies. At the end of the task, the participants were rewarded with an "I survived Spy School" sticker. This Spy School wrapper for the TOWRE-2 reading tasks is archived publicly (Woon & Styles, 2021).

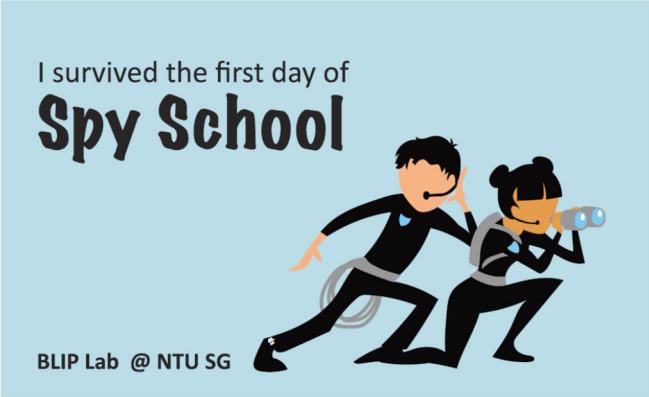

Figure 2: Sticker given to participants who completed our Spy School wrapper for TOWRE-2 reading tasks.

Predictions

In our previous report, the children were above chance when choosing congruent sound-shape matches in the 16-trial Alien Zoo task, demonstrating some tendency to respond with congruent choices, albeit less strongly than adults (Woon et al., 2018). In addition, the range of the children's scores was wide and normally distributed, providing a good range for analysis of individual differences.

If weak multisensory preferences are related to weak reading abilities as seen in Dutch dyslexic adults (Drijvers et al., 2015), we predicted a positive relationship between the number of congruent sound-symbolic choices an individual made in the Alien Zoo task at age 6 and their scores in reading tasks in TOWRE-2 at age 8.5. We also expected reading scores to be higher for Words than for Non-Words, and as dyslexic children struggle more with decoding unfamiliar non-words than reading familiar real words (Elbro et al., 1994; Rack et al., 1992), we predicted that the relationship between multisensory sound-symbolic matching and reading scores might be stronger in the non-word reading task. We also predicted a positive relationship between the earlier predictors of reading (ages 2 years to 4 years) and early reading outcomes in the Spy School TOWRE-2 reading tasks (Preregistration: Woon, Yap, & Styles, 2021).

Data-handling and Analysis Plan

In our previous analysis of the data collected at age 2y and age 4y, strong correlations were found between all four measures which are known in the literature to predict reading. To prevent multicollinearity in a generalized linear modelling approach, we followed a preregistered analysis pathway and performed dimensional reduction using Principal Components Analysis to extract factors with shared sources of variance among the four measures. Significant factors arising from this analysis were then used as predictors in our linear mixed model analysis. In our **Primary analysis**, we planned to test a positive relationship between Spy School TOWRE-2 reading scores at age 8.5 years and Alien Zoo sound-symbolic matching score at 6 years using a general linear mixed effects model in R (R Core Team, 2020). The fixed effects include the number of congruent matches in the Alien Zoo task, Spy School TOWRE-2 task type (Word or Non-word), and factor(s) arising from the PCA analysis of predictors of reading. The random factors are participants and the precise age at which when the Spy School TOWRE-2 tasks were completed. Particularly, we are interested in the interaction between the Alien Zoo score and the Non-word TOWRE-2 task. Our **secondary analysis** involves examining the correlations between the Spy School TOWRE-2 scores and the significant predictor(s) identified by the analysis.

Given the longitudinal nature of the data used in analyses and variable completion rates for different test waves, all children for whom the complete set of measures were available are included in the modelling analysis reported here.

Results

118 children completed all seven tasks over four testing waves, and are included in the current analysis, with no exclusions. Figure 2 shows the scores for the Alien Zoo task and the two tasks of the TOWRE-2 test of reading. Alien Zoo scores ranged between 5 and 16 ($M = 9.6$, $SD = 2.45$), and their scores on the Word reading lists were higher than their scores on the Non-Word reading lists (Words: $M = 62.6$, $SD = 10.1$; Non-Words: $M = 35.2$, $SD = 11.9$; $t(117) = 19.13$, $p = 2.2e^{-16}$).

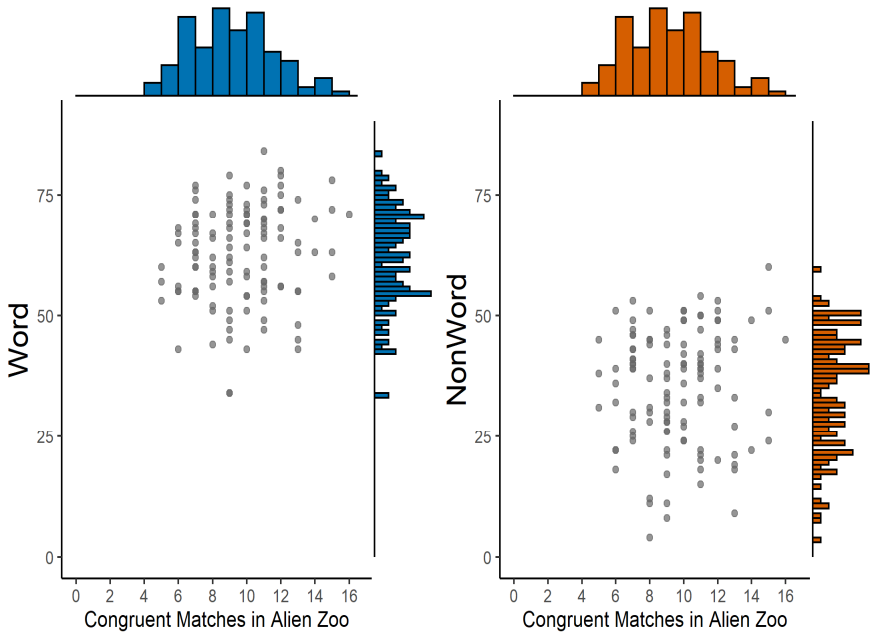

Figure 3: Alien Zoo multisensory congruent matches at 6y and Spy School TOWRE-2 reading tasks raw scores at 8.5y.

Relationships among Reading Predictors

The exploratory Principal Component Analysis identified four factors, with the first component accounting for 58% of the variance, with an Eigenvalue of 2.19. The first component has positive associations with English vocabulary sizes at 2y (.39) and 4y (.58), phonological awareness at 4y (.55), and letter knowledge at 4y (.45), so we characterise this Factor as 'General pre-reading skills' in English.

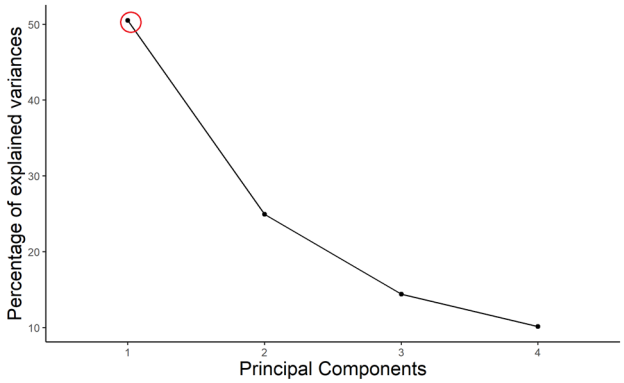

Figure 4: Scree plot of the components identified by the PCA on early predictors of reading.

The second factor does not equivocally account for a substantial proportion of extra variance in this small model, as assessed by Scree plot (Figure 1), and limited explanatory value (Eigenvalue: .91; variance: 22.8%, inconsistent factor loadings). In line with our preregistered analysis plan, we included Component 1 as a predictor of reading scores in subsequent models. Supplementary analysis (omitted for space constraints) reveals Factor 2 does not significantly add to the GLM of reading scores.

Reading Outcomes & Sound Symbolism

To test whether sound symbolic matching on the Alien Zoo task at age 6 was a predictor of reading skills in early readers at age 8 we conducted a preregistered GLM of reading scores with Task type, Alien Zoo score and the PCA derived General prereading skills score as fixed factors, along with the interaction between task type and alien zoo, and participant and age as random factors. Contrary to the primary hypothesis, we did not find a significant relationship ($t(168.9) = .98$, $p = .33$) between the number of congruent sound-symbolic matches in the Alien Zoo task and reading outcomes from the Spy School TOWRE-2 reading tasks, nor did the relationship between these scores differ for the two kinds of task ($t(116) = -.86$, $p = .39$).

In line with the existing literature on predictors of early reading, a strong significant relationship between the general prereading skills component identified by PCA and later reading outcomes was found ($t(115) = 4.66$, $p < .001$). The model which included the General prereading skills component was shown to have a better fit compared to the model without ($\chi^2(1) = 20.62$, $p < .001$). As the component of General prereading skills was identified as a significant predictor, planned correlations were run. General prereading skills were strongly correlated with Word reading ($R = .47$, $p < .0001$) and Non-word reading ($R = .28$, $p < .0001$).

Table 2: Results of fixed effects from GLM predicting Spy School TOWRE-2 reading task.

Fixed effects	Estimate	Std Error	*t*	*p*
Intercept	58.91	3.90	15.08	<.001
Alien Zoo	0.38	0.39	0.98	0.33
Task Type	-24.55	3.49	-7.04	<.001
Prereading	2.70	0.58	4.67	<.001
Alien Zoo* Task Type	-0.30	0.35	-0.86	0.391

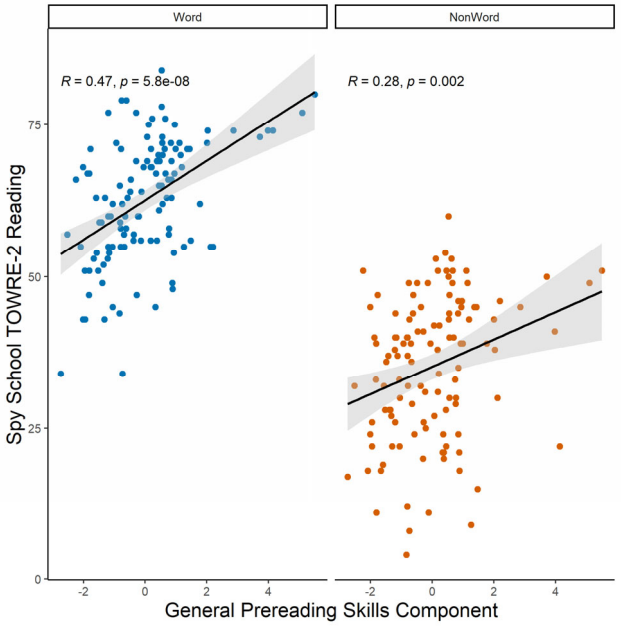

Figure 5: Correlations between Word/Non-Word reading at 8.5y and early general prereading skills derived from tests at ages 2-4y.

Discussion

We wanted to know if the number of congruent sound-symbolic choices made by children at 6 years of age predicted their English reading outcomes at 8.5 years of age. There is evidence that sound-symbolism precedes reading (Maurer et al., 2006) and research has suggested that sound-symbolism may provide a bootstrap for vocabulary learning in infancy (Imai et al., 2015). By following the design of a longitudinal cohort study, we were able to examine, for the first time, the developmental trajectory of early pre-reading skills, sound-symbolism during pre-schooling years, and reading outcomes after the onset of formal instruction in a large cohort of children. In a previous study, Dutch dyslexic adults made fewer congruent sound-symbolic choices (60%) compared to normal readers (73%) (Drijvers et al., 2015) leaving open the question of whether deficits in multisensory processing precede and/or cause deficits in reading. We posited that there are three possible developmental relations between multisensory matching and reading: (i) weaker multisensory processing precedes (and possibly causes) later reading difficulties, (ii) weaker reading precedes (and possibly causes) later weakness in multisensory processing, (iii) weaker multisensory processing and weaker reading are discrete deficits but may arise from a common deficit of neural representations of language.

In a previous preregistered study, we examined early relationships between multisensory processing and prereading skills, using known predictors of reading at 2y and 4y as proxies for reading ability. At this early age, we did not find significant relationships (Woon et al., 2018), but this could be because the children had not yet begun formal reading instruction so differences may not have been apparent. In the current preregistered study, we were able to follow up the same children in a clearer examination of the first developmental relationship (above) using sound-shape multisensory matching scores in the Alien Zoo task at 6 years of age and reading scores at 8.5 years of age. As English reading involves a multisensory link between vision (letters) and audition (phonemes), if multisensory deficit seen in dyslexic adults in Drijvers et al.'s study (2015) precedes the onset of reading instruction, we expected that the children who made fewer congruent sound-symbolic choices would also score lower for the reading tasks, in particular, for non-word reading. We were optimistic that the Alien Zoo adaptation of the bouba-kiki task could provide a valuable language-independent screening tool, complementary to existing tools for identifying children at risk of developing language-related difficulties, if a relationship had been found. However, our predictions were not borne out for this cohort of Singaporean children.

Aside from multisensory processing and in line with existing literature on early reading, bilingual children in the current study showed a robust predictive relationship between early pre-reading skills and later reading outcomes, particularly for real words of English. This finding aligns with existing literature on the language acquisition trajectory for monolingual English speaking children (Kuhl, 2011).

The reading tasks were conducted during the year 8.5 testing wave for the children in the GUSTO cohort from March 2019 to August 2020. Due to the COVID-19 lockdown measures in Singapore from April to June 2020, testing schedules were affected which resulted in a much smaller group of the cohort who completed both the Alien Zoo task at 6y and Spy School TOWRE-2 reading tasks at 8.5y. However, given our analysis, we are not confident that the Alien Zoo multisensory matching task at 6y will predict reading outcomes at 8.5y even with all the children in the cohort included.

In line with standard administration of the TOWRE-2 reading task, the scoring of the children was done online when the children were reading the words aloud. All the scorers were speakers of Singapore English and from the same linguistic community as the children. The TOWRE-2 answer manual provides pronunciation guides for each word using vowels spoken by US English speakers. Recent research showed that fluent Singapore English speakers'

perception of English vowels differs from standard UK English or US English pronunciations; highlighting the possibility of vowel mergers in Singapore English (e.g., a merger between long and short vowels /i/) (Leung, 2019). Given the pronunciation differences in the local variety of English, scorers were asked to judge the pronunciation of the children by comparing it to their own Singapore English pronunciations. Thus, understandably, there may be individual scoring differences.

In a related project, we collected audio recordings of several skilled adult Singaporean readers articulating all the items in the TOWRE-2 test sets (Woon, Yogarrajah, et al., 2021). We plan to model skilled Singapore English readers' pronunciations of each word on the TOWRE-2 task to create a standard scoring guide for researchers using the task in Singapore and in regions where English speakers have a vowel repertoire similar to Singapore English. Once the scoring guide is complete, offline scoring of the children's responses reported here will be possible.

Since offline scoring requires clear audio recordings, all children in the study wore a headset mic while performing the task. Wearing the headset was embedded in our novel story framework called 'Spy School'. With the 'Spy School' narrative, the children were happy to wear the headset and enjoyed the task. No children dropped out of the task before completion. We believe the 'Spy School' wrapper may be a useful addition to the standard TOWRE-2 administration especially in contexts where a) children experience performance anxiety during reading assessments and b) offline scoring is desirable – for example, where children speak different regional varieties of English.

Acknowledgements

This research is made possible by a NAP grant to Suzy J Styles: *The Shape of Sounds in Singapore* and NRF grant to Suzy J Styles: *How language mixes contribute to effective bilingualism and biliteracy in Singapore* (NRF2016-SOL002-011). It is also supported by the Singapore National Research Foundation (NRF), Prime Minister's Office (PMO), under its "Translational Clinical Research (TCR) Flagship Programme on Developmental Pathways to Health and Disease: Metabolic, Neurodevelopmental and Related Outcomes" (DevOS), administered by the Singapore Ministry of Health's National Medical Research Council – NMRC/TCR/004-NUS/2008;NMRC/TCR/012-NUHS/2014 and additionally funded by the Singapore Institute for Clinical Sciences (SICS), Agency for Science, Technology & Research (A*STAR).

Author Contributions

Designed the GUSTO longitudinal study: GUSTO; Designed and coded the Alien Zoo Sub-task and created Spy School TOWRE-2 Wrapper: FTW & SJS; Tested the participants: GUSTO & FTW; Contributed earlier language measures resources: GUSTO; Pre-registered the analysis, and analysed the data: FTW, DFY & SJS; Wrote the paper, FTW & SJS.

References

Blau, V., Reithler, J., van Atteveldt, N., Seitz, J., Gerretsen, P., Goebel, R., & Blomert, L. (2010). Deviant processing of letters and speech sounds as proximate cause of reading failure: a functional magnetic resonance imaging study of dyslexic children. *Brain*, *133*(3),868–879. https://doi.org/10.1093/brain/awp308

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

Drijvers, L., Zaadnoordijk, L., & Dingemanse, M. (2015). Sound-Symbolism is Disrupted in Dyslexia : Implications for the Role of Cross-Modal Abstraction Processes. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, 602–607.

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

Elbro, C., Nielsen, I., & Petersen, D. K. (1994). Dyslexia in adults: Evidence for deficits in non-word reading and in the phonological representation of lexical items. *Annals of Dyslexia*, *44*(1), 203–226. https://doi.org/10.1007/BF02648162

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

Harrar, V., Tammam, J., Pérez-Bellido, A., Pitt, A., Stein, J., & Spence, C. (2014). Multisensory integration and attention in developmental dyslexia. *Current Biology*, *24*(5),531–535. https://doi.org/10.1016/j.cub.2014.01.029

Imai, M., Miyazaki, M., Yeung, H. H., Hidaka, S., Kantartzis, K., Okada, H., & Kita, S. (2015). Sound Symbolism Facilitates Word Learning in 14-Month-Olds. *PLOS ONE*,*10*(2). https://doi.org/10.1371/journal.pone.0116494

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

Kuhl, P. K. (2011). 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

Leung, W. T. (2019). *Challenges of teaching phonics-based reading remediation in Singapore* [Nanyang Technological University]. https://doi.org/10.32657/10356/137123

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

R Core Team. (2020). *R: A language and environment for statistical computing*. (4.0.3). R Foundation for Statistical Computing. https://www.r-project.org/

Rack, J. P., Snowling, M. J., & Olson, R. K. (1992). The

Nonword Reading Deficit in Developmental Dyslexia: A Review. *Reading Research Quarterly*, *27*(1), 28–53. https://doi.org/10.2307/747832

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

Soh, S.-E., Tint, M. T., Gluckman, P. D., Godfrey, K. M., Rifkin-Graboi, A., Chan, Y. H., Stunkel, W., Holbrook, J. D., Kwek, K., Chong, Y.-S., Saw, S. M., Sheppard, A., Chinnadurai, A., Ferguson-Smith, A., Goh, A. E. N., Biswas, A., Chia, A., Leutscher-Broekman, B., Shuter, B., … 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

Styles, S. J., & Gawne, L. (2017). When does maluma/takete fail? Two key failures and a meta-analysis suggest that phonology and phonotactics matter. *I-Perception*, *8*(4).

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

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). *Test of Word Reading Efficiency (TOWRE-2)* (Second). PRO-ED.

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

Woon, F. T., Chong, Y.-S., Daniel, L. M., Broekman, B. F. P., Cai, S., & Styles, S. J. (2018). Pre-Readers at the Alien Zoo: A Preregistered Study of the Predictors of Dyslexia and Linguistic Sound Symbolism in 6-year-olds. In C. Kalish, M. Rau, J. Zhu, & T. T. Rogers (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society (CogSci 2018)* (pp. 2699–2704). Cognitive Science Society. http://mindmodeling.org/cogsci2018/papers/0512/0512.pdf

Woon, F. T., & Styles, S. J. (2017). Linguistic Sound Symbolism and Learning to Read: Preliminary results from a large-scale prospective study. *Society for Research in Child Development Biennial Meeting*.

Woon, F. T., & Styles, S. J. (2021). Spy School: TOWRE-2 Reading Task Wrapper. *OSF Respository*. https://osf.io/stzj6/

Woon, F. T., Yap, D., & Styles, S. J. (2021). Multisensory Alien Zoo Task and links to Reading and Dyslexia. *OSF Preregistration*. osf.io/8e574

Woon, F. T., Yogarrajah, E. C., & Styles, S. J. (2021). *Skilled Singapore Readers for TOWRE-2*. DR-NTU. https://doi.org/10.21979/N9/ORFGEM