Title	Animated electronic storybook and children's Mother Tongue development: Tracing the process and the outcome with eye-tracking
Author(s)	Sun He

OER FINAL REPORT SERIES

The OER Final Report series includes final reports from funds managed by Office of Education Research, National Institute of Education, Nanyang Technological University.

Reports are submitted as part of the funding review process and intended for the funding agency, local schools and educators, teacher educators, policymakers, and education scholars. They do not take the place of scholarly, peer-reviewed articles but report on the background, procedures, and major findings of the project.*

This study was funded by Singapore Ministry of Education (MOE) under the Education Research Funding Programme (OER 13/16 HS) and administered by National Institute of Education (NIE), Nanyang Technological University, Singapore. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Singapore MOE and NIE.

*In some cases reports show coloured font or highlights. These are an artifact of the review process and not intended to have any special weight or meaning within the report itself.

Version: 17 Jul 2020, NIE OER

EDUCATION RESEARCH FUNDING PROGRAMME

PROJECT CLOSURE REPORT

Animated electronic storybook and children's Mother Tongue development: Tracing the process and the outcome with Eye-tracking

By

Sun He

National Institute of Education
Singapore

EXECUTIVE *SUMMARY (NO MORE THAN 5 PAGES)*

Introduction/Background

Storybook reading to children is considered an efficient way to provide a meaningful context for exposure to unfamiliar words and grammar (Weizman & Snow, 2001). Nonetheless, children with limited language knowledge (e.g., child MTL learners in Singapore) may benefit less from the reading activities, due to the gap between their skills and those required for processing the narration. They often fail to derive the meaning of unknown words/grammar from the verbal context and consequently have trouble figuring out the story plots (Verhallen & Bus, 2010). Children's electronic storybook (e-storybook) seems to hold great promise to assist in developing children's emerging literacy as such reading formats are favored by children due to its entertaining elements (e.g., sound and interactive games) (Hio, 2015).

Compared to traditional print books, well-designed animated e-storybooks can stimulate readers' visual, auditory and even kinaesthetic senses to comprehend a story and unfamiliar language via the match between nonverbal sources (motion pictures, images, sound and music) and the narration (de Jong & Bus, 2002; 2004; Neuman, 1997; Verhallen, Bus, & de Jong, 2006). It is especially promising for second language learners/bilingual child learners, because these children with less language knowledge at hand may need extra information sources to digest the story plots and grasp the new words and grammar.

Statement of Problems

According to Singapore National Library Board, the number of electronic books borrowed has reached eleven million in 2015, increasing approximately four times since 2009. Despite the changing landscape of children's storybook in Singapore, little is known how these children e-storybooks influence bilingual language learning and whether children can ultimately benefit from such exposure, for their Mother Tongue language (MTL) in particular. Although bilingualism is a corner stone of the Singapore education system, recent years have witnessed an inclination for English use in bilingual families due to the change of the sociolinguistic landscape.

To strengthen Singapore children's MTL development as emphasized by the Ministry of Education Singapore, the investigation into approaches that motivate children to better master MTL seems more urgent and necessary than ever. Against this social backdrop, research on animated e-storybooks is especially meaningful because of their promising effect on children's emergent literacy skills. This study will help us understand whether Singapore pre-schoolers would benefit from animated electronic storybook exposure for their vocabulary learning and reading comprehension in MTL. Moreover, by following children's attention foci and span with eye trackers, we could tap into children's learning process over repeated reading and explore the potential of different e-storybook features for a better learning outcome. This proposed research would broaden the scope of inquiry for related projects pertaining to children's MTL learning, as well as the integration of new learning modalities to enhance early childhood language education.

Purpose of Study

This study will explore the efficacy of animated e-storybooks on pre-schoolers' Chinese development. The relation between features of the animated e-storybook (e.g., sound and motion) and children's visual attention (as measured with eye tracking equipment) will be explored with mixed effects model. Children's vocabulary and reading comprehension will be compared after using an animated e-storybook and a corresponding static version. There are three objectives in the current study:

• The first objective is to examine whether animated electronic storybooks would promote children's learning outcomes (vocabulary learning and reading comprehension) compared to the traditional book format.
• The second objective is to find out whether the special features of e-storybooks (i.e., sound effects and motion pictures) could help retain more and longer visual attention from children.
• The third objective is to examine children's visual attention when processing the e-storybook features as described above. We focus on how their visual attention span changes from initial to subsequent repeated readings of the book paying less attention to key details in pictures, and whether the changing patterns vary between animated and static e-book reading groups.

Participants

102 English-Chinese bilingual learners (4-5 years old; 49 boys and 53 girls) were selected from Singapore preschools for the experiment. In order to control the impact of variables such as language proficiency, social economic status and e-book reading preference, which might affect children's picture book processing and comprehension, a larger group of children (N=188) were tested for Chinese proficiency and investigated via parental questionnaire before the experiment. Consent forms were sent to parents for their willingness to have their children participate in this project. Participants should not have a language impairment record, had Chinese exposure at home and attained preliminary Chinese proficiency before the project started.

Methodology / Design

Description of the methods and design used, including data collection and analysis

The study was a randomized control trial (RCT) including four experimental conditions, we estimated the optimal sample size ($n=1.962\sigma 2/E2$ where σ = standard deviation and E = error rate):

• Static e-books, with neither sound nor motion (N=28): Children independently read three e-books four times spread over two weeks, in which static illustrations were accompanied by only an oral narration of the story.

• Static e-books, with sound only (N=29): Children independently read three e-books four times spread over two weeks, in which static illustrations were accompanied by background music, motion sounds and oral narration of the story.

• Animated e-books (N = 32): Children independently read three animated versions four times spread over two weeks, in which animated illustrations (including motion, music and sound) accompanied an oral narration of the story.

• Control condition (N = 13): children independently played non-literacy related math games four times spread over two weeks during the intervention period.

Children were pre-tested for their general Chinese receptive and productive vocabulary size and the target vocabulary in the storybooks. As indicators of general cognitive proficiency, their nonverbal intelligence, and phonological short-term memory were tested with standardized tests. The receptive vocabulary component of the Bilingual Language Assessment Battery (BLAB; Sze & Rickard Liow, 2009) was used to estimate children's general Chinese proficiency. Digit span and non-word repetition, the sub-tests of the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen & Rashotte, 1999) was administered to assess the children's short-term phonological memory. Raven's Colored Matrices (Raven, Raven & Court, 1998, sets A, B and AB) was used to estimate children's analytical reasoning, which is assumed to be an important component of language aptitude (Paradis, 2011).

Three children stories, Little Kangaroo (Genechten, 2007), Imitators (Veldkamp, 2006) and Cycling with Grandpa (Boonen, 2004) were used as reading materials. Static versions of these three stories consisted of scanned pictures from the original print storybooks and were presented on screen. The story text was automatically read aloud and the stories continued without any action. In the animated versions, the static illustrations representing the story events are dramatized by using motion sound and background music. Each reading session includes three stories and lasted about 15 minutes including the preparation time.

The Tobii X3-120, a remote eye tracker, was used to measure the length and frequency of eye fixation while the recorded voice read the stories to children. The eye tracker was mounted to the bottom of the laptop screen. The total fixation duration on the storybooks as a whole and the pictorial locations of 23 target words in storybook was being analyzed. We specified the pictorial location of the target words as Areas of Interest (AOI) with the software - Tobii Studio 3.4.5. The duration of children fixating the AOI and the frequency of such fixations was recorded by Tobii Studio's fixation filter with the default settings from velocity and distance threshold. We were particularly interested in longer fixations (i.e., gaze duration longer than 150 ms) because previous literature revealed that it takes such amount of time to process complex visual information as those in e-storybooks (Rayner, Smith, Malcolm & Henderson, 2009).

During the experimental session, a research assistant was present to operate the eye tracking machine (e.g., calibration) and to instruct the children to read the story. No interaction was allowed once children start to read the e-books, other than reminders to pay attention to the stories. Children in the control group spent the same amount of time on screen as those in experimental groups. Within one week after the last experimental session, children was tested again for the target words (post-test).

23 words were selected from the 3 stories for reception and production tests. They were tested at pre-test and at post-test in random order. Story sentence repetition was considered an indicator of grammatical understanding. Fifteen sentences were selected from the stories based on the following criteria: each sentence included 5 or more words, and these sentences were composed with increasing difficulty levels. They were applied at the post-test in a random order. Story recall were also conducted directly after the first and fourth experimental sessions. The three static e-storybooks were shown on screen page by page and children were asked to retell each story based on their understanding and memory.

Findings / Results

Productive Vocabulary

There was a significant effect of Group on productive vocabulary ($F(3,68.03) = 4.05$, $p=.010$). Pairwise comparisons revealed that the animated eBook condition resulted in significantly higher productive vocabulary scores than the sound-only ($p=.021$, $d=0.64$), static ($p=.015$, $d=0.69$), and control ($p=.005$, $d=1.23$) conditions. In comparison, there were no significant differences between the sound-only, static, and control conditions. Higher total pre-test vocabulary scores ($p<.001$) and Mandarin proficiency ($p<.001$) also significantly increased productive vocabulary.

Receptive Vocabulary

Mandarin proficiency again had an effect, with higher scores resulting in significantly higher receptive vocabulary ($p=.017$). There was also a significant effect of Group ($F(3,68.90) = 3.62$, $p =.017$). Pairwise comparisons (figure 2b) revealed no significant differences between the eBook types, but significantly higher receptive vocabulary scores in all eBook conditions in

comparison to control (animated, p=.002, d =1.31; sound-only, p=.031, d=0.90; static, p=.018, d=1.01).

Context Integration

There was a significant effect of group on context integration (F(3,59.90) = 3.03, p = .036). Pairwise comparisons revealed that the animated eBook condition significantly improved context integration in comparison to the static (p=.011, d=0.70) and control (p=.035, d=0.85) conditions, but not in comparison to the sound-only condition. Higher total pre-test vocabulary scores (p=.001) and Mandarin proficiency (p=.004) also significantly increased context integration.

Meaning Recognition

Higher Mandarin proficiency again had a significant effect (p< .001), giving higher meaning recognition scores. There was also a significant effect of group on meaning recognition, (F(3,93.00) = 7.65, p< .001). Pairwise comparisons (figure 2d) revealed no significant differences between the eBook types, but significantly higher meaning recognition in all eBook conditions in comparison to control (animated, p<.001, d=1.59; sound-only, p<.001, d =1.28; static, p<.001, d=1.27).

Story Retelling

There were significant effects of group, assessment, and general Mandarin proficiency for all three stories. Children were better at retelling in assessment 2 and those with higher Mandarin proficiency performed better. For Little Kangaroo, there was also a significant interaction of group x assessment (F(3,151,78) = 4.29, p=.006). Post-hoc pairwise contrasts were used to explore the effect of eBook condition on the retelling of each story. For Cycling with Grandpa, the animated condition significantly improved retelling ability in comparison to all other conditions (sound-only, p=.047, d=0.55; static, p=.029, d=0.63; control, p < .001, d=3.26). For Little Kangaroo, there were no differences between the animated, sound, or static conditions, but all were significantly higher than control (animated, p< .001, d=3.30; sound-only, p< .001, d=3.29; static, p<.001, d=3.03). Similarly, for Imitators, there were only significant differences between the eBook conditions and control (animated, p<.001, d=2.60; sound-only, p<.001, d=2.56; static, p<.001, d=2.25).

When examining development rate, for Little Kangaroo, all eBook conditions showed significantly greater assessment 1 to assessment 2 improvement than control (animated, p<.001; sound-only, p=.004; static, p=.022). However, for the other two stories, only the animated (Cycling with Grandpa, p=.025; Imitators, p=.024) and sound (Cycling with Grandpa, p=.014; Imitators, p=.016) conditions showed greater improvement than control.

Visual Attention

ANOVA summaries for the three stories are in table 5. There was a significant effect of group for Cycling with Grandpa (F(2,85)=3.41, p=.038), for Little Kangaroo (F(2,82.34)=12.89, p<.001), and for Imitators, (F(2,85)=4.18, p=.019). Pairwise comparisons (figure 4) revealed that the animated condition significantly increased visual attention in comparison to the static condition for all stories (Cycling with Grandpa, p=.015, d = 0.54; Little Kangaroo, p<.001, d = 1.01; Imitators, p=.031, d=0.48), and also significantly increased visual attention in comparison to the sound-only condition for Little Kangaroo (p<.001, d=0.83), and for Imitators (p=.008, d=0.58).

Contributions

- The findings will present at OER Monday Talk series.
- One paper has been accepted by *AERA open* and another paper is under review by *Frontier in Psychology*

- Bring MOE insight whether e-storybook should be promoted for enhancing emerging literacy development
- Inform the parents in Singapore how to use ebook on pre-schoolers

Conclusion

The current study confirmed previous findings that animated eBooks might facilitate children's productive vocabulary learning and attract better attention from them via listening to the stories. We have extended the beneficial scope from Germanic languages to Mandarin Chinese. Animated illustrations, by adding motion, could enhance the congruity of the auditory reading of the story and the visual illustrations. This could direct children's attention to the rich details of the story, thereby enhancing story comprehension and word learning. This finding is generally in line with the multimedia learning hypothesis that children could use dual channel resources to effectively process the input as long as the verbal and nonverbal information are coherently designed.

Acknowledgements

This study was funded by the Education Research Funding Programme, National Institute of Education (NIE), Nanyang Technological University, Singapore, project no. OER 13/16 HS. The views expressed in this paper are the author's and do not necessarily represent the views of NIE.

Keywords

Animated eBook, multimedia theory, Mandarin learning, storytelling, total fixation time, motion

Animated electronic storybook and children's Mother Tongue development: Tracing the process and the outcome with Eye-tracking

Sun He

National Institute of Education

Introduction

Storybook reading is one of the most effective approaches for children to acquire novel words and grammar in a meaningful context (Weizman & Snow, 2001). Nevertheless, children with limited language knowledge, such as emergent bilinguals, may benefit less from reading activities, due to the gap between their skills and those needed for processing the story. They may fail to derive the meaning of new words from the verbal context and consequently have difficulties in figuring out the story plots (Verhallen & Bus, 2010). Well-designed animated electronic books (eBooks) hold great promise for children's emerging literacy in this case, as such books may stimulate readers' visual, auditory and even kinesthetic senses to comprehend and digest a story via the match between the animated features (motion pictures, hotspot, and sound) and the read-aloud (de Jong & Bus, 2002; 2004; Neuman, 1997; Verhallen, Bus, & de Jong, 2006). In animated eBooks, the visual elements that are typically compressed into one page of static illustration could be decomposed into several parts with highlights or zoom-in effects. These features, together with story read-aloud, may direct children's attention to the essential details of the story and provide them richer sensory information to process and retain the story, and ultimately enhance their language acquisition with repetitive readings (Bus, Takacs, Kegel, 2015; Verhallen & Bus, 2009). The goal of this study was to identify whether animated features enhanced bilingual children's Mandarin reading outcomes and total fixation time in comparison to static eBook reading.

Statement of Problems

The Mechanism of Animated EBooks from the Perspective of Multimedia Theory

Animated eBooks equipped with multimedia features (such as video, sound, and music) have been found to support early language acquisition (Verhallen, Bus, & de Jong, 2006), and the mechanism

of such effects may lie in the presentation of the information via multiple channels. According to Paivio's (1986) dual-coding theory, visual and auditory information is processed in two separate but interconnected channels. When presented simultaneously in a coherent way, the brain is able to interpret the words, images, and auditory information in an inclusive manner, leading to better learning effects compared to a single-channeled manner of information delivery. The cognitive theory of multimedia learning (Mayer, 2003, 2005) explicitly proposes that deeper learning occurs when information is presented both verbally and nonverbally (Figure 1). Specifically, the enriched messages could scaffold learners to pick up the target information more easily and establish a coherent mental representation. Mayer (2009) stated that the major goal of multimedia learning is to manage essential processing (i.e., to draw learner's attention to the target information), to reduce extraneous processing (i.e., to avoid trivial and confusing information) and to foster generative processing (i.e., to trigger the progress that leads to conceptual knowledge). When illustrations and narration are able to complement each other in picture storybooks, the nonverbal information may support comprehension of verbal information and, vice versa, verbal information may support the interpretation of illustrations and other nonverbal information (Sipe, 1998).

< Insert Figure 1 here>

Conventional educational instructions rely heavily on verbal information (i.e., one system only); while animated eBooks offers children multiple verbal and nonverbal information to understand concepts that are difficult to grasp with words and static pictures alone. Building upon Mayer's theory, the theory of Synergy (Neuman, 2008) suggests that each medium's physical features, structure, and method of handling material, may add a new dimension to children's knowledge and bestow on them an approach to obtain novel knowledge. Therefore, rather than distracting from literacy, multimedia may expose children to an additional set of processing tools to interpret events, allowing them to benefit from the "redundancy effect" (Neuman, 2008), where information has been delivered with multiple channels. This could be especially true when one delivery system is "blocked" due to the learners' unfamiliarity of such a format. Given that knowledge and information are so central to children's comprehension of a story, the "redundant" information may ensure they understand the content and acquire the language within.

The potential benefits of the specific features of animated eBooks (e.g., motion and sound) might be inferred from these theories. They could provide rich verbal and nonverbal information to optimize temporal congruity of narration and pictures, facilitating children's selection of content for story processing and strengthening their story recall afterward (Bus et al., 2015; Mayer, 2001). The static illustrations in the traditional paper storybooks represent the complete event(s) on a single page, and it might be a challenge for children with little knowledge in the language to figure out which part of the illustration to focus on to form explicit associations between words and visual details in the pictures. Children may remember some visual contexts where they heard a word and acquire the novel word receptively (i.e., the ability to identify the semantic content) (Sénéchal, 1997). However, such an association might not be precise enough to produce the words orally (i.e., producing the correct word for an image; Sénéchal & Cornell, 1993). Under such a circumstance, motion in animated storybooks could be useful to direct children's attention to the specific details, to enhance their comprehension of complex expressions, and to strengthen their memory of unfamiliar words. Takacs and Bus (2016) found evidence for this hypothesis and revealed that well-designed motions in animated eBooks are able to guide children's attention to the target information and support story and language comprehension. Sound (e.g., onomatopoeia and background music) in animated eBooks may lead to similar effects, though the current findings are inconclusive. Sounds such as crying and humming would concretize scenes and word meaning, adding more information for children to understand the story (Schnotz & Rasch, 2005). Background music, on the other hand, would highlight a character's mood (e.g., anger or happiness) and the tone of the story, and scaffold children to comprehend the text. However, it is worth noting that sound might disrupt the perception of speech for children who have difficulties with verbal processing (Smeets, van Dijken, & Bus, 2014), leading to poorer outcomes.

A series of studies have demonstrated that the mentioned features in animated eBooks (i.e., motion and sound) may be beneficial to child bilingual's early vocabulary acquisition (e.g., Verhallen & Bus, 2009), phonological awareness (e.g., Van der Kooy-Hofland et al., 2011), and grammar development (e.g., Smeets & Bus, 2015). They may also facilitate children's story comprehension, better awareness of the goals, motivations, and emotions of the story figures (Verhallen et al., 2006). For instance, Takacs and Bus (2016) followed 4-6-year-old children's visual attention with eye tracking and

found that children in the animated eBook group recalled significantly more story language with the help of motion-powered illustrations than their peers in the static eBook group. Although there are increasing numbers of studies on the efficacy of animated e-storybooks in recent years, most of them have been conducted in western countries, mainly focusing on Germanic languages like English or Dutch, and few have paid attention to other languages in different contexts (e.g., Chinese learning in Asia). The current study aims to extend the scope of this area by focusing on the effect of animated e-storybooks on bilingual preschoolers' Mandarin learning in Singapore.

Preschooler's Mandarin Learning in Singapore

Singapore is a multilingual society with three main ethnic groups (Chinese 74.3%, Malays 13.3%, and Indians 9.1%) and four official languages (English, Mandarin, Malay, and Tamil) (Singapore Department of Statistics, 2016). English is the language of inter-ethnic communication of education, government, and commerce while the other three official languages (i.e., Mother Tongue languages, MTL) are for ethnic identity and heritage maintenance. Although children are encouraged to develop their MTL and English simultaneously (Ministry of Education, 2013), recent years have witnessed a discrepancy between English and ethnic languages in both learning outcomes and environment. By following 805 Singaporean preschoolers, aged four to five years, Sun, Yin, Amsah, and O'Brien (2018a) found that children's ethnic language vocabulary size was substantially smaller than their English vocabulary size across all the three ethnic groups, and their input quantity and quality of ethnic language was significantly lower than that for English learning. As Cavallaro and Ng described (2014, p.36), "English is increasingly becoming the mother tongue for more and more Singaporeans, and their ethnic languages are technically more like second languages."

Against this social backdrop, research on animated e-storybooks is especially relevant because of their promising effects on young second language learners' emergent literacy skills. According to the Singapore National Library Board, the number of electronic books borrowed has reached eleven million in 2015, increasing almost four times since 2009, and such a reading format is favoured by children because of its entertaining elements (e.g., sound and interactive games) (Hio, 2015). Despite the changing landscape of children's reading format, little is known how these eBooks may influence bilingual language learning in Singapore, and whether children can ultimately benefit from such exposure. To

strengthen Singapore children's Mandarin language development as emphasized by the Ministry of Education Singapore, an investigation into approaches that motivate our children to better master the Mandarin language seems more essential than ever. This study explored the efficacy of animated e-storybooks on pre-schoolers' Chinese language development and acquisition. With this aim, we compared children's vocabulary acquisition and reading comprehension in four conditions: (1) animated stories with motion and sound, (2) corresponding static stories, with sound and soft-copy printed illustrations, (3) corresponding static stories with soft-copy illustration only, and (4) a control condition (no reading exposure). In addition, we explore the relation between features of animated e-storybooks (i.e., motion and sound) and preschoolers' visual attention with eye trackers. It is becoming increasingly common to use eyetracking to investigate visual attention. Eye trackers measure gaze direction and saccadic eye movements by measuring infra-red light reflected off the surface of the eye. This non-invasive technology is quick and easy to administer to children of all ages. Eye movement is a direct measure of overt visual attention (Kulke, Atkinson, & Braddick, 2016), and has been used to examine novel word learning (Mather & Plunkett, 2012), where gaze duration is increased when novel objects are paired with novel names, and eBook reading in Germanic languages (Takacs & Bus, 2016). In the context of eBook reading, eyetracking allows the possibility of analyzing whether attention is allocated to the presented material, or to other distracting stimuli in the environment through fixation duration (Wass, Smith, & Johnson, 2012).

Purpose of Study (including Research Questions and/or Objectives)

Two questions were addressed in the current study:

Question 1: Do children from the animated eBook group outperform their counterparts from the static eBook groups in terms of vocabulary acquisition and story retelling?

Hypothesis 1: Children from the animated eBook group would score higher than those from static eBook groups with the help of the animated features. Motion and sound are assumed to be valuable additions to support vocabulary learning and story understanding, being in line with the arguments proposed by the cognitive theory of multimedia learning that deeper learning occurs when a coherent message is demonstrated via both verbal and nonverbal channels. Children's vocabulary acquisition has been operationalized as productive vocabulary, receptive vocabulary, context integration, and meaning

recognition. As children's general language proficiency and cognitive status would affect vocabulary acquisition from the intervention, children's general Mandarin proficiency (indicated by receptive vocabulary size, receptive grammar knowledge, and verbal fluency in Mandarin) and cognitive factors (nonverbal intelligence and phonological short-term memory) have been controlled.

Question 2: Compared with children in the static eBook groups, would children in the animated eBook condition show longer eye-fixations time?

Hypothesis 2: Children from the animated eBook group are expected to outperform their counterparts from the static eBook groups in terms of total duration of eye fixations to the eBook. Similar to the first hypothesis, motion and sound are assumed to optimize temporal congruity of narration and pictures, and eventually guide and maintain children's attention to the target information for story comprehension.

The findings of this study may not only benefit young learners in the Singapore context but also Mandarin learners worldwide. The Ministry of Education of the People's Republic of China estimates that over 40 million people outside China are presently learning Mandarin and that the number is growing annually (Student Travel Planning Guide, 2014). However, even with such rapid increase in Mandarin language learners worldwide, studies on teaching Chinese to second or foreign language speakers are still limited (Duff et al., 2013; Han, 2014) and the limited empirical studies have largely focused on adult learners in the western context (Ruan, Zhang, & Leung 2016).

Participants

The current experiment recruited 129 kindergarten-1 children at the age of four to five years old from 21 preschools. The parents of each child were asked basic information about their children's Mandarin learning at home and at school. Three criteria were included for participant selection in the current study. First, children needed to be English-Mandarin emergent bilinguals and those who had recently migrated from China were excluded. Secondly, participants needed to have no history of developmental delays or impairment, based on parental report and teacher's observations. Thirdly, participants had to complete most of the readings and the outcome assessments. Among the 129 children, 27 children were eventually removed from the final analysis due to sickness/holiday leave, recent immigration, or atypical language development. The final sample comprised of 102 children, 49 boys and 53 girls. The socio-economic status varied among children, but most of them were from middle-class families. On average, parents

possessed a polytechnic or bachelor's degree as the highest degree (e.g., mother's education; M=5.34, SD=1.28, range=2-8, ranking from "No qualification" to "Doctorate degree"), with approximately S$7500 to S$7999 family income per month (M=15.01, SD=5.07, range=0-19, with S$500 increment for each higher level).

Methodology/Design

The study was a randomized control trial (RCT) including four experimental conditions, we estimated the optimal sample size ($n=1.962\sigma2/E2$ where σ = standard deviation and E = error rate):

• Static e-books, with neither sound nor motion (N=28): Children independently read three e-books four times spread over two weeks, in which static illustrations were accompanied by only an oral narration of the story.

• Static e-books, with sound only (N=29): Children independently read three e-books four times spread over two weeks, in which static illustrations were accompanied by background music, motion sounds and oral narration of the story.

• Animated e-books (N = 32): Children independently read three animated versions four times spread over two weeks, in which animated illustrations (including motion, music and sound) accompanied an oral narration of the story.

• Control condition (N = 13): children independently played non-literacy related math games four times spread over two weeks during the intervention period.

Experiment procedure

Children were pre-tested for their general Chinese receptive and productive vocabulary size and the target vocabulary in the storybooks. As indicators of general cognitive proficiency, their nonverbal intelligence, and phonological short-term memory were also tested with standardized tests.

The intervention of storybook reading will take 4 sessions over 2 weeks (two sessions per week). Children read 3 stories per session with randomized order in a quiet room individually during their free reading time in the afternoon. Two research assistants were present to operate the eye-tracking machine (e.g., calibration) and to instruct the children to read the story. No interaction were allowed once children start to read the storybook, other than reminders to pay attention to the stories. Children in the control group spent the same amount of time on screen as those in experimental groups.

Within one week after the last treatment session, children were tested again for the target words (post-test). This also included tests of sentences repetition and story retelling. Test and survey for participants' selection

Demographic questionnaire

A parental questionnaire was used to estimate children's language background information (Appendix 1). The questionnaire was designed based on the Language Exposure Questionnaire (Sun, Steinkrauss, Tenderio & de Bot, 2016) and the Utrecht Bilingual Language Exposure Calculator (Unsworth, 2013). In contrast to these questionnaires, the current one explored children's e-storybook preference and reading history in added depth.

General Chinese proficiency

Bilingual Language Assessment Battery (BLAB; Sze & Rickard Liow, 2009) was used to estimate children's general Chinese proficiency. The receptive vocabulary component is a standardized auditory-picture matching task. Norms had been developed based on children of preschool level in Singapore.

Phonological short-term memory

Digit span and non-word repetition, the sub-tests of the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen & Rashotte, 1999) was administered to assess the children's short-term phonological memory. The two tests composed of a list of digits or non-words, which increases in length.

Nonverbal intelligence

Raven's Colored Matrices (Raven, Raven & Court, 1998, sets A, B and AB) was used to estimate children's analytical reasoning, which is assumed to be an important component of language aptitude (Paradis, 2011). Children were required to choose the missing part of a presented pattern from 6 options. In total, there were 36 items that increased in complexity.

Materials

Storybooks

Three prize winning child stories, Little Kangaroo (Genechten, 2007), Imitators (Veldkamp, 2006), and Cycling With Grandpa (Boonen, 2004) were used as reading materials. These books had been translated into many languages, including Mandarin. They have been successfully adapted to animated

e-storybooks for educational purpose in Dutch and used in several language acquisition experiments with children with limited language proficiency (e.g., Smeets & Bus, 2015). The static versions consisted of pictures from the original print storybooks and were presented on screen. The story text was automatically read aloud and the story continued without any action. In the animated versions, the static illustrations representing the story events were dramatized (congruent with the story) by using motion (e.g., to run), sound (e.g., the sound of bell) and background music (e.g., low-spirit music to reflect character's frustration). Each reading session included three stories and lasted about 15 minutes including the preparation time. Children in the control group played math games on the computer during the intervention sessions.

Outcome tests

18 words were selected from each story for reception and production tests. They were tested at pre-test at post-test in random order.

Story sentence repetition is considered an indicator of grammatical understanding. 15 sentences were selected from the three stories based on the following criteria: each sentence should include five or more words, and these sentences composed increase in difficulty. They were applied at post-tests in a random order. Story recall were conducted directly after the first and fourth experimental sessions. The e-storybooks in static versions were shown on screen page by page and children were asked to retell each story based on their understanding and memory.

Eye-tracking system

The Tobii X3-120, a remote eye-tracker, was used to measure the length and frequency of eye fixation while the recorded voice narrated the stories to children. With a sampling rate of 120, the eye-tracker could get an observation of eye movement approximately every 8 ms. Since children moved around quite often, it was hard to let them wear a head-mounted eye-tracker. Instead, the Tobii eye-trackers was mounted to the base on the laptops, which should result in more accurate measurements. In order to ensure the optimal registration of eye movements, children were asked to be seated at a distance of 60-70 cm from the eye tracker, and the machine was calibrated for each child at the beginning of each reading session by asking them to fixate five dots shown on the screen. No additional effort was needed by children during the procedure and the preparation took approximately 2 minutes. Fixation time

on the target words in a storybook was divided by the total time of visual attention for the whole narration because the three books are varied in length. We specifed the pictoral location of the target words as Areas of Interest (AOI) with the software of Tobii studio. The duration of children fixating the AOI and the frequency of such fixations was recorded by Tobii Studio's fixation filter with the default settings from velocity and distance threshold. In the current study, we were particularly interested in longer fixations (i.e., gaze duration longer than 200 ms) because previous literature revealed that it takes such amount of time to process complex visual information as those in e-storybooks (Henderson, 2006, Watson, Gunlogson & Tanenhaus, 2006).

Analysis

The first and second research question would be analyzed with mixed effects modelling. Researchers conducting longitudinal studies might encounter difficulties such as within-group variation, small sample sizes and participant attrition. In the current study, preschoolers might vary in affective factors, such as the level of language anxiety, which would be beyond our research focus. Moreover, some children might miss a reading session due to miscellaneous reasons (such as illness) resulting in an imbalanced dataset. Compared with traditional approaches such as ANOVA, LMER is more useful to control for these aspects. It is a powerful statistical technique that is robust to outliers and missing values and is able to take both by-item and by-participant variability into account, thereby yielding generalizable results (Baayen, 2008, Ch. 7). The random-effect factor in our study are participant, book and location of AOI in the picture. Blom, Paradis and Duncan (2012) explain random effects as follows: "the effect of the participating children is typically considered a random-effect variable. This is because the sample of children is drawn from a larger population, and each participating child has unknown properties that will influence the measurements. Hence, the effect of participating children cannot be measured without error and if a new study is undertaken, other participants will be included, with again unknown properties" (p. 978). The same holds true in the case of random effects brought by the choice of e-storybook and location of AOI. By including random-effect factors, type-I errors are prevented (Baayen, 2008).
The third question concerning children's visual attention and the occurrence of the special features (i.e., motion or background music) will be analysed with CRQA, which is a particular type of non-linear time series analysis to assess the attunement of two interacting systems (Shockley et al., 2002). This

approach has proven to be effective in the study of the coordination (i.e., shared dynamics) of verbal and nonverbal behavior in dyadic interactions, e.g., during a conversation (Reuzel et al., 2013; 2014). In such cases, recurrence reflects that the behavioral state of one component (e.g., activation of background music) is matched by the other component (e.g., activation of long visual attention) across all possible time scales, from the sample rate to the duration of the observation. That is, these matches are not only detected if they occur at the same time, but also earlier or later in time (for more details on this approach, see Dale and Spivey, 2006 and Reuzel et al., 2013; 2014).

Findings / Results

Analysis

The three Mandarin proficiency scores (verbal fluency, receptive vocabulary, and receptive grammar) were combined using factor analysis. This gave a simplified score of general Mandarin proficiency for each child. A mixed effects model was used for data analysis, as this statistical technique is robust to outliers and missing values, being able to take both by-item and by-participant variability into account, thereby yielding generalizable results (Baayen, 2008, Ch. 7). The random-effect factors in our study are children's class and school. By including random-effect factors, type-I errors are prevented (Baayen, 2008).

Vocabulary tests. Results were analyzed in R (R Development Core Team, 2010) as mixed models, using the lme4 package (Bates, et al., 2011). For each vocabulary outcome variable, a model was created including the fixed effects PreTotal (total vocabulary score in pretest), Group (animated vs sound only vs static vs control), PhoMem (phonological short-term memory score), NonInte (nonverbal intelligence score) and Man.EFA (General Mandarin proficiency). Class and School IDs were included as interacting random effects. Models were compared to simplified models (including only PreTotal main effect and Class / School ID random effects) using Akaike Information Criterion (AIC) scores, where they were only used for further analysis if the AIC score was a minimum of 2 points lower. These models were then recalculated, using scaled and centred values for the fixed effects PreTotal, PhoMem, NonInte, and Man.EFA, and the Restricted Maximum Likelihood (REML) procedure. Standard ANOVA tables were produced using the lmerTest package (Kuznetsova et al, 2017), using Satterthwaite's method for

denominator degrees of freedom and F statistics to estimate p values. A summary of all F statistics is presented in Table 3. Finally, to explore question 1, pairwise differences of estimated marginal means for all levels of Group (animated, sound-only, static, control) were calculated (see Appendix E for full pairwise tests).

Story retelling. Total story retelling scores for each story were again analysed using mixed models. For each story, a model was created including the fixed effects Group, Assessment (Tel1: after the first reading, Tel2: after the fourth reading), Man.EFA, PhoMem, NonInte, and the interaction of Group x Assessment. Class and School IDs were again included as interacting random effects using the REML procedure. Pairwise differences of estimated marginal means for all levels of Group (animated, sound-only, static, control) were calculated (see Appendix F) to examine final retelling ability across conditions. To examine whether development rate of retelling ability was affected by the eBook conditions, pairwise difference-of-differences (assessment Tel1 vs Tel2) of estimated marginal means for all levels of group were calculated (Appendix G).

Visual attention. Total fixation time was recorded for each child for each story at the four readings. The four readings were averaged to give a mean total fixation time per child for each story, and the mean fixation was then divided by the total length of specific story, yielding a percentage score for children's attention for each book. Again, mixed models were used. For each story, a model was created including the fixed effects of Group and (scaled and centred) Man.EFA. Class and School IDs were included as interacting random effects, using the REML procedure. To explore question 2, pairwise differences of estimated marginal means for the eBook conditions within Group (animated, sound-only, stati***C) WERE CALCULATED (APPENDIX G)***

<Insert Table 3 here>

Productive Vocabulary

There was a significant effect of Group on productive vocabulary ($F(3,68.03) = 4.05$, $p=.010$). Pairwise comparisons (figure 2a) revealed that the animated eBook condition resulted in significantly higher productive vocabulary scores than the sound-only ($p=.021$, $d=0.64$), static ($p=.015$, $d=0.69$), and control ($p=.005$, $d=1.23$) conditions. In comparison, there were no significant differences between the

sound-only, static, and control conditions. Higher total pre-test vocabulary scores (p<.001) and Mandarin proficiency (p<.001) also significantly increased productive vocabulary.

Receptive Vocabulary

Mandarin proficiency again had an effect, with higher scores resulting in significantly higher receptive vocabulary (p=.017). There was also a significant effect of Group (F(3,68.90) = 3.62, p =.017). Pairwise comparisons (figure 2b) revealed no significant differences between the eBook types, but significantly higher receptive vocabulary scores in all eBook conditions in comparison to control (animated, p=.002, d =1.31; sound-only, p=.031, d=0.90; static, p=.018, d=1.01).

Context Integration

There was a significant effect of group on context integration (F(3,59.90) = 3.03, p = .036). Pairwise comparisons (figure 2c) revealed that the animated eBook condition significantly improved context integration in comparison to the static (p=.011, d=0.70) and control (p=.035, d=0.85) conditions, but not in comparison to the sound-only condition. Higher total pre-test vocabulary scores (p=.001) and Mandarin proficiency (p=.004) also significantly increased context integration.

Meaning Recognition

Higher Mandarin proficiency again had a significant effect (p< .001), giving higher meaning recognition scores. There was also a significant effect of group on meaning recognition, (F(3,93.00) = 7.65, p< .001). Pairwise comparisons (figure 2d) revealed no significant differences between the eBook types, but significantly higher meaning recognition in all eBook conditions in comparison to control (animated, p<.001, d=1.59; sound-only, p<.001, d =1.28; static, p<.001, d=1.27).

<Insert Figure 2 here>

Story Retelling

There were significant effects of group, assessment, and general Mandarin proficiency for all three stories (table 4). Children were better at retelling in assessment 2 and those with higher Mandarin proficiency performed better. For Little Kangaroo, there was also a significant interaction of group x assessment (F(3,151,78) = 4.29, p=.006). Post-hoc pairwise contrasts were used to explore the effect of eBook condition on the retelling of each story (figure 3). For Cycling with Grandpa, the animated condition significantly improved retelling ability in comparison to all other conditions (sound-only, p=.047, d=0.55;

static, p=.029, d=0.63; control, p < .001, d=3.26). For Little Kangaroo, there were no differences between the animated, sound, or static conditions, but all were significantly higher than control (animated, p< .001, d=3.30; sound-only, p< .001, d=3.29; static, p<.001, d=3.03). Similarly, for Imitators, there were only significant differences between the eBook conditions and control (animated, p<.001, d=2.60; sound-only, p<.001, d=2.56; static, p<.001, d=2.25).

When examining development rate, for Little Kangaroo, all eBook conditions showed significantly greater assessment 1 to assessment 2 improvement than control (animated, p<.001; sound-only, p=.004; static, p=.022). However, for the other two stories, only the animated (Cycling with Grandpa, p=.025; Imitators, p=.024) and sound (Cycling with Grandpa, p=.014; Imitators, p=.016) conditions showed greater improvement than control.

<Insert Table 4 here>

<Insert Figure 3 here>

Visual Attention

ANOVA summaries for the three stories are in table 5. There was a significant effect of group for Cycling with Grandpa (F(2,85)=3.41, p=.038), for Little Kangaroo (F(2,82.34)=12.89, p<.001), and for Imitators, (F(2,85)=4.18, p=.019). Pairwise comparisons (figure 4) revealed that the animated condition significantly increased visual attention in comparison to the static condition for all stories (Cycling with Grandpa, p=.015, d = 0.54; Little Kangaroo, p<.001, d = 1.01; Imitators, p=.031, d=0.48), and also significantly increased visual attention in comparison to the sound-only condition for Little Kangaroo (p<.001, d=0.83), and for Imitators (p=.008, d=0.58).

Contributions of study

In the current study, we explored the effects of widely used animated features in eBooks on children's vocabulary acquisition, story retelling, and total fixation time during Mandarin storybook reading. We used a combination of pre- and post-test vocabulary questions and fixation duration, as measured by eyetracking to examine whether these animated and acoustic features could enhance bilingual children's reading outcomes. Children's performances in the animated eBook condition (i.e., illustrations with sound and motion) were compared with that of their peers in the static eBook condition (i.e., illustrations neither sound nor motion). As previous findings on the effect of sound are inconclusive,

we added a sound-powered static eBook condition (i.e., illustrations with sound but no motion) to investigate its independent contribution to children's learning in the eBook reading. The animated eBook condition improved productive vocabulary, context integration, and visual attention as measured by gaze duration, as well as improving story retelling for Cycling with Grandpa. Our findings are in line with the previous studies in Germanic languages (Verhallen et al., 2006; Takacs, Swart, & Bus, 2015; Takacs & Bus, 2016) that animated eBooks could facilitate children's emergent language development. As our study focuses on Mandarin, which is one of the "notorious" languages that is difficult to learn as a second language (Moser, 1991; MacBride, 2016), our positive findings from animated books implies such a reading format might be useful for child bilingual learners to acquire this tonal language. This is particularly timely, given that computer and internet use in this age group is rapidly increasing. In 2018, 77% of Singapore residents under the age of seven were internet users, up from 40% in 2016 (IMDA, 2018). In terms of the specific features, when motion and sound were used together (animated condition), their effects on children's vocabulary learning and story comprehension were significantly higher than when sound was used alone or none of the features were employed. We did not find a consistent effect of the addition of sound on learning (sound-only condition, -motion +sound) in comparison to a purely static eBook condition (static condition, -motion -sound). This suggests a lack of utility in sound-enriched eBooks. However, our research design forbids us to conclude whether it is the interaction between motion and sound or purely motion alone that benefits children's Mandarin acquisition. The findings by Takacs and Bus (2016) on motion in the illustrations suggest that motion could promote story comprehension independently, with a medium effect size ($\eta 2 = 0.14$) similar to the advantage found for multimedia stories in comparison to print stories. Future studies should aim to examine the effects of motion and sound independently, so that these features can be separated. Such findings would help to "create clearer guidelines for designers of multimedia stories" (Takacs & Bus, 2016, p. 9).

This study broadly supports Neuman's (2008) theory of Synergy, with some caveats. Despite the favorable effects in general, the animated eBook worked on various aspects of language learning differently. The children in the animated eBook condition were found to obtain more target words productively (to produce the target words with the illustrations from the storybooks or in a novel setting; i.e., "productive vocabulary" and "context integration") but not receptively (to comprehend the target

words with the illustrations from the storybooks or in a novel setting; i.e., “receptive vocabulary” and “meaning recognition”). This keeps in line with Sénéchal’s arguments that the static illustrations in the storybooks may facilitate children’s word memory as the visual cues to identify the semantic content, nevertheless, to establish an explicit association between the visual image and the target word, the learners may need additional scaffolding like motion to produce the word (Sénéchal, 1997; Sénéchal & Cornell, 1993). Synergy theory would suggest that all aspects of language learning would improve as more redundant information was added. Our lack of evidence for improvement in the sound-only condition suggests that redundant visual information may be more useful than redundant sound. On average, children’s production and comprehension of the target words are better in the condition with the illustrations than in the novel settings, indicating that transfer of the semantic knowledge requires more usage of the words in various contexts (Nagy & Scott, 2000). The effect of the animation also worked differently on the retellings of the three books. Children produced significantly more details in the animated condition than their counterparts in the story retell after the fourth reading of the book Cycling with Grandpa. There were no significant differences in the retells of the other two books, Little Kangaroo, and Imitators. We attribute the discrepancy to the extent of language complexity of the stories and propose that the benefits of the animated eBook reading might be more distinguishable when children need to process a larger amount of linguistic input. The specific animated features might assist them to focus on the main plots while accumulating the details to enrich their understanding. Further studies are needed to verify this hypothesis.

Children in the animated condition showed longer total fixation time (in percentage) across the repetitive readings than their peers in the static groups (sound-only and static conditions). It might be that the animated features attracted children to explore the illustrations in detail. In other words, watching and listening to an animated storybook might be more engaging than reading its static version, leading to longer lingering on the illustration and higher motivation to explore the details, and resulting in better word learning and story comprehension. The animated feature “motion” might cause a longer total fixation time in particular. Previous studies demonstrate that there is a close link between current fixation and spatial attention (Liversedge & Findlay, 2001), and a combination of exogenous and endogenous factors could affect the selection of the target for fixation. Regions of space that are distinct from the rest because of

motion may draw people's attention in a quick and automatic manner (Franconeri & Simons 2003). Results from Takacs' and Bus' study (2016) suggest that motion could change children's processing strategy. When children were provided with motion powered illustrations, their average fixations were longer, despite the contents being the same in the animated eBook condition and in the static eBook condition. Their fixations were also found to be steadier as they moved less between various visual elements in the animated condition. Such longer and steadier attention suggests in-depth processing of the essential details (Rayner, 2009) and integration of the story information, which might lead to better learning results. Similarly, in our study, the animated eBook condition increased fixation duration in comparison to the static or sound-only conditions, suggesting that our participants were engaging in a different processing strategy.

Limitations and Implications

The current studies have three limitations. First of all, it remains unclear whether it is motion per se or the interaction between motion and sound in the animated condition that made children have better learning outcomes and longer total fixation time. Future studies may unravel the specific characteristics of the animated eBook to explain the learning effects. Secondly, a within-subject design would be better as children's general Mandarin proficiency and children's initial scores of the target words were found to be significant control variables for most of the outcomes. Furthermore, the timing of data collection should be considered. In our case, we have conducted the experiments in the last quarter of the school year (October to December) and some children missed sessions due to holiday leave. Future studies should have a better plan to avoid attrition. Finally, yet importantly, more book genres should be included in the experiment. The current study focused on fiction, and future studies should consider exploring animated features in nonfiction books to examine whether features like motion would function equally well in different types of reading. Despite the limitations, the current findings are considered useful to child Chinese language learners. Keeping in line with the previous studies (e.g., Verhallen & Bus, 2011), animated eBooks have been found to be beneficial to children's second language learning, compared to the effects of the static versions. The animated features, motion in particular, may draw children's longer fixation time to digest the details of the illustration and scaffold their story comprehension and novel vocabulary learning. It implies that well-designed animation would create congruency between the

narratives and animated illustrations, and eventually promote children's second language learning. This tip should be recommended to eBook developers, parents, educators, and policy makers.

Conclusion

The current study confirmed previous findings that animated eBooks might facilitate children's productive vocabulary learning and attract better attention from them via listening to the stories. We have extended the beneficial scope from Germanic languages to Mandarin Chinese. Animated illustrations, by adding motion, could enhance the congruity of the auditory reading of the story and the visual illustrations. This could direct children's attention to the rich details of the story, thereby enhancing story comprehension and word learning. This finding is generally in line with the multimedia learning hypothesis that children could use dual channel resources to effectively process the input as long as the verbal and nonverbal information are coherently designed.

Acknowledgements

The project is sponsored by OER 13/16 HS. The author would like to thank Jieying Loh, Shi Lin Au Yong, and Bitong Bai for data collection and coding. The author would also express thanks to Jieying Loh for literature review for the manuscript.

References

Baayen, R. H. (2008). Analyzing linguistic data: A practical introduction to statistics using R. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511801686

Bak, X. Y. (2012) Mandarin Receptive Grammar Test for children: An analysis of aspect, connective and passive grammatical markers (Unpublished master dissertation). National University of Singapore, Singapore.

Bates, D., Maechler, M., & Bolker, B. (2011). lme4: Linear mixed-effects models using S4 Classes (R package version 0.999375-38). Retrieved from http://CRAN.R-project.org/package=lme4.

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. Journal of Educational Psychology, 98(1), 44. http://dx.doi.org/10.1037/0022-0663.98.1.44

Bishop, D. (1982). TROG: Test of receptive grammar. Abingdon, Oxon: Thomas Leach (for Medical Research Council).

Boonen, S. (2004). Met opa op de fiets [Cycling with grandpa]. Amsterdam: Clavis.

Bus, A. G., Takacs, Z. K., & Kegel, C. A. (2015). Affordances and limitations of electronic storybooks for young children's emergent literacy. Developmental Review, 35, 79-97. https://doi.org/10.1016/j.dr.2014.12.004

Cavallaro, F., & Ng, B. C. (2014). Language in Singapore: From multilingualism to English plus. In J. Hajek & Y. Slaughter (Eds.), Challenging the monolingual mindset, 156, 33-48. https://doi.org/10.21832/9781783092529-005.

Clark, E.V. (1993). The lexicon in acquisition. New York, NY: Cambridge University Press.

De Jong, M. T., & Bus, A. G. (2002). Quality of book-reading matters for emergent readers: An experiment with the same book in a regular or electronic format. Journal of Educational Psychology, 94(1), 145. https://doi.org/10.1037/0022-0663.94.1.145

De Jong, M. T., & Bus, A. G. (2004). The efficacy of electronic books in fostering kindergarten children's emergent story understanding. Reading Research Quarterly, 39(4), 378-393. https://doi.org/10.1598/rrq.39.4.2

De Temple, J., & Snow, C. E. (2003). Learning words from books. In A. van Kleeck, S. A. Stahl, & E. B. Bauer (Eds.), Center for Improvement of Early Reading Achievement, CIERA. On reading books to children: Parents and teachers (pp. 16-36). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Duff, P., Anderson, T., Ilnyckyj, R., Van Gaya, E., Wang, R., & Yates, E. (2013). Learning Chinese: Linguistic, sociocultural, and narrative perspectives. Berlin/ Boston: DeGruyter.

Dunn, L.M., & Dunn, D.M. (2007). The Peabody Picture Vocabulary Test (4th edition). Bloomington, MN: NCS Pearson, Inc.

Evans, M. A., & Saint-Aubin, J. (2013). Vocabulary acquisition without adult explanations in repeated shared book reading: An eye movement study. Journal of Educational Psychology, 105(3), 596–608. https://doi-org.libproxy.nie.edu.sg/10.1037/a0032465

Franconeri, S. L. & Simons, D. J. (2003). Moving and looming stimuli capture attention. Perception & Psychophysics, 65, 999–1010. https://doi.org/10.3758/bf03194829

Genechten, G. V. (2007). Kleine Kangoeroe [Little Kongaroo]. Amsterdam: Clavis.

Han, Z. (2014). Preface. In Z. Han (Ed.), Studies in second language acquisition of Chinese (pp.ix-xi). Buffalo, NY: Multilingual Matters. https://doi.org/10.21832/9781783092093

Hio, L. (2015, July 13). E-books click with more library users in Singapore. The Straits Times. http://www.straitstimes.com/tech/e-books-click-with-more-library-users-in-singapore

Infocomm Media Development Authority. (2018). Annual Survey on Infocomm Usage in Households and by Individuals for 2018. Retrieved from www.imda.gov.sg

Justice, L. M., Meier, J., & Walpole, S. (2005). Learning new words from storybooks: An efficacy study with at-risk kindergartners. Language, Speech, and Hearing Services in Schools, 36(1), 17- 32. https://doi.org/10.1044/0161-1461(2005/003)

Korat, O., Shamir, A., & Segal-Drori, O. (2014). E-books as a support for young children's language and literacy: the case of Hebrew-speaking children. Early Child Development and Care, 184(7), 998-1016. https://doi.org/10.1080/03004430.2013.833195

Kulke, L. V., Atkinson, J., & Braddick, O. (2016). Neural differences between covert and overt attention studied using EEG with simultaneous remote eye tracking. Frontiers in Human Neuroscience, 10, 592. https://doi.org/10.3389/fnhum.2016.00592

Kuznetsova, A., Brockhoff, P., & Christensen, R. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software, 82(13), 1 - 26. http://dx.doi.org/10.18637/jss.v082.i13
Liversedge, S. & Findlay, J. (2001). Saccadic eye movements and cognition. Trends in Cognitive Science, 4, 6–14. https://doi.org/10.1016/S1364-6613(99)01418-7
Mather, E., & Plunkett, K. (2012). The role of novelty in early word learning. Cognitive Science, 36(7), 1157-1177. https://doi.org/10.1111/j.1551-6709.2012.01239.x
Mayer, R. (2005). Cognitive Theory of Multimedia Learning. In R. Mayer (Ed.), The Cambridge handbook of multimedia learning (pp. 31-48). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.004
Mayer, R. E. (2001). Multimedia learning. Cambridge, UK: Cambridge University Press.
Mayer, R. E. (2003). Elements of a science of e-learning. Journal of Educational Computing Research, 29(3), 297-313. https://doi.org/10.2190/yjlg-09f9-xkax-753d
Mayer, R. E. (2009). Multimedia learning (2nd edition.). New York, NY, US: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511811678
McBride, C. A. (2016). Is Chinese Special? Four Aspects of Chinese Literacy Acquisition That Might Distinguish Learning Chinese from Learning Alphabetic Orthographies. Educational Psychology Review, 28(3), 523–549. https://doi.org/10.1007/s10648-015-9318-2
Ministry of Education. (2013). Nurturing early learners – A curriculum for kindergartens in Singapore. Framework for Mother Tongue Languages. Republic of Singapore.
Moser, D. (1991). Why Chinese is so damn hard. Sino-Platonic Papers, 27, 59–70. Retrieved from http://www.pinyin.info/readings/texts/moser.html
Nagy, W., & Scott, J. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), Handbook of Reading Research (Vol.3., pp. 269–284). Mahwah, NJ: Erlbaum.
Neuman, S.B. (1997). Television as a learning environment: A theory of synergy. In J. Flood, S. Brice Heath, & D. Lapp (Eds.), Handbook of research on teaching literacy through the communicative and visual arts (pp. 15– 30). New York: Simon & Schuster.
Neuman, S. B. (2008). The case for multimedia presentations in learning: A theory of synergy. In A.G. Bus and S. B. Neuman (Eds.), Multimedia and literacy development: Improving achievement for young learners (pp. 44-56). New York, NY: Routledge
Paivio, A. (1986). Mental representations: A dual coding approach. Oxford, UK: Oxford University Press.
Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child external factors. Linguistic Approaches to Bilingualism, 1, 213–237. https://doi.org/10.1075/lab.1.3.01par
R Development Core Team (2010). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Retrieved from http: //www.R-project.org/.
Raven, J. C., Court, J. H., & Raven, J. (1995). Coloured progressive matrices manual. London: Lewis.
Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. Quarterly Journal of Experimental Psychology, 62, 1457–1506. https://doi.org/10.1080/17470210902816461
Rickard Liow, S. J., Sze, W. P and Lee L.C. (2013). Bilingual Language Assessment Battery (BLAB) Manual (Unpublished measure). Singapore: National University of Singapore, Department of Psychology and Division of Graduate Medical Studies.
Ruan, J., Zhang, J., & Leung, C. B. (2016). Chinese Language Education in the United States. Switzerland: Springer.
Schnotz, W., & Rasch, T. (2005). Enabling, facilitating, and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning. Educational Technology Research and Development, 53(3), 47. https://doi.org/10.1007/bf02504797
Schwartz, M., Moin, V., & Leikin, M. (2012). Lexical knowledge development in first and second languages: A role of early bilingual education. Bilingualism and Bilingual Education, 15(5), 549-571. doi: https://doi.org/10.1080/13670050.2011.650332
Sénéchal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. Journal of Child Language, 24(1), 123-138.
Sénéchal, M., & Cornell, E. H. (1993). Vocabulary acquisition through shared reading experiences. Reading Research Quarterly, 28(4), 360-374. https://doi.org/10.2307/747933
Singapore Department of Statistics. (2016). Population Trends 2016.

Sipe, L. R. (1998). How picture books work: A semiotically framed theory of text-picture relationships. Children's Literature in Education, 29(2), 97-108. https://doi.org/10.1023/a:1022459009182
Smeets, D. J., & Bus, A. G. (2015). The interactive animated e-book as a word learning device for kindergartners. Applied Psycholinguistics, 36(4), 899-920. https://doi.org/10.1017/s0142716413000556
Smeets, D. J., Van Dijken, M. J., & Bus, A. G. (2014). Using electronic storybooks to support word learning in children with severe language impairments. Journal of Learning Disabilities, 47(5), 435-449. https://doi.org/10.1177/0022219412467069
Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. Review of Educational Research, 56(1), 72-110. https://doi.org/10.3102/00346543056001072
Student Travel Planning Guide. (2014). Global trends in foreign language demand and proficiency. Retrieved from http://studenttravelplanningguide.com/global-trends-in-foreign-language-demand-and-proficiency/
Sulzby, E. (1985). Children's emergent reading of favorite storybooks: A developmental study. Reading Research Quarterly, 20(4), 458-481. https://doi.org/10.1598/rrq.20.4.4
Sun, H., Steinkrauss, R., Tendeiro, J., & De Bot, K. (2016). Individual differences in very young children's English acquisition in China: Internal and external factors. Bilingualism: Language and Cognition, 19(3), 550-566. https://doi.org/10.1017/s1366728915000243
Sun, H., Steinkrauss, R., Wieling, M., & De Bot, K. (2018b). Individual differences in very young Chinese children's English vocabulary breadth and semantic depth: internal and external factors. International Journal of Bilingual Education and Bilingualism, 21(4), 405-425. https://doi.org/10.1080/13670050.2016.1178706
Sun, H., Yin, B., Amsah, N. F. B. B., & O'Brien, B. A. (2018a). Differential effects of internal and external factors in early bilingual vocabulary learning: The case of Singapore. Applied Psycholinguistics, 39(2), 383-411. https://doi.org/10.1017/s014271641700039x
Takacs, Z. K., & Bus, A. G. (2016). Benefits of motion in animated storybooks for children's visual attention and story comprehension. An eye-tracking study. Frontiers in Psychology, 7, 1591. https://doi.org/10.3389/fpsyg.2016.01591
Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and Pitfalls of Multimedia and Interactive Features in Technology-Enhanced Storybooks: A Meta-Analysis. Review of Educational Research, 85(4), 698- 739. https://doi.org/10.3102/0034654314566989
Unsworth, S. (2013). Utrecht Bilingual Language Exposure Calculator (UBiLEC): Excel spreadsheet. (Unpublished material).
Van der Kooy-Hofland, V. A. C., Kegel, C. A. T., & Bus, A. G. (2011). Evidence-based computer interventions targeting phonological awareness to prevent reading problems in at-risk young students. Handbook of Early Literacy Research, 3, 214-227.
Veldkamp, T. (2006). Na-apers [Imitators]. Amsterdam: Clavis.
Verhallen, M. J., & Bus, A. G. (2009). Video storybook reading as a remedy for vocabulary deficits: outcomes and processes. Journal for Educational Research Online, 1(1), 172-196.
Verhallen, M. J., & Bus, A. G. (2010). Low-income immigrant pupils learning vocabulary through digital picture storybooks. Journal of Educational Psychology, 102(1), 54-61. https://doi.org/10.1037/a0017133
Verhallen, M. J., & Bus, A. G. (2011). Young second language learners' visual attention to illustrations in storybooks. Journal of Early Childhood Literacy, 11(4), 480-500. https://doi.org/10.1177/1468798411416785
Verhallen, M. J., Bus, A. G., & de Jong, M. T. (2006). The promise of multimedia stories for kindergarten children at risk. Journal of Educational Psychology, 98(2), 410-419. https://doi.org/10.1037/0022-0663.98.2.410
Wagner, R., Torgesen, J., & Rashotte, C. (1999). Comprehensive test of phonological processing. Austin, TX: Pro-Ed.
Wass, S.V., Smith, T.J. & Johnson, M.H. (2012). Parsing eyetracking data to provide accurate fixation duration estimates in infants and adults. Behavior Research Methods. https://doi.org/10.3758/s13428-012-0245-6

Weizman, Z. O., & Snow, C. E. (2001). Lexical output as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. Developmental Psychology, 37(2), 265-279.

Figures

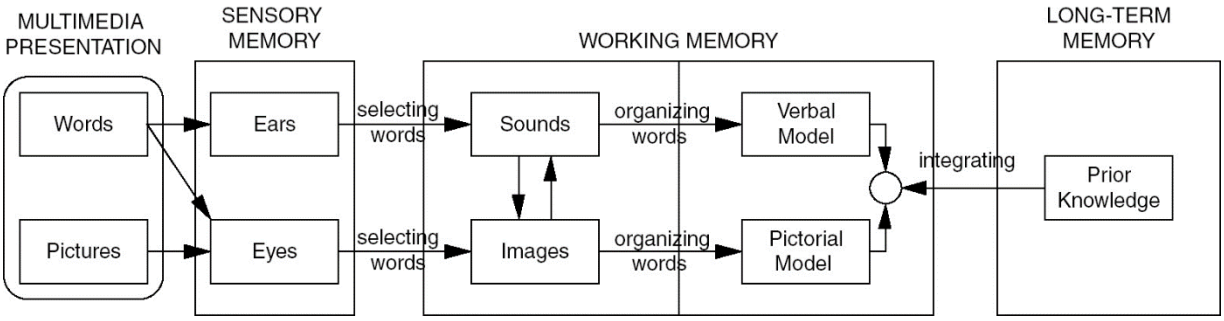


Figure 1. *Conceptual model of the cognitive theory of multimedia learning. Adapted from Mayer (2009).*

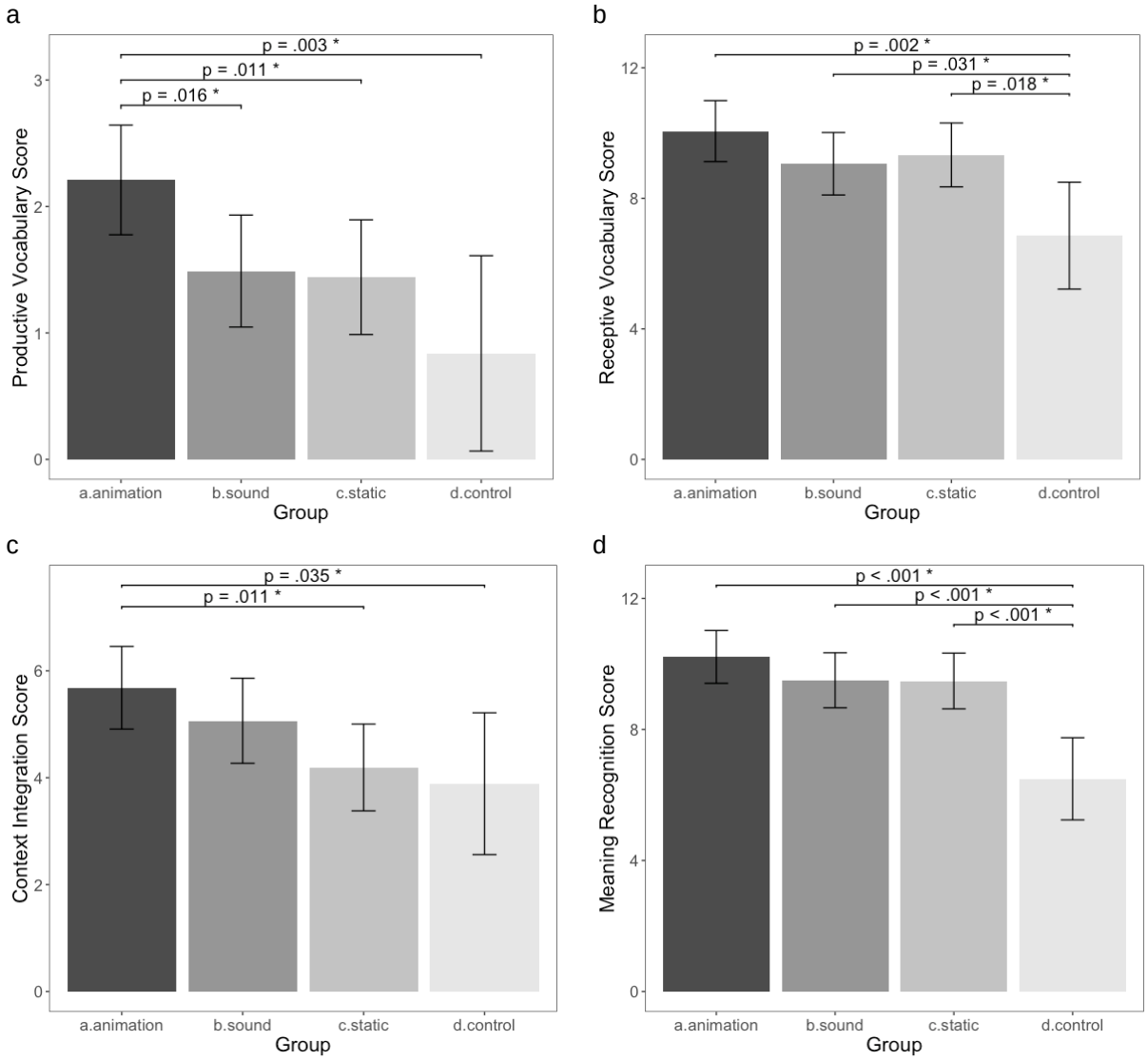
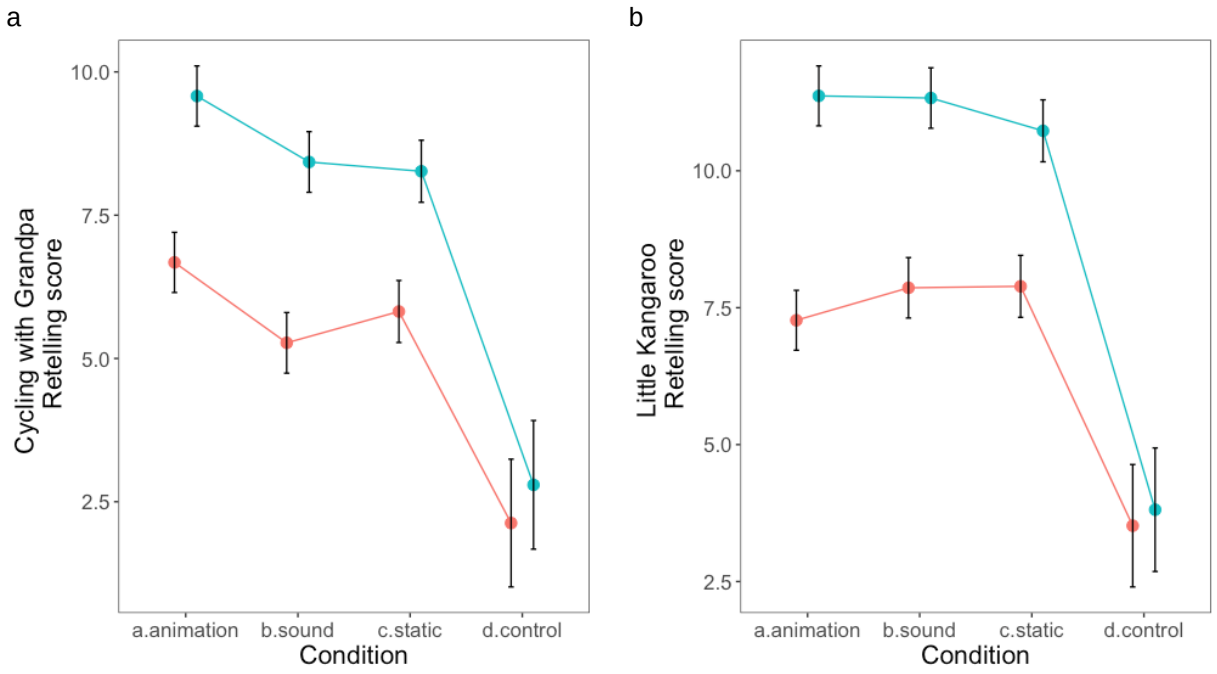


Figure 2: *Pairwise contrasts for vocabulary tests: a) productive vocabulary, b) receptive vocabulary, c) context integration and d) meaning recognition. Error bars denote SEM. Significant pairwise contrasts are highlighted.*



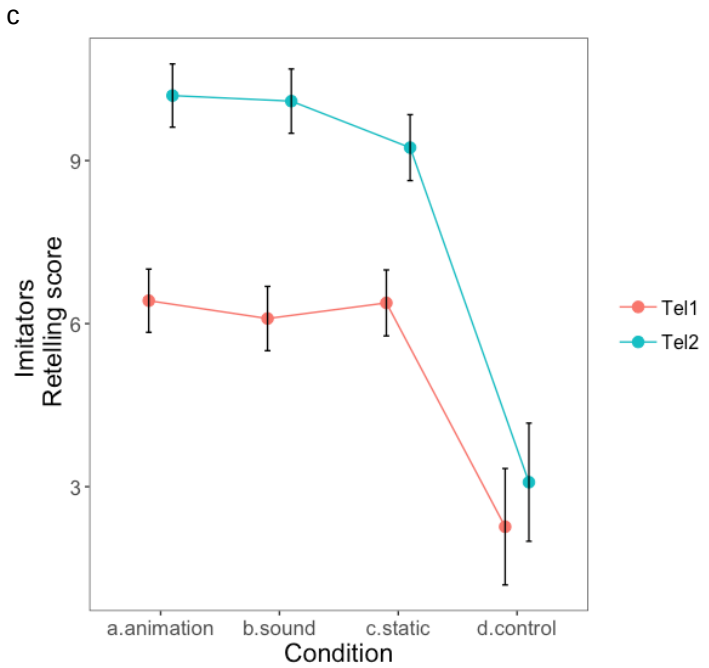


Figure 3: *Retelling scores for a) Cycling with Grandpa, b) Little Kangaroo, and c) Imitators, in Assessment Tel1 (after first reading) and Assessment Tel2 (after fourth reading). Error bars denote SEM.*

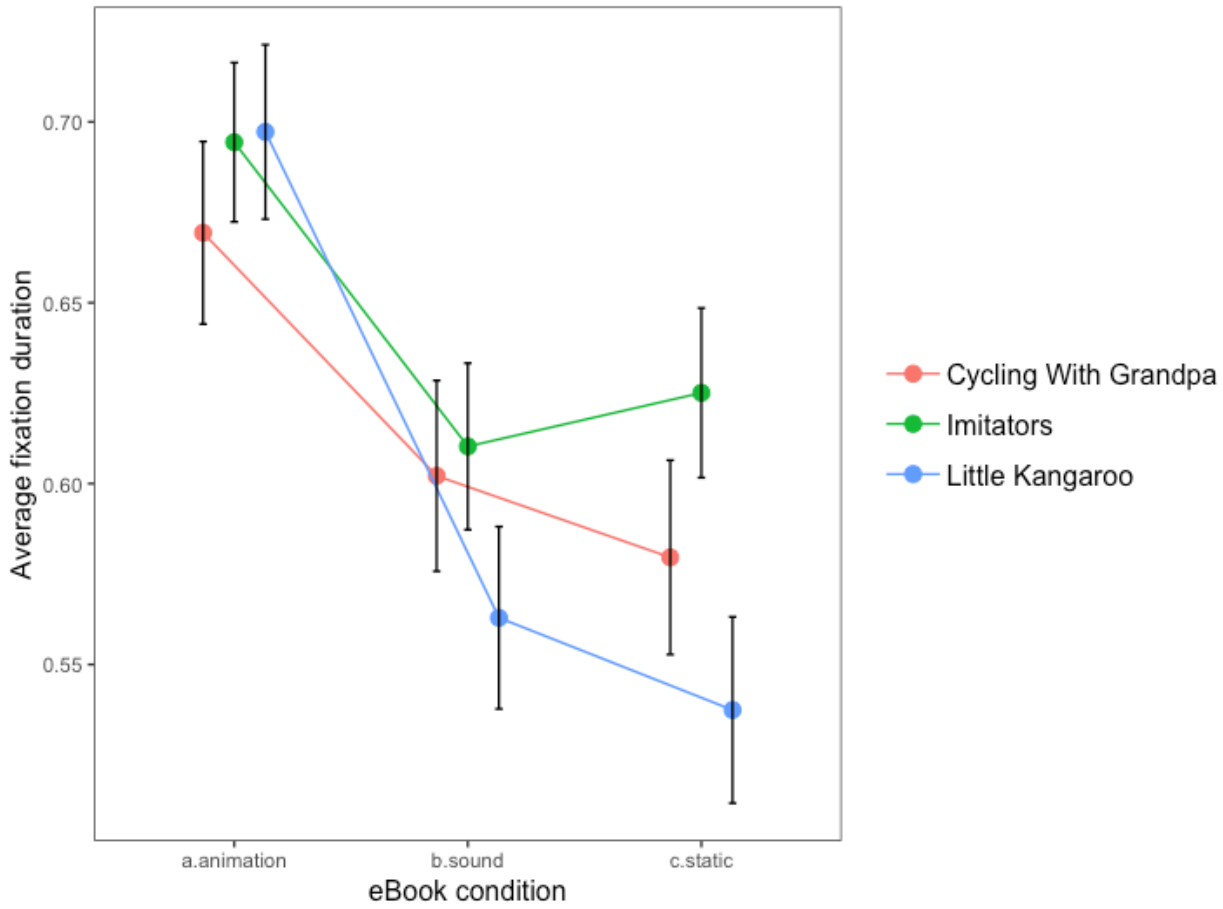


Figure 4: *Average fixation durations for each eBook (Cycling with Grandpa, Little Kangaroo, and Imitators) for the three eBook conditions. Error bars denote SEM.*

Tables

Table 1.
Experimental design and details

Condition	Description
Animation (+motion+sound)	Children independently read three animated eBooks four times spread over two weeks, in which animated illustrations (including motion and sound) accompany an oral narration of the story. There are 32 participants in this condition.
Sound-only (-motion+sound)	Children independently read three eBooks for four times spread over two weeks, in which static illustrations are accompanied with sound, accompanied by an oral narration of the story. There are 29 participants in this condition.
Static (-motion-sound)	Children independently read three eBooks for four times spread over two weeks, in which static illustrations are accompanied by an oral narration of the story only. There are 28 participants in this condition.
Control (math game)	Children independently play non-literacy related math games for four times spread over two weeks during the intervention period. There are 13 participants in this condition.

Table 2.
The procedures of the study

Stage I	Stage II	Stage III				
Parental questionnaire	Screening tests + Pretests	Reading session I + Retelling	Reading session II	Reading session III	Reading session IV+ Retelling	Posttest
Demographics & home language environment	General mandarin, cognitive, & target vocabulary tests	Story 1 Story 2 Story 3	Story 2 Story 3 Story 1	Story 1 Story 3 Story 2	Story 2 Story 1 Story 3	Target vocabulary tests

Table 3.
ANOVA summaries for vocabulary tests

	Productive vocabulary				Receptive vocabulary				Context integration				Meaning recognition			
Effect	*F*	*df*	*df(E)*	*p*	*F*	*df*	*df(E)*	*p*	*F*	*df*	*df(E)*	*p*	*F*	*df*	*df(E)*	*p*
Group	**4.05**	**3**	**68.03**	**.010**	**3.62**	**3**	**68.90**	**.017**	**3.03**	**3**	**59.90**	**.036**	**7.65**	**3**	**93.00**	**<.001**
PreTotal	**37.24**	**1**	**90.89**	**< .001**	2.77	1	91.59	.100	**11.52**	**1**	**92.26**	**.001**	0.05	1	93.00	.818
PhoMem	0.17	1	88.62	.685	1.54	1	89.72	.218	0.00	1	90.51	.955	3.42	1	93.00	.067
NonInte	0.03	1	87.40	.859	0.00	1	88.46	.981	0.57	1	88.49	.451	0.62	1	93.00	.434
Man.EFA	**21.81**	**1**	**92.18**	**< .001**	**22.14**	**1**	**92.61**	**<.001**	**8.73**	**1**	**92.99**	**.004**	**5.85**	**1**	**93.00**	**.017**

Table 4.
ANOVA summaries for retelling

	Cycling with Grandpa				Little Kangaroo				Imitators			
Effect	*F*	*df*	*df(E)*	*p*	*F*	*df*	*df(E)*	*p*	*F*	*df*	*df(E)*	*p*
Group	**9.71**	**3**	**76.58**	**<.001**	**9.41**	**3**	**69.25**	**<.001**	**10.18**	**3**	**78.38**	**<.001**
Assessment	**52.97**	**1**	**157.09**	**<.001**	**59.79**	**1**	**151.81**	**<.001**	**47.91**	**1**	**154.95**	**<.001**
Man.EFA	**56.62**	**1**	**181.45**	**<.001**	**30.97**	**1**	**183.57**	**<.001**	**59.28**	**1**	**189.98**	**<.001**
PhoMem	3.73	1	181.87	.055	0.80	1	183.61	.373	0.32	1	189.78	.575
NonInte	.080	1	180.14	.343	3.13	1	181.73	.078	0.84	1	189.12	.362
Group x Assessment	2.26	3	157.06	.084	**4.29**	**3**	**151.78**	**.006**	2.28	3	154.89	.081

Table 5.
ANOVA summaries for visual attention

	Cycling with Grandpa				Little Kangaroo				Imitators			
Effect	*F*	*df*	*df(E)*	*p*	*F*	*df*	*df(E)*	*p*	*F*	*df*	*df(E)*	*p*
Group	**3.41**	**2**	**85**	**.038**	**12.89**	**2**	**82.34**	**<.001**	**4.18**	**2**	**85**	**.019**
Man.EFA	3.15	1	85	.079	2.07	1	84.99	.154	1.85	1	85	.177