

Motion and Sound in Animated Storybooks for Preschoolers’ Visual Attention and Mandarin Language Learning: An Eye-Tracking Study With Bilingual Children

He Sun
Jieying Loh
Adam Charles Roberts

Nanyang Technological University, Singapore

This study examined the role of the “animated eBook advantage” in child bilingual’s Mandarin learning, which has tended to be examined in the acquisition of Germanic languages. With this aim, 102 4- to 5-year-old preschoolers in Singapore were assigned to one of four conditions: (a) animated eBooks (+sound+motion), (b) static eBooks with sound, (c) static eBooks only, and (d) a control condition where children played a math game on an iPad. Three stories were displayed to children each for four times over 2 weeks, while visual attention was traced with an eye tracker. Children’s target words and story comprehension were assessed for the effects of the intervention conditions. The results revealed that children in the animated condition outperform their counterparts in total fixation duration, target word production, and storytelling of one of the stories (Cycling With Grandpa). There were no consistent differences between the two static conditions. Our results indicate the importance of motion in animated eBook design, in line with previous findings.

Keywords: *animated eBook, multimedia theory, Mandarin learning, storytelling, total fixation time, motion*

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 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.

The Mechanism of Animated eBooks From the Perspective of Multimedia Theory

Animated eBooks equipped with multimedia features (e.g., video, sound, and music) have been found to support early language acquisition (Verhallen et al., 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., draw learner’s attention to the target



information), reduce extraneous processing (i.e., avoid trivial and confusing information), and foster generative processing (i.e., 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).

Conventional educational instructions rely heavily on verbal information (i.e., one system only), while animated eBooks offer children multiple verbal and nonverbal information to understand concepts that are difficult to grasp with words and static pictures alone. Building on 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, enhance their comprehension of complex expressions, and 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 has 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, Kegel, & Bus, 2011), and grammar development (e.g., Smeets & Bus, 2015). They may also facilitate children's story comprehension, such as better awareness of the goals, motivations, and emotions of the story figures (Verhallen et al., 2006). For instance, Takacs and Bus (2016) followed 4- to 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.

Preschoolers' Mandarin Learning in Singapore

Singapore is a multilingual society with three main ethnic groups (74.3% Chinese, 13.3% Malays, and 9.1% Indians) and four official languages (English, Mandarin, Malay, and Tamil) (Singapore Department of Statistics, 2016). English is the language of interethnic 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 4 to 5 years, Sun, Yin, Amsah, and O'Brien (2018) 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

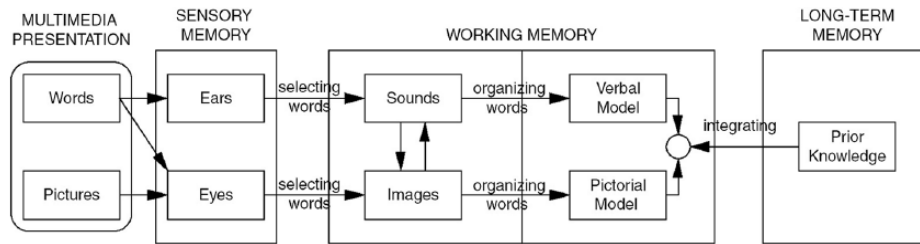


FIGURE 1. *Conceptual model of the cognitive theory of multimedia learning. Adapted from Mayer (2009).*

significantly lower than that for English learning. As Cavallaro and Ng described (2014), “English is increasingly becoming the mother tongue for more and more Singaporeans, and their ethnic languages are technically more like second languages” (p. 36).

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 11 million in 2015, increasing almost four times since 2009, and such a reading format is favored 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 preschoolers’ Chinese language development and acquisition. With this aim, we compared children’s vocabulary acquisition and reading comprehension in four conditions: (a) animated stories with motion and sound; (b) corresponding static stories, with sound and soft-copy printed illustrations; (c) corresponding static stories with soft-copy illustration only; and (d) 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 eye tracking to investigate visual attention. Eye trackers measure gaze direction and saccadic eye movements by measuring infrared light reflected off the surface of the eye. This noninvasive 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, eye tracking allows the possibility of analyzing whether attention is allocated to the presented material or other distracting stimuli in the environment through fixation duration (Wass, Smith, & Johnson, 2012).

Two questions were addressed in the current study:

Research 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.

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

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).

Methods

We utilize data collected as part of a project that examines the process and outcome of children's Mandarin learning with electronic storybooks in Singapore. Children were mainly recruited from kindergartens of the PAP Community Foundation, which is the largest operator of kindergarten and child care centers in Singapore. The project has several research goals, including exploring children's variation in attention span over repetitive readings. The current paper focuses on children's differences in vocabulary learning, story retelling performance, and total fixation time across various reading formats.

Participants

The current experiment recruited 129 kindergarten-1 children at the age of 4 to 5 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. Second, participants needed to have no history of developmental delays or impairment, based on parental report and teacher's observations. Third, participants had to complete most of the readings and 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 102 children, 49 boys and 53 girls. The socioeconomic 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\$7,500 to S\$7,999 family income per month ($M = 15.01$, $SD = 5.07$, range = 0–19, with S\$500 increment for each higher level).

Design

The study was a between-subject design in which the participants were randomly assigned to one of the following

four conditions: (a) the animated eBook reading group, (b) the static eBook with sound, (c) the static eBook with no motion and no sound, and (d) the no reading exposure control group (Table 1). The control condition is considered necessary to make sure that the vocabulary learning and story comprehension rely on the storytelling in the experiment but not the exposure of the illustrations in the posttest. Each of the storybooks was exposed to children four times as previous studies revealed that such repetitions are necessary to provide children enough opportunities to digest the story plot and retain the language (Biemiller & Boote, 2006; Justice, Meier, & Walpole, 2005; Verhallen & Bus, 2009). Children may only derive a partial meaning of a new word after the first encounter with it in a meaningful context (Clark, 1993) and cannot integrate this word into their vocabulary system until after several occasions of exposure (De Temple & Snow, 2003).

The design of the animated eBooks and the static eBooks are the same as those used in Takacs and Bus's (2016) study. For each of the same stories, the content of the story, voice of the narrator, and sceneries are identical in the animated and static versions. The aesthetic and artistic qualities of the visual representations are similar. That being said, none of the readings included distracting animations or games unrelated to the story content. The notable differences in the animated version are the cinematic techniques, such as zoom, pan, edits, and sound, being used and motions being employed to represent events described by the text (Stahl & Fairbanks, 1986). One of the storybooks used in our experiment, *Little Kangaroo*, includes an event of the mother kangaroo dancing to the tunes sung by surrounding birds and inviting the little kangaroo to join her. The little kangaroo rejects with the reason of finding the birdsongs noisy, although her foot unwittingly swings to the rhythm of songs. In the static eBook versions, the scene is depicted by the mother kangaroo in a dancing pose with birds in the background. The static eBook with sounds condition contains the addition of background sounds of birdsongs with the same static illustration. Using motion and zoom, the animated version successively depicts the whole event, further encapsulating the narration of the little kangaroo's foot moving to the rhythm of the birdsongs as well.

Materials

Three prize-winning child stories, *Little Kangaroo* (Genechten, 2007), *Cycling With Grandpa* (Boonen, 2004) and *Imitators* (Veldkamp, 2006) were chosen as the reading materials. The first two books have been translated into many languages, including Mandarin, while *Imitators* has been translated into Mandarin by the authors of the paper. *Little Kangaroo* is about how a kangaroo mother encourages her daughter to walk independently and explore the world by herself. *Cycling With Grandpa* is an adventure of a group of

TABLE 1
Experimental Design and Details

Condition	Description
Animation (+motion+sound)	Children independently read three animated eBooks four times spread over 2 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 four times spread over w 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 2 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 four times spread over 2 weeks during the intervention period. There are 13 participants in this condition.

children cycling with their grandfather on one bike. *Imitators* is a story of how a boy (and his parents) make friends with their monkey neighbors and help them avoid the zookeepers. *Cycling With Grandpa* contains the most Chinese characters and utterances among the three books (N characters = 2,034; N utterances = 97), and *Little Kangaroo* contains the least (N characters = 723; N utterances = 30), leaving *Imitators* in the middle (N characters = 1,031; N utterances = 68). These books have been successfully adapted to animated eBooks for educational purposes in the Netherlands and used in several studies on emergent language learners (e.g., Smeets & Bus, 2015). The static versions consisted of soft copy scanned illustrations from the original hardcopy storybooks and were slightly edited to create the final eBook format presented on screen. The story text was automatically read aloud, and the story would continue automatically. In the animated version, 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 a bell), and background music (e.g., low-spirit music to reflect character's frustration). For further details on the development of the animated version, see Smeets et al. (2014) and Takacs and Bus (2016). There were slight variations between the three reading conditions in terms of total reading time. For instance, *Little Kangaroo* lasted for 245 seconds in the animated condition but lasted for 249 seconds in the static conditions. We have corrected the differences in length by dividing total fixation durations for each book by the length of the story. Eighteen words (verbs, adjectives, and adverbs) were selected from the three books as the target words after consulting the Chinese teachers in the preschools, and these words are assumed to be beyond children's current Chinese knowledge in our sample.

Procedure

Before the experiment, children were assessed for their general Mandarin language proficiency (e.g., Mandarin vocabulary), cognitive abilities (e.g., phonological short-term

memory), and comprehension of target words selected from the storybooks. All the tests were conducted in a quiet room at the children's schools within school hours. Experimenters were three trained Mandarin-English bilingual research assistants who majored in psychology or linguistics. Each of the tests was introduced in Mandarin, however, the English explanation was provided if the children demonstrated difficulties in understanding the tasks or requirements. After the screening tests and the pretests, the participants underwent four reading sessions based on their assigned condition. Children had four sessions in 2 consecutive weeks, with two sessions per week, and listened to three stories in each session (Table 2). Each reading session lasted for 15 to 20 minutes, including preparation and eye-tracker calibration. The sequence of the storybooks presented was randomized by an online randomizer and was double-checked to ensure that it was not repeated. The research assistants were present to operate the eye-tracking machine (e.g., calibration) and instruct the children to read the story. The books were presented on laptop screens with the eye trackers mounted to the laptop. Once the children started to read, the experimenters were not allowed to interrupt their reading. Children in the control group spent the same amount of time playing math games on iPads (about 15 minutes) as the reading of those in experimental groups during each session. Story retelling was conducted right after the first and fourth sessions, and the posttests on vocabulary were administered on the following school day after the fourth session. The order of the four vocabulary tests was the same for all children, starting with the productive vocabulary tests to avoid any possible learning from the receptive tests.

Measures Used in Screening Tests, Pre- and Posttests

Demographic survey. A parental questionnaire was used to estimate children's language background. The questionnaire was designed based on the Language Exposure Questionnaire (Sun, Steinkrauss, Tenderio, & de Bot, 2016; Sun, Yin, et al., 2018) and the Utrecht Bilingual Language Exposure Calculator (Unsworth, 2013). In contrast to these questionnaires, the

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 and home language environment	General Mandarin, cognitive, and 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

current one explored children’s eBook preference and reading history at home in more detail.

Children’s general language proficiency has been found to affect their novel word acquisition and fixation on illustrations (Evans & Saint-Aubin, 2013). To account for this, we measured our participants’ Chinese proficiency, including receptive vocabulary size, receptive grammar knowledge, and verbal fluency in Mandarin. In addition, we assessed cognitive skills related to early language learning (i.e., nonverbal intelligence and phonological short-term memory; Sun et al., 2016).

Mandarin receptive vocabulary. The Bilingual Language Assessment Battery (BLAB; Rickard-Liow, Sze, & Lee, 2013) is a locally developed vocabulary test for Singaporean children, which is similar to Peabody Picture Vocabulary Test in format (PPVT; Dunn & Dunn, 2007). The 80-trial receptive vocabulary component is a computerized auditory-picture matching task that assesses single-word receptive vocabulary. Children were presented with four pictures and were asked to select the picture that corresponds to the given word (see Appendix B for example). BLAB receptive vocabulary tests have been reported to be reliable in the context of Singapore within the original norming sample (alphas = .75–.77) (Rickard-Liow et al. 2013).

Mandarin Receptive Grammar Test. The Mandarin Receptive Grammar Test (MRGT; Bak, 2012) is also a locally developed test for Singapore preschoolers. It comprises 60 trials and assesses preschoolers’ grammar knowledge in six aspects. Modeled after the Test for Reception of Grammar developed by Bishop (1982), children saw four images and heard a spoken sentence at the same time. They were asked to select an image based on their understanding of the sentence (see Appendix B for example). The MRGT has been found to demonstrate good external validity ($r = .64$) and internal reliability (Cronbach’s $\alpha = .75$) (Bak, 2012).

Mandarin verbal fluency. Children’s Mandarin productive vocabulary was measured with a semantic fluency task. Semantic categories “food” and “animals” have been used in the current study as they have been found to be effective in assessing child bilingual’s language ability in previous

studies (Schwartz, Moin, & Leikin, 2012; Sun, Steinkrauss, Wieling, & De Bot, 2018). Children were asked to name as many items as possible in 1 minute for each semantic category. Each appropriately named item scored 1 point (see Appendix B for example).

Nonverbal intelligence. Raven’s Colored Progressive Matrices (CPM; Raven, Court, & Raven, 1998) was used to estimate children’s analytical reasoning (Paradis, 2011). The test contains 36 items, each consisting of a big patterned picture with one piece missing. The child participant was expected to choose the missing part of a presented pattern from six options (see Appendix B for example). The task would be stopped if children made five consecutive mistakes. A raw score was calculated by adding up the total number of correct items.

Phonological short-term memory. Digit span and nonword repetition, the subtests of the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen & Rashotte, 1999), were administered to assess the children’s phonological short-term memory. The two tests compose of a list of digits or nonwords, which increase in length. They were played to participants from a computer, and children were subsequently asked to repeat them immediately after (see Appendix B for example). The scores of the two tasks were summed up, yielding a composite score for children’s short-term memory.

The effects of the interventions were examined via children’s performances in the target vocabulary tests and story retelling.

Productive vocabulary. The productive vocabulary test measured whether children could orally produce the target words in the context of the storybooks. Children were asked to complete a sentence with the target word missing while seeing the corresponding illustration on the screen. Only answers with the target words were awarded 1 point, any other answers would score 0 points (see Appendix C for example).

Receptive vocabulary. The receptive vocabulary test measured children’s comprehension of the target words in the

context of the storybooks. The multiple-choice format of the test is similar to PPVT and BLAB, and children were asked to choose the corresponding picture from four options. Target pictures and distractors were selected from the same storybooks (see Appendix C for example).

Context integration test. The context integration test used open-ended questions to examine whether children could transfer the knowledge of the target words into a novel context. For instance, children were asked to indicate the direction of the verb *jump into* as moving up and down or from left to right. Experimenters also acted out the actions of verbs to make them clear to participants. Only answers reflecting the meanings of the target words were awarded 1 point (see Appendix C for example).

Meaning recognition test. The meaning recognition test used yes-no questions to assess whether children comprehend the target word meaning in a novel context. Each question relates to one target word, and all items were presented in a quasi-random manner. See Appendix C for example.

Story retelling. Children were asked to retell the three stories using the static illustrations of the stories. The task resembles the common practice of independently “reading” a familiar storybook at home or at school (Korat, Shamir, & Segal-Drori, 2014; Sulzby, 1985; Takacs & Bus, 2016). The experimenter would first prompt the children with the more generic question of “What happened in this picture/scene?” and followed up with more pointed prompts like “How does the character feel?” or “What are the characters doing?” if the children required more assistance. Children’s retellings were then scored based on presence or absence of essential details preselected by experimenters. Important details of the narration on each page of the stories were denoted based on having at least one detail being depicted in each static illustration page and at least one detail being narrated but not shown through the static illustration. Each detail scored 0.5 marks (see Appendix D). A total score was summed up for each story. This was repeated twice, first after the child’s first reading of the story and then after the fourth reading of the story.

Visual Attention at the Illustrations

Eye-tracking system. The Tobii X3-120, a remote eye tracker, was used to measure the length of eye fixation when children listen to the stories. With a sampling rate of 120 Hz, the eye tracker could get an observation of eye movement and pupil direction approximately every 8 milliseconds. As children tend to move around quite often, it can be difficult to have them wear a head-mounted eye tracker. The Tobii X3-120 is mounted to the bottom of the laptop screen and detects eye movements using infrared reflectance, leading to

more accurate measurements. To ensure the optimal registration of eye movements, children were asked to be seated at a distance of 60 cm to 70 cm from the eye tracker, and the machine was calibrated for each child at the beginning of the reading session by asking them to fixate five dots shown on the screen. On average, the preparation took approximately 3 to 5 minutes. Total fixation durations on the storybooks were recorded. The fixation qualities of three participants were low, with registered eye movements less than 50% of the time in at least one session. Their data have been excluded, leaving 1,064 sessions from the 89 children in the three experimental conditions to be analyzed in the end.

Data Analysis and 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). 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, Maechler, & Bolker, 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 centered 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, Brockhoff, & Christensen, 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 Research 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).

TABLE 3
ANOVA Summaries for Vocabulary Tests

Effect	Productive Vocabulary				Receptive Vocabulary				Context Integration				Meaning Recognition			
	<i>F</i>	<i>df</i>	<i>df(E)</i>	<i>p</i>	<i>F</i>	<i>df</i>	<i>df(E)</i>	<i>p</i>	<i>F</i>	<i>df</i>	<i>df(E)</i>	<i>p</i>	<i>F</i>	<i>df</i>	<i>df(E)</i>	<i>p</i>
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

Note. PreTotal = total vocabulary score in pretest; PhoMem = phonological short-term memory score; NonInte = nonverbal intelligence score; Man.EFA = general Mandarin proficiency.

Story retelling. Total story retelling scores for each story were again analyzed 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 × 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 centered) Man.EFA. Class and school IDs were included as interacting random effects, using the REML procedure. To explore Research Question 2, pairwise differences of estimated marginal means for the eBook conditions within group (animated, sound only, static) were calculated (Appendix G).

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 pretest 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 pretest 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, $p < .001, d = 1.27$).

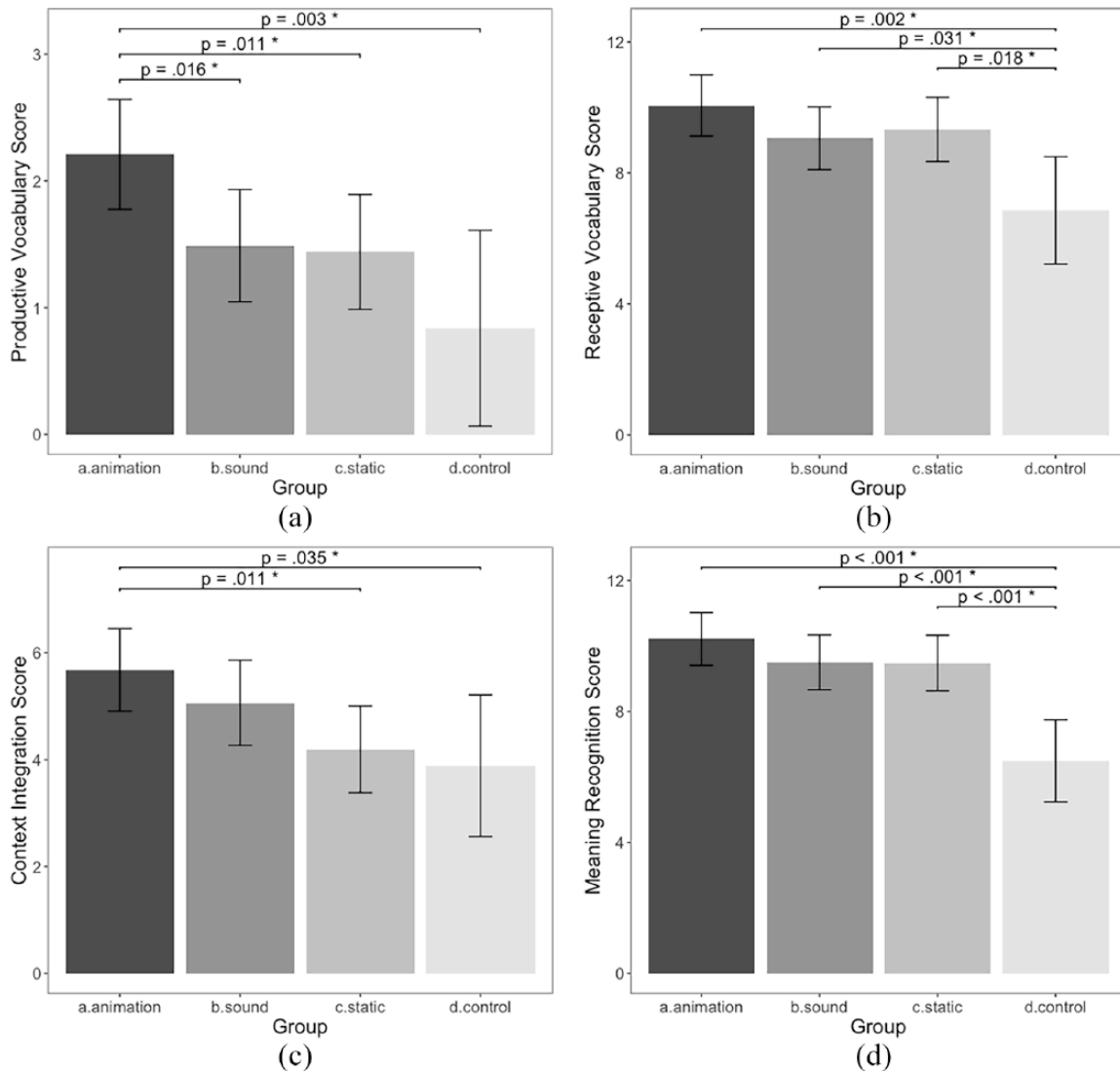


FIGURE 2. Pairwise contrasts for vocabulary tests. (a) Productive vocabulary. (b) Receptive vocabulary. (c) Context integration. (d) Meaning recognition. Error bars denote standard error of the mean. Significant pairwise contrasts are highlighted.

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 \times 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.

TABLE 4
ANOVA Summaries for Retelling

Effect	<i>Cycling With Grandpa</i>				<i>Little Kangaroo</i>				<i>Imitators</i>			
	<i>F</i>	<i>df</i>	<i>df(E)</i>	<i>p</i>	<i>F</i>	<i>df</i>	<i>df(E)</i>	<i>p</i>	<i>F</i>	<i>df</i>	<i>df(E)</i>	<i>p</i>
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 × Assessment	2.26	3	157.06	.084	4.29	3	151.78	.006	2.28	3	154.89	.081

Note. Man.EFA = general Mandarin proficiency; PhoMem = phonological short-term memory score; NonInte = nonverbal intelligence score.

Visual Attention

ANOVA summaries for the three stories are in Table 5. There was a significant effect of group for *Cycling with Grandpa*, $F(2, 85) = 3.41, p = .038$; *Little Kangaroo*, $F(2, 82.34) = 12.89, p < .001$; and *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 *Imitators* ($p = .008, d = 0.58$).

Discussion

Mandarin Vocabulary Learning and Story Comprehension From the Interventions

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 posttest vocabulary questions and fixation duration, as measured by eye tracking, 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 with 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 (Takacs &

Bus, 2016; Takacs, Swart, & Bus, 2015; Verhallen et al., 2006) 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 (McBride, 2016; Moser, 1991), 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 7 were Internet users, up from 40% in 2016 (Infocomm Media Development Authority, 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

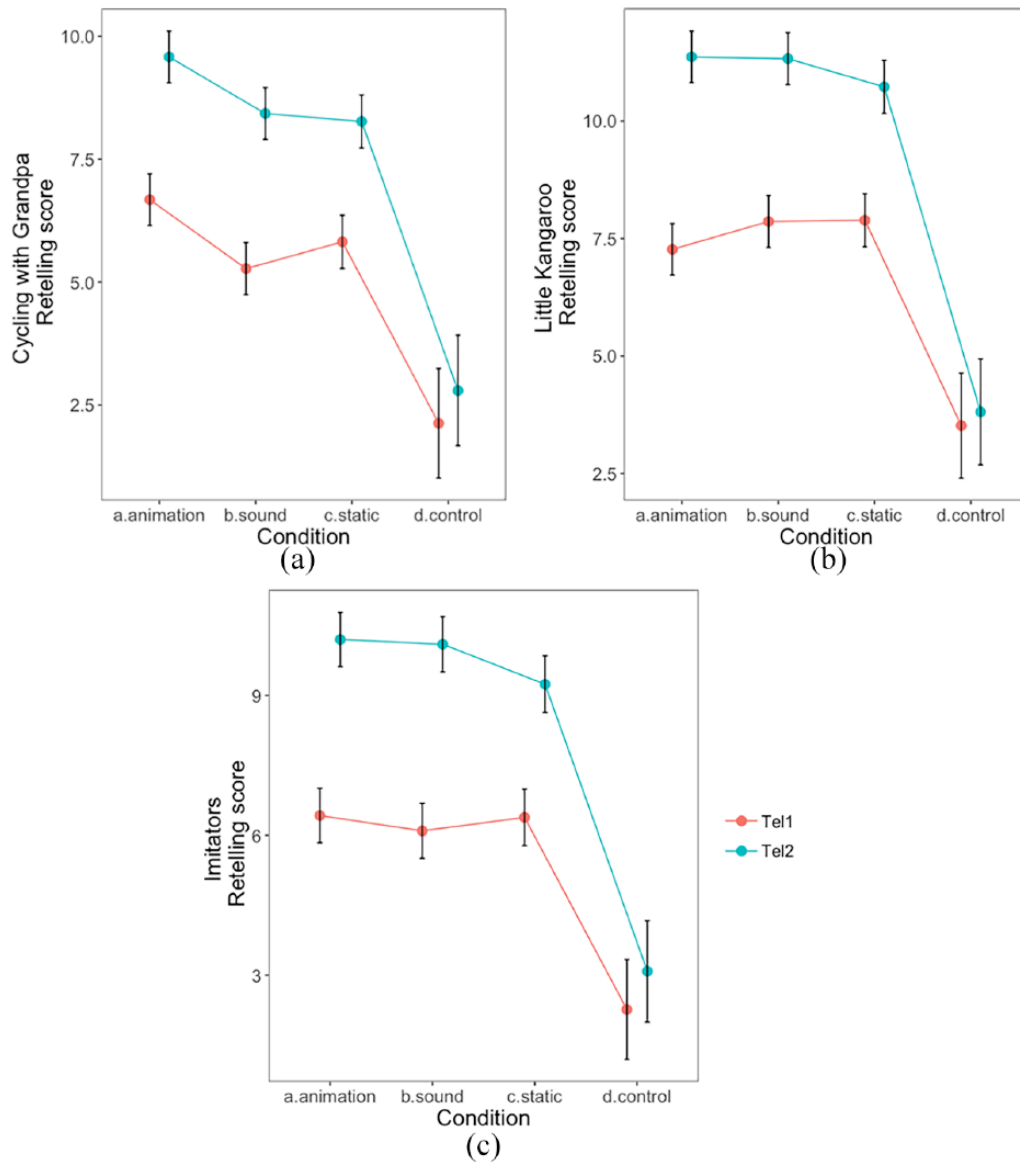


FIGURE 3. Retelling scores for (a) *Cycling With Grandpa*, (b) *Little Kangaroo*, and (c) *Imitators*, in *Assessment Tel1* (after first reading) and *Assessment Tel2* (after fourth reading). Error bars denote standard error of mean.

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 (1997; Sénéchal & Cornell, 1993) arguments that while 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. 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*

TABLE 5

ANOVA Summaries for Visual Attention

Effect	Cycling With Grandpa				Little Kangaroo				Imitators			
	<i>F</i>	<i>df</i>	<i>df(E)</i>	<i>p</i>	<i>F</i>	<i>df</i>	<i>df(E)</i>	<i>p</i>	<i>F</i>	<i>df</i>	<i>df(E)</i>	<i>p</i>
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

Note. Man.EFA = general Mandarin proficiency.

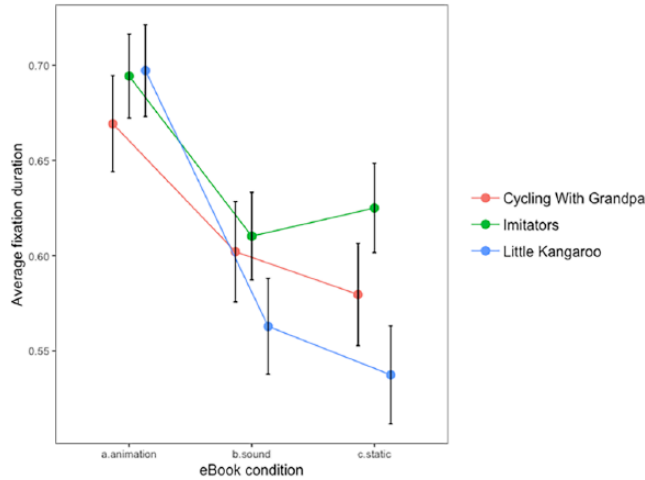


FIGURE 4. Average fixation durations for each eBook (Cycling With Grandpa, Little Kangaroo, and Imitators) for the three eBook conditions. Error bars denote standard error of mean.

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.

Total Fixation Time During the Interventions

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’s (2016) study 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 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. Second, 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–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 policymakers.

Concluding Remarks

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.

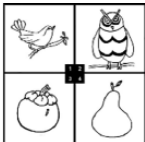
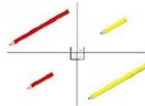
APPENDIX A

The Target Verbs and Adjectives/Adverbs From the Three Stories

Target Verbs	English Translation	Target Verbs	English Translation
荡	Swing	挥手	Wave hand
闯进	Barge in	皱鼻子	Wrinkle the nose
钻进	Burrow into	驮着	Carry
瞪眼	Stare at	耸肩	Shrug shoulders
奔向	Jump towards		
Target Adjective/Adverb	English Translation	Target Adjective/Adverb	English Translation
严厉地	Severely	愉快地	Happily
疯狂的	Crazy	骄傲的	Proudly
飞快地	Quickly	强壮的	Strong
兴高采烈地	Elatedly	无聊的	Bored
筋疲力尽地	Exhaustedly		


APPENDIX B

Examples of Screening Tests Used; Participants Were Shown Stimuli Depicted Below on the Computer

Measure	Test Used	Item Example	Demo	Correct Answer
Mandarin receptive vocabulary	Bilingual Language Assessment Battery (BLAB)	猫头鹰 (Owl)		2
Mandarin receptive grammar test	Mandarin Receptive Grammar Test (MRGT)	这支红色的铅笔 <u>不</u> 长。 The red pencil is <u>not</u> long.		3
Mandarin verbal fluency	Semantic Fluency Task	食物 (food) 动物 (animals)	NIL	面 (Noodle) 老虎 (Tiger)



(continued)

APPENDIX B (CONTINUED)

Measure	Test Used	Item Example	Demo	Correct Answer
Nonverbal intelligence	Raven’s Colored Progressive Matrices (CRM)	This is a pattern with a missing piece. Only one of these pieces is the right pattern.		4
Phonological short-term memory	Comprehensive Test of Phonological Processing (CTOPP)	Digit Span repetition 1 7 9 Nonword repetition Jup	NIL	1 7 9 JupJup (jōōp)


APPENDIX C

Examples of Tests Assessing Knowledge of Keyword Burrow Into, Including Pictures Shown to Participants to Elicit the Keyword

Measure	Item Example	Picture Demo	Correct Answer
Productive vocabulary (cloze test)	小袋鼠_____妈妈的育儿袋。 The little kangaroo doesn’t like to walk on her own, she would always _____ Mother Kangaroo’s pouch.		钻进 burrow into
Receptive vocabulary	在小袋鼠的故事里的这几副图中，哪副你可以看到“钻进去”？ In the story of Little Kangaroo, which one depicts “burrow into”?		2
Context integration test	当你“钻进”车子时，是低头弯腰还是抬头挺胸？ When you “burrow into” a car, do you lower your body or keep upright?	NIL	低头弯腰
Meaning recognition test	小朋友们在玩捉迷藏，小明立刻“钻进”了衣橱里。 The children are playing hide and seek and Xiao Ming immediately “burrow into” the closet.	NIL	Lower your body

APPENDIX D

Example of Story Retelling Scene Shown to Participants With Prompts

<p><i>Story Retelling</i></p> 	<p>妈妈指向了正在花丛里飞舞的蝴蝶。 Mummy Kangaroo points at the dancing <u>butterflies</u> among the flowers. (Score: 0.5)</p> <p>小袋鼠不喜欢蝴蝶，觉得它们很烦人。 Little Kangaroo <u>does not like the butterflies</u> and feels that they are <u>annoying</u>. (Scores: 1)</p>	<p>Prompts: 他们是什么？ What are those?</p> <p>小袋鼠喜欢吗/觉得怎么样？为什么？ Does Little Kangaroo like them/how does she feel about them? Why?</p>
---	--	---

APPENDIX E

Post Hoc Pairwise Comparisons Within Conditions for Vocabulary Tests

Contrast	Productive Vocabulary			Receptive Vocabulary			Context Integration			Meaning Recognition						
	Estimate (SE)	t	df	p	Estimate (SE)	t	df	p	Estimate (SE)	t	df	p				
Animation – sound	0.72 (0.31)	2.35	82.81	.021	1.01 (0.66)	1.52	84.03	.131	0.62 (0.57)	1.10	82.28	.274	0.71 (0.62)	1.16	93.00	.249
Animation – static	0.77 (0.31)	2.48	90.97	.015	0.75 (0.67)	1.11	91.54	.269	1.48 (0.57)	2.60	92.01	.011	0.74 (0.61)	1.20	93.00	.233
Animation – control	1.37 (0.47)	2.97	33.02	.005	3.22 (0.99)	3.27	34.32	.002	1.78 (0.80)	2.23	26.28	.035	3.72 (0.79)	4.72	93.00	<.001
Sound – static	0.05 (0.32)	0.17	83.16	.868	-0.26 (0.68)	-0.39	84.34	.701	0.86 (0.58)	1.48	82.51	.143	0.02 (0.63)	0.03	93.00	.973
Sound – control	0.66 (0.46)	1.43	33.92	.162	2.21 (0.98)	2.25	35.40	.031	1.16 (0.80)	1.45	27.59	.157	3.01 (0.79)	3.81	93.00	<.001
Static – control	0.61 (0.47)	1.30	34.74	.203	2.47 (1.00)	2.48	36.16	.018	0.30 (0.81)	0.37	28.12	.713	2.99 (0.80)	3.72	93.00	<.001

APPENDIX F

Post Hoc Pairwise Comparisons Within Conditions For Retelling Tel2

Contrast	Cycling With Grandpa			Little Kangaroo			Imitators					
	Estimate (SE)	t	df	p	Estimate (SE)	t	df	p	Estimate (SE)	t	df	p
Animation – sound	1.15 (0.58)	2.00	170.08	.047	0.04 (0.63)	0.06	168.54	.951	0.10 (0.74)	0.14	175.32	.890
Animation – static	1.31 (0.60)	2.20	176.53	.029	0.64 (0.65)	0.98	176.75	.328	0.96 (0.76)	1.26	184.10	.209
Animation – control	6.78 (1.24)	5.47	48.42	<.001	7.56 (1.26)	6.02	44.82	<.001	7.11 (1.23)	5.77	59.21	<.001
Sound – static	0.16 (0.59)	0.28	169.29	.784	0.60 (0.64)	0.93	167.68	.354	0.86 (0.76)	1.13	174.91	.262
Sound – control	5.63 (1.24)	4.55	48.49	<.001	7.52 (1.25)	6.00	45.03	<.001	7.01 (1.23)	5.68	60.27	<.001
Static – control	5.47 (1.25)	4.39	49.42	<.001	6.92 (1.26)	5.48	45.99	<.001	6.16 (1.24)	4.95	61.39	<.001

APPENDIX G
Post hoc Pairwise Comparisons Within Conditions Across Assessments (T1 – T2) for Retelling

Contrast	Cycling With Grandpa			Little Kangaroo			Imitators			
	Estimate (SE)	t	df	Estimate (SE)	t	df	Estimate (SE)	t	df	p
Animation – sound	0.25 (0.76)	0.33	156.62	-0.63 (0.84)	-0.76	156.88	0.23 (1.00)	0.23	158.03	.822
Animation – static	-0.46 (0.77)	-0.60	156.62	-1.26 (0.84)	-1.49	156.88	-0.92 (1.01)	-0.91	158.03	.365
Animation – control	-2.24 (0.99)	-2.26	156.75	-3.80 (1.09)	-3.51	157.03	-2.96 (1.30)	-2.28	158.30	.024
Sound – static	-0.71 (0.78)	-0.91	156.62	-0.63 (0.86)	-0.73	156.88	-1.14 (1.03)	-1.12	158.03	.267
Sound – control	-2.49 (1.00)	-2.49	156.75	-3.17 (1.10)	-2.90	157.03	-3.18 (1.31)	-2.43	158.29	.016
Static – control	-1.78 (1.00)	-1.77	156.75	-2.55 (1.10)	-2.31	157.03	-2.04 (1.32)	-1.55	158.29	.124

APPENDIX H
Post Hoc Pairwise Comparisons Within Conditions for Visual Attention

Contrast	Cycling With Grandpa			Little Kangaroo			Imitators			
	Estimate (SE)	t	df	Estimate (SE)	t	df	Estimate (SE)	t	df	p
Animation – sound	0.07 (0.04)	1.88	85	0.13 (0.03)	3.98	79.5	0.08 (0.03)	2.69	85	.008
Animation – static	0.09 (0.04)	2.48	85	0.16 (0.03)	4.66	84.9	0.07 (0.03)	2.19	85	.031
Sound – static	0.02 (0.04)	0.61	85	0.03 (0.03)	0.73	81.2	-0.01 (0.03)	-0.46	85	.648

References

- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- 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 Eigen and Eigenfaces* (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), 44–62. doi:10.1037/0022-0663.98.1.44
- Bishop, D. (1982). *TROG: Test of receptive grammar*. Abingdon, UK: 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. doi: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* (pp. 33–48). Bristol, UK: Multilingual Matters.
- 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), 145–155. doi: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. doi: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: Lawrence Erlbaum Associates Publishers.
- Duff, P., Anderson, T., Illyckyj, R., Van Gaya, E., Wang, R., & Yates, E. (2013). *Learning Chinese: Linguistic, sociocultural, and narrative perspectives*. Berlin: DeGruyter.
- Dunn, L. M., & Dunn, D. M. (2007). *The Peabody Picture Vocabulary Test* (4th ed.). 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. doi:10.1037/a0032465
- Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics, 65*, 999–1010. doi: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.
- Hio, L. (2015, July 13). E-books click with more library users in Singapore. *The Straits Times*. Retrieved from <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. doi: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. doi: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. doi: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. doi:10.18637/jss.v082.i13
- Liversedge, S., & Findlay, J. (2001). Saccadic eye movements and cognition. *Trends in Cognitive Science, 4*, 6–14. doi: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. doi:10.1111/j.1551-6709.2012.01239.x
- 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. doi:10.2190/yjlg-09f9-xkax-753d
- 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.
- Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York, NY: Cambridge University Press.
- 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. doi: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, NY: Simon & Schuster.
- Neuman, S. B. (2008). The case for multimedia presentations in learning: A theory of synergy. In A. G. Bus, & 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. doi:10.1075/lab.1.3.01par
- R Development Core Team. (2010). *R: A language and environment for statistical computing*. Retrieved from <http://www.R-project.org/>
- Raven, J. C., Court, J. H., & Raven, J. (1998). *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. doi:10.1080/17470210902816461
- Rickard-Liow, S. J., Sze, W. P., & 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*. New York, NY: 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. doi: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: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. doi:10.2307/747933
- Singapore Department of Statistics. (2016). *Population trends 2016*. Singapore: Author.
- 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. doi: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. doi: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. doi: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. doi: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. doi: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. doi: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. doi: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. doi: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. doi: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. doi: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. doi: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. doi: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. doi: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*, 45, 229–250. doi: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.

Authors

HE SUN is a research scientist at National Institute of Education, Nanyang Technological University, Singapore; email: he.sun@nie.edu.sg. Her major interests are shared (e)book reading and emergent bilingual's language and literacy development and

individual differences in early bilingualism and social emotional learning from the perspective of cognitive and environmental factors.

JIEYING LOH is a research assistant at National Institute of Education, Nanyang Technological University, Singapore. Her major interest is early childhood intervention and children's language and literacy development.

ADAM CHARLES ROBERTS is a senior research fellow at School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore. He has extensive research experience with eye tracker on child and adult learners.