The Magic of Cartoons?
How English-learning Animations Enhance Child L2 Acquisition

by

Ruoyu Huang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Psychology)
in the University of Michigan
2023

Doctoral Committee:

Professor Ioulia Kovelman, Chair
Professor Twila Tardif, Co-Chair
Professor William Baxter
Professor Susan Gelman

Ruoyu (Lexie) Huang

ryhuang@umich.edu

ORCID iD: 0000-0001-7389-8796

Dedication

To my very first angel cat, Acorn.

Acknowledgments

As I'm finishing up this dissertation, the summer in Ann Arbor is waning, with beautiful trees and flowers adorning the entire town, just like the delightful summer that greeted my arrival six years ago. These past years have been a great journey for me to grow professionally as a psycholinguistic researcher and developmentally as a better person. I cannot appreciate enough how lucky and grateful I am to work with the brilliant minds in research and meet with great people who eventually become my second family in this country.

I would like to, first, express my deepest gratitude to my primary supervisor, Professor Twila Tardif, who has been the most patient and supportive along the journey. This dissertation would not have been possible without her tremendous knowledge and contributions to the field of child language development. I would also like to thank her for her trust and faith in me when I went out alone to collect data for this project and thank her for the freedom and encouragement she provided me to explore what's good and challenging in science. Most importantly, I sincerely appreciate her support for me in learning how to be authentic and confident in this overwhelmingly competitive and intense world.

Next, I would like to extend my great appreciation to my dissertation committee. A huge thank you to the chair, Professor Ioulia Kovelman, who (with all her lovely lab members: Kehui Zhang, Valeria Caruso, Rachel Eggleston, Isabel Hernandez and many more others) offered me enormous help to graduate in my last year of the program. I felt welcomed and valued when I presented my research ideas in the lab, and I would have felt lost and aimless if she was not there to support me to the end. Another big thank you goes to my committee member, Professor Susan

Gelman. Susan always strikes me as an intelligent, open-minded and heartwarming mentor. I cherish all the “aha moments” in my conversations with her and all the insightful feedback she has given me on this project. They taught me to be critical and creative and to think from children’s perspectives when I study children. Next, I want to express my gratitude to the last important member of my committee, Professor William Baxter in the Linguistics Department. The detailed thoughts Bill put on the target exposures children received in my study truly reflect a new view of how we should look at what children *hear* in the linguistic input, which inspired me to want more rigorous stimulus design in my future career.

Lastly, I would like to give appreciation from the bottom of my heart to my dearest friends who have unconditionally loved and supported me since day one in my happy and dark times. The exhaustive list of names will be too long for this short acknowledgment section, so I won’t show them all, but will give a special thank you to my two wonderful officemates and the closest friends in this program: Michael Demidenko and Yeonjee Bae. Although we studied very different things in developmental psychology, we laughed, cried, traveled and grew together like a family. My journey would never be the same without them—they brought color, meaning and strength to my life.

Thank you, everyone, for your kind words, insightful critiques, marvelous inspirations, and unconditioned love throughout my six-year journey. I shall look forward to the path ahead with pride and hope.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

Growing up in the digital era, young children are exposed to numerous educational applications and programs, some with well-designed curricula and others focused more on entertainment or game-like qualities. With the increasing popularity of mobile language-learning applications among young children, this dissertation raises a crucial question: do these learning materials actually enhance children's acquisition of language? A total of 71 preschool-aged Mandarin-speaking children (M_{age} = 5.76, SD_{age} = .59) participated in a one-week-long intervention study in Nanjing, China. Across a total of two visits, we assigned children to two experimental conditions to watch targeted English-learning animations focusing first on verbs, in one condition, and on possessives in the other. Children from each condition were encouraged to watch their condition-related animations as many times as they wanted between visits, along with a new take-home animation that was neither verb- nor possessive-focused. Children's performance on target verb and possessive items (vocabulary and grammar) for both comprehension and production was measured after each animation viewing session, in addition to assessing general English verb tense and possessive marking. Overall, our results suggest that preschool-aged Mandarin-speaking L2 learners can benefit from targeted English-learning animations. Moreover, consistent with first language acquisition results in both English and Chinese, some target language categories appear to be easier to acquire than others (e.g., possessive grammar - *'s* > verb past tense grammar). In addition, comprehension measures for both vocabulary and grammar appear to be easier than production measures. Finally, our results demonstrate that standardized English tests that reflect children's general language knowledge

might not be sensitive enough to capture children's limited exposure to targeted language-learning materials. Moreover, different amounts of exposure might be acquired for successful learning of different target language categories in L2 depending on their difficulty (e.g., easy L1 to L2 transfer facilitated learning of possessive -*'s* marking among Mandarin-speaking children). These findings not only demonstrate the potential for well-designed language-learning applications, and they also speak to a number of limitations, challenges, and future directions in how we might examine the efficacy of language-learning media.

Chapter I Introduction

English is the most widely spoken non-native language in the world (Rao, 2019). Pre-pandemic estimates by the British Council suggested that by 2020, one in four people across the globe, roughly two billion people, would be speaking English "at a useful level" (British Council, 2013, p.5). Whether this specific number has been reached or not, English has become a global *lingua franca* such that learning English as a non-native language has become essential for economic success and global reach at both a national and an individual level (Rao, 2019). Parents of young children are particularly interested in English-language learning (Rao, 2019; Xu et al., 2020). However, this rise in the demand for English-as-foreign-language instruction for young children has far exceeded the supply of qualified teachers with formal training or fluency to provide the best opportunities for students (Shin, 2008).

One solution has been to develop educational materials and programs with newer technology (e.g., digital devices, web-based learning platforms) that allow instruction to reach individual users at low cost and with maximal convenience—literally, to put learning "in the hands" of learners. Following this trend, numerous applications (APPs) and programs have been developed worldwide for young second-language learners. One challenge, in addition to ensuring access to such technology, is that very few of the products have been examined for their efficacy in helping children actually *learn* a second language (Kokla, 2016; Martins & Booth, 2022). In fact, there are minimal findings regarding the efficacy of educational APPs in general on digital devices (e.g., tablets, mobile phones) or on traditional media platforms (e.g., TVs), and even fewer focused on the very youngest learners.

Do Digital Media Facilitate Early Child Development?

For decades, there has been a scientific debate on whether digital media exposure enhances or hinders children's social and/or cognitive development skills (e.g., Fisch, 2004a; Hassinger-Das et al., 2020; Hurwitz, 2019), including those important for school readiness (e.g., emotion regulation, literacy, and numeracy skills). One line of research examined the association between the *quantity* of media exposure (e.g., screen time) and early child development. The findings from this line of research are mixed in the efficacy of media exposure but show a rather consistent impact of age and viewing experience—whereas children younger than two or three years old rarely pay attention to or understand anything from screen media (Anderson & Pempek, 2005; Huston & Wright, 1983; Richert et al., 2010), preschoolers and older children demonstrate more variations in learning outcomes depending on how the media are used as educational tools (Puzio et al., 2022; Jing et al., 2023).

For example, some scholars (e.g., Krcmar et al., 2007) suggested that without human interaction, prelinguistic children were not likely to learn new words from passive screen-viewing (e.g., watching *Teletubbies*). In addition, a negative association was found between young children's screen time consumption and early development, including executive functioning, expressive language, and motor development (van den Heuvel et al., 2019; Lin et al., 2015; Madigan et al., 2020). However, some researchers argued that young children could have positive learning results from media if they were exposed to well-designed educational programs embedded with pedagogical teaching strategies. For instance, Krcmar and colleagues (2007) found that vocabulary learning from television was possible for toddlers older than 22 months if the target program was designed with live interaction (e.g., joint referencing by calling the child's name) and child-oriented content. These features help children focus on the delivered

content and facilitate vocabulary development. Furthermore, in a longitudinal study comparing several popular real-life children's TV programs, Linebarger and Walker (2005) found that children at 30 months of age showed more vocabulary learning and higher expressive language scores from watching *Dora the Explorer, Blue's Clues, Arthur, Clifford* and *Dragon Tales* than from watching *Teletubbies*, a child TV program with a poor language model and less interactive viewing experience. This is to say, the *quantity* of media exposure alone does not offer a sufficient explanation for whether children can or cannot learn. Instead, the *quality* of media content should be considered when examining their learning efficiency (Gola et al., 2012).

A growing body of research has started to pay attention to the quality of media with most examining television viewing with cross-sectional samples on children's cognitive, vocabulary and literacy improvement or overall academic achievement (Espinosa et al., 2006; Ferguson & Beresin, 2017; Li & Brand, 2009; Stracker et al., 2018). These studies did not measure children's screen time on mobile devices, nor did they capture the materials children now watch on these digital and portable devices (Viner, 2019). Additionally, many studies did not distinguish educational content from regular media content (Puzio et al., 2022).

In the current era of digital natives and the post-COVID learning environment, there have been rapid increases in the number of educational APPs and programs being developed and used by children in and out of school, raising concerns about the educational values of these so-called "educational" APPs (Hirsh-Pasek et al., 2015). According to *Common Sense Census* of media (Rideout & Robb, 2020), children aged from zero to eight spend an average of two and half hours on screen media daily. These data were collected right before the start of the COVID-19 pandemic, and it is reasonable to anticipate even higher amounts of screen time post-COVID. Thus, it is necessary to investigate the efficacy of current mobile educational APPs or programs

as this can help educators and parents choose the right programs for young L2 learners. Moreover, many educational language-learning APPs claim to follow the decades of research in the Science of Learning, as well as on First (L1) and Second language (L2) acquisition. Thus, examining how young children learn from these APPs will also advance our knowledge of how the Science of Learning principles really work in real-life language-learning products. Finally, this examination has the potential to increase the quality of educational APPs across media platforms.

General Aims of the Current Study

The overarching research question for this dissertation is whether preschool-aged children are able to learn English as a second language with brief exposure to targeted educational animations. If so, then what *exactly* can they learn? Is it vocabulary, grammar, or both? Do the animations facilitate comprehension only, production only, or both? To the best of our knowledge, few studies have rigorously manipulated viewing exposure among children. For those that did, home viewing experience outside of the laboratory was not well-controlled as it is difficult to monitor and measure this in today's conditions unless specific content is given to the parents on a specific device (simply asking parents to estimate children's general media exposure time does not account for types of media or the multiple devices and sources that children may be able to view). Moreover, very few studies have gone beyond an assessment of general language ability to examine item-specific learning that was targeted in the educational media. Furthermore, there is insufficient research on young children's grammar and spoken language learning through educational media (Gola et al., 2012). This suggests a necessity for better experimental design and language measures as well as more in-depth data analyses when it

comes to examining whether a language-learning program is effective for the aspects of language that are targeted.

The current study aimed to fill some of the gaps in the literature. Specifically, it was designed to focus on non-native English-speaking preschool-aged children's comprehension and production of vocabulary and grammar in English. after viewing audiovisual media designed for L2 learners of English.

The three main research questions and hypotheses are specified in Chapter II. In addition, Chapter II gives a brief overview of how children's educational TV programs and targeted language-learning APPs adhere to various New Science of Learning principles. In addition, it presents empirical findings on children's language learning through educational TV programs and APPs as well as on children's development of their first and second language. The English-learning animations selected for the current intervention study were designed to follow as many principles as possible from the New Science of Learning and these features are also presented and discussed in Chapter II. Chapter III proceeds to provide a brief overview of children's language learning from a developmental-psycholinguistic perspective, which provides theoretical support for the hypotheses introduced later in the dissertation. Chapters IV and V provide detailed experimental design methodologies and analytical results for each research question. Chapter VI provides a general discussion for the current study and raises considerations for future research.

Chapter II How the New Science of Learning Principles Guide Language-Learning Media

The science of learning aims to offer theoretical explanations for how people learn by synthesizing scientific work from psychology, education, cognitive science, and many other disciplines. It was first applied to real-life schooling in the early 20 century to prepare students for the industrialized economy (Sawyer, 2014). As researchers realized the flaws in traditional schooling methods of *instructionism* (e.g., rote learning), the New Science of Learning was established in the 1970s. It encourages close interdisciplinary collaboration on children's learning behaviors with additional empirical support from neuroscience and computer science. The New Science of Learning reveals children's capabilities for statistical learning and analytical reasoning (Melzoff et al., 2009), and it advocates a learner-oriented learning experience, focusing on deeper conceptual understanding and scaffolding on learners' prior knowledge (Swayer, 2014).

Two of the most cited papers on the Science of Learning (Hirsh-Pasek et al., 2015; Melzoff et al., 2009) discuss both the foundation of the new learning science and the practical applications to real-life digital learning. This dissertation synthesizes their highlighted principles that should be considered when designing educational material for children as the following: (1) learning is computational, (2) learning is neuro-cognitively linking self and other, (3) learning is and should be social, (4) learning should be active, (5) engaging, (6) meaningful, (7) scaffolded and (8) goal-oriented. Take language learning as an example, the first three principles acknowledge a child's capability to be a universal learner for learning any target language with an adequate amount of social interaction. The last six reveal the important pedagogical design

that can provide children with a successful language-learning experience through digital media in the 21st century.

This chapter begins by reviewing the empirical literature on children's language learning through educational programs or APPs to explicate the aforementioned principles. Although these studies have provided valuable insights into the ways in which media can facilitate children's language learning, they are constrained by various limitations, including issues with target language measures and data analyses. Addressing these limitations is one of the driving forces behind the current dissertation.

Empirical Studies with Language-Learning Media

Principle 1: Learning is computational.

According to Meltzoff et al. (2009), the computational skills infants are equipped with enable them to detect and apply patterns by utilizing frequency and probabilistic distribution in the language input. This affords infants the cognitive capacities to learn any language in their living environment. One of the most well-known studies by Saffran et al. (1996) demonstrated 8-month-olds' statistical learning ability on artificial word segmentation after two minutes of audio exposure. Fulkerson and Waxman (2007) showed that infants at 6 and 12 months were able to form conceptual object categories when the same novel word was used to name a set of distinct objects. Moreover, among multiple ambiguous word-reference situations (given audio and visual stimuli), Smith and Yu (2008) reported 12- and 14-month-old infants could take advantage of the statistical regularities in the exposed word-scene co-occurrences for rapid word learning. This experiment reflects children's rapid lexical knowledge building in the early years, indicating their capability to accrue statistical evidence for successful word-meaning fast mapping from the overwhelming linguistic input in real life. Although children's universal language learning

abilities narrow to their native language with age, Chen (2013) suggested this ability consistently assisted L2 word learning across 30-month-olds, young school-aged children and college students. In addition to lexical acquisition, Godfroid and Kim (2021) also provided evidence for statistical learning in implicit L2 morphosyntactic knowledge among adults with a mean age of 24.

When this principle is applied to language learning from media, we would expect children to extract linguistic regularities from the naturalistic language input in the media exposure. However, this does not suggest that children can learn a target language from any type of media with naturalistic language content—many studies, in fact, reported there was no positive association between children's language learning and media exposure (Anderson & Pempek, 2005; Linebarger & Walker, 2005)—but suggests that quality educational media designed to assist language learning should provide sufficient linguistic structures for children to have a better and more efficient learning experience. This is because both children and adults can have better lexical acquisition when there are multiple word-referent pairs than a single or a few (e.g., 18 vs. 9 pairs in adults. Yu & Smith, 2007; Yu et al., 2007). Another related linguistic benefit applied in educational media could be having the visual representation and the sound of a target word occur at the same time to emphasize their association (Gogate et al., 2000) and having the same new words appear across multiple audio-visual examples (Silverman et al., 2013).

In a recent content analysis across five bilingual educational TV programs, Rivera Pérez et al. (2021) revealed various pedagogical approaches used for bilingual children's English and Spanish learning. Compared to three other shows (i.e., *Maya & Miguel*, *Handy Manny* and *Nina's World*), *Dora the Explorer* and *Go, Diego, Go!* appear to embed more vocabulary

learning cues (e.g., more visual presentation, repetition and various examples for new words) that theoretically facilitate children's lexical development. Although numerous studies have provided evidence of lexical development through these real-life educational media (Comstock & Scharrer, 2010), they did not directly test how effective the pedagogical approaches adapted from the Science of Learning were. One of the important reasons was that most studies were correlational, measuring children's vocabulary gains either by giving a standard vocabulary test or by collecting parents' reports to reflect children's general vocabulary improvement, not directly assessing the exact content children viewed (Gola et al., 2012). In other words, as we have a growing body of detailed content analysis for educational programs, which discuss specific learning targets, there is not much corresponding work done on the empirical side of the research.

Principle 2: Learning is neuro-cognitively linking self and others.

The second principle sheds light on children's neurocognitive systems that can link the action of self and others, which enables children to adapt themselves to the complicated world they are born into by observing, learning and imitating real-life models. This ability is particularly important for language learning, which requires enormous tries to apply observed linguistic rules, hand gestures, facial and oral muscle coordination, as well as correcting grammar errors to become a native speaker of a target language (Meltzoff et al., 2009). Researchers like Möttönen et al. (2005) argue that there might be a mirror neuron system supporting language learning in Broca's area and primary motor cortex. In their experiment, eight adults with a mean age of 24 showed related motor neuron activation by viewing articulatory gestures of speakers' lip movements. Moreover, Imada and colleagues (2006)

provided MEG evidence to support that infants' Broca's area could be activated when listening to human speech.

It is not a difficult extension to assume that the same areas would be activated when children listen to and watch native speakers in cartoons, although it is possible that they may not show as much mirroring or selectively show the same neuro activity (e.g., show similar brain activity with the face of cartoon female faces than cartoon character faces) when watching cartoons vs. human facial movements (Yamanaka et al., 2022). Nevertheless, a number of researchers have pointed out that children as young as five months can perceive cartoon human faces as faces in both behavioral and neuroimaging experiments (Kobayashi et al., 2020; Yamanaka et al., 2022). Furthermore, Hyde et al. (2014) suggested that children between four to ten years old interacted with animated cartoon characters in a similar way as they would with real humans. This evidence implies the potential for children to learn a target language with animated human characters, apart from other benefits children can get from cartoons in general—attracting more interest and visual attention from children and facilitating knowledge learning, including lexical and grammar learning with autistic children (Bosseler & Massaro, 2003; Kilicgun, 2015; Snyder & Mares, 2021; Valkenburg & Vroone, 2004).

Similar to the first principle, Principle 2 is created based on a plausible social-cognitive learning mechanism rather than a pedagogical feature that can be directly tested with large-scale real-life educational programs. As we move on to the rest of the principles that are more related to content design, we will introduce empirical studies on language learning with educational media.

Principle 3: Learning is/should be social and interactive.

Whereas Principles 1 and 2 acknowledge children's inborn capacities and powers to learn enormous information from the world they are exposed to, the third principle we highlight here emphasizes the selective process of children's learning. According to Meltzoff et al. (2009), children's computational skills are not applied indiscriminately. Instead, they tailor their target learning (e.g., *what* and *when*) to the needs of real-life social interactions. In this learning process, social-cognitive skills, such as joint attention, social pointing and gestures, are fundamental for child development (Talbott et al., 2020; Tomasello, 2003). Children who have deficits in social functions (e.g., autistic children) were commonly reported to be delayed in multiple aspects of language development, including delays in vocabulary and overall literacy achievements (See a recent review by Hwang & Lee, 2022). In addition, there is a plethora of studies suggesting that children learn a language more efficiently when they are engaged in a certain level of social interaction with either a co-viewing partner or a teaching agent that can provide contingent feedback (Gola et al., 2012; Kirkorian et al., 2016; Richert et al., 2011; Roseberry et al., 2009; Roseberry et al., 2014). The younger the learner, the more essential roles of social interaction and clear communication play (Baldwin, 2000; Mares & Acosta, 2010). As children get older to preschool ages (between three and seven), they start to view screen content as different or unreal from reality and learn more from digital content if there is a credible source of information and if there is a good parasocial relationship with the on-screen characters, even if there is no contingent feedback (Fisch, 2004b; Richards & Calvert, 2017; Richert et al., 2011).

As Hirsh-Pasek et al. (2015) proposed that social interactive features should be included in children's learning with educational APPs, empirical studies on language learning with educational television have provided evidence to support this principle. For example, Chiong and Shuler (2010) reported significant vocabulary and literacy gains among children between three

and seven years old, who used the APPs *Martha Speaks: Dog Party* and *SuperWhy!* respectively. These two educational APPs both provided contingent feedback during children's play. For audio-visual programs that seemingly only involve watching, embedding pseudo interaction between the viewers and the onscreen character facilitates children's learning (Calvert & Richards, 2014). This one-way active interaction from child viewers to onscreen characters, although might not be as effective or natural as two-way face-to-face interaction, still enhances children's learning experience and helps build children's parasocial relationships with TV characters. For example, by utilizing interactive techniques like having onscreen characters look at the audience, pause after asking a question, singing to the audience etc., *Sesame Street, Blue's Clues* and *Dora the Explorer* made preschool-aged children feel closer to the main characters, which leads to more emotional and cognitive participation as well as improved comprehension as educational content is delivered on screen (Calvert et al., 2007; Crawley et al., 1999; Miron et al., 2001).

Principle 4: Learning should be mentally active.

Active learning is essential for knowledge-building for both adults and children. As defined in Hirsh-Pasek et al.'s article (2015), the active learning that should be applied to educational APPs is not only about making the learners do physical activities (e.g., clicking buttons, turning pages) but requires constant minds-on learning experience that involves "thinking and intellectual manipulation" (p. 8) on the spot of learning. Compared to traditional lecturing, where students passively listen to the lecture, the active learning approach that involves in-class activities and discussions leads to better conceptual understanding and examination scores among college students in the STEM areas (Freeman et al., 2010). Consistent with adult learners' data, children who demonstrated chemical processes through drawing while

learning chemistry performed better than those who only read the same learning material (Fiorella & Zhang, 2018; Van Meter, 2001).

When it comes to children's learning with educational media programs, many studies were guided under the framework of the capacity model (Fisch, 2004a), with the assumption that children's limited attentional capacity needs to be directed appropriately. Active learning was proposed as one way to help children comprehend media content with participatory cues embedded to break the fourth wall (Anderson et al., 2000; Piotrowski, 2014). Some researchers believe that by asking children content-related questions and giving them time to think before responding to the screen (e.g., *Mister Roger's Neighborhood*, *Blue's Clues*), children are offered an opportunity to actively rehearse the learning content, which increases their attention and mental effort invested in the educational content (Crawley et al., 1999). However, the empirical studies reported inconsistent results. For example, by showing preschoolers an episode from *Dora the Explorer* with or without participatory cues, Piotrowski (2014) showed there was no statistically significant difference between the two viewing groups on content comprehension, even though the children who received participatory cues indeed engaged more during viewing. Nevertheless, there is very little evidence supporting or against the effects of active learning in young children's language learning through real-life educational media—the existing positive learning results in vocabulary learning through educational media (e.g., *Sesame Street;* Rice et al., 1990) are not sufficient to show it is the active participation children have during the viewing that contribute to the final language improvement. This is an inevitable and ongoing challenge researchers have to face because most child-centered educational programs combine multiple pedagogical features in curriculum design such that is impossible to tell them apart for individual

principal testing unless the experimenters strategically manipulate the viewing content (e.g., Piotrowski, 2014).

Principle 5: Learning should be engaging and focused.

While active learning provides learners with the steering wheel to learn how to drive, engagement in learning clears a path for them to focus on *driving* without getting distracted by irrelevant objects on the road. Compared to traditional books and lectures, modern technology and media enable more multi-channel information delivery (e.g., embedding audio, video, and interactive activities in learning). However, multiple studies suggest that adult learners can learn more deeply if they are exposed to essential learning materials without extraneous material (Mayer et al., 2001; Mayer et al. 2008). For child learners who have smaller cognitive capacities to focus, balancing between the amount of child-oriented entertainment and the level of on-task content is more challenging and important. For example, Kannass and Colombo (2007) showed that any type of distraction could significantly impair three-and-half-year-olds' problem-solving skills. Although doing better with intermittent distractors, the four-year-olds still struggled in the tasks when exposed to continuous distraction.

Furthermore, studies on children's language learning with media showed that compared to foreground TVs, children did not learn from background TVs because they rarely paid attention or engaged in the content (Anderson & Pempek, 2005). With foreground screen viewing, Kuhl and colleagues (2003) demonstrated the variation in learning when children were exposed to different types of screen content. In their experiments, English-speaking infants only learned Mandarin phonetic units with a live Mandarin speaker but did not learn the target phonetic contrasts through prerecorded audio-visual or audio-only materials. It is possible that even when children paid attention to the latter two recordings, they did not focus on the target

phonetic units but on something else in the video or audio. Thus, having an interactive live speaker could potentially guide their attention to the right place. Moreover, studies on literacy development with media showed that picture books with too many pop-ups did not result in better learning than traditional books, and it's difficult for children to stay focused if there is no feedback given for their performance (Tare et al., 2010).

Principle 6: Learning should be meaningful by linking the new to the old

Whether a learner can acquire target knowledge immediately after a learning session or gradually over multiple sessions, the ultimate goal is to understand the learning content conceptually and to apply it to the real world. As a large set of memorized but disconnected facts do not suffice for a competent learner (Bransford et al., 1999; Hirsh-Pasek et al., 2005), it is important to make connections between the learning targets and learners' existing knowledge. The two indicators of meaningful learning, according to Hirsh-Pasek et al (2005), are (1) the depth of processing the new information, which affects the learner's retention of that piece of information (Craik & Tulving, 1975), and (2) how much the learner can apply it to solve novel problems (Chi & Wylie, 2014; Goldstone & Day, 2012). To take letter learning as an example, showing a target word over and over again to a child may enhance their memory of spelling that word, but it is not meaningful learning. It is also not as effective as having the child watch a TV character solving a missing-letter problem in a target word spelling task (Fisch, 2004a), which links children's prior knowledge of a letter to a newly introduced target word.

Furthermore, the old information in meaningful linking does not only include prior knowledge but also includes familiarity with onscreen characters and with the learning content (Barr & Linebarger, 2017; Calvert & Richards, 2014; Gola et al., 2013). Krcmar (2010) reported in a study with 6-24-month-olds that young children imitated more actions from a video of a

socially meaningful character (i.e., mom) than from a video of a stranger. Calvert et al. (2007) suggested that children who related themselves more to the main character Dora in *Dora the Explorer,* achieved higher divergent thinking scores after viewing a target episode from *Dora.* However, it is worth noting that the social meaningfulness featured in both studies did not show any significant benefits for children's language and literacy-related learning. Children across all age cohorts in Krcmar's study (2010) showed the same amount of target word comprehension from a video of their mom and from a video of a stranger. In Calvert et al.'s study (2007), children who identified themselves more with Dora did now show more understanding of the central storyline. This is an important note to researchers and educators, as learning different subjects from screen media may require a different pedagogical approach. It might not be always appropriate to apply screen-media learning models for other areas (e.g., simple imitation, divergent thinking) to language learning.

Principle 7: Learning should be scaffolded.

Principles 3, 4, 5 and 6 were systematically discussed by Hirsh-Pasek et al. (2005) and Hassinger-Das et al. (2020) as the four psychological pillars that can best facilitate children's learning through educational media. However, learning could still be ineffective even when a child mentally actively participates and stays engaged in a meaningful and interactive learning session if the overall learning environment is not tailored to his/her learning level and pace. Hence Principle 7 emphasizes the structured and strategic assistance children can get throughout a target educational curriculum—scaffolding. The concept of scaffolding was originally discussed in Vygotsky's Zone of Proximal Development theory (1978), which proposed that an ideal learning situation puts children in between what they can independently accomplish and the support of more capable others.

When adapted to teaching/learning in practice, scaffolding can be manifested through verbal prompts, interactive activity with guidance and tailored learning materials (Mamun & Wright, 2020; Salmon et al., 2007; Wasik & Jacobi-Vessels, 2017). The benefits of scaffolded learning have been widely explored in children learning science and language. For example, Pratt and colleagues (1992) reported that the fifth graders whose parents utilized more scaffolding skills in math tutoring outperformed those whose parents did not. Baldwin et al. (1996) showed that 18-month to 20-month toddlers who had an adult joining the object labeling activity had better word comprehension scores than those who only listened to labeling audios. Similar findings were demonstrated in Roseberry et al.'s study (2009), where children between 30 and 42 months could only learn a novel verb presented by either a *Sesame Street* puppet or a real baby in a video when a live adult labeled the target verb with them. These studies highlighted the enormous support children could get from a scaffolding-focused partner.

Apart from the assistance from a live co-viewer, scaffolding can be applied to curriculum design to facilitate teaching. Crawley et al. (2002) and Mares (2006) highlighted the importance of repetition in educational curriculum design—multiple exposures to the same target (e.g., viewing the same episode in *Blue's Clues* and the same Disney film *The Sword in the Stone*, respectively) help children gradually understand but eventually master the content that they may not comprehend in the initial exposure. However, in Crawley et al's (2002) study, children with more exposure to the *Blue's Clues* episode appeared to only benefit in comprehending "series-typical" content (e.g., *Where does Steve keep the notebook?* P. 269), not episode-specific content. Moreover, Mares' study (2006) only revealed the facilitation of repetition on easy tasks (e.g., main character identification) but not complex ones (e.g., inferring a character's emotions).

These data suggest the potential of scaffolded repetition in educational programs to overcome contingent-feedback-deficits and facilitate children's learning from screen content, but they also reveal the caution of interpreting media learning results—preschool-aged children can learn (or not learn) differently from different types of screen content. Unfortunately, there is not enough empirical data on the effect of target repetition in language learning through educational media (Neuman et al., 2020), except for Linebarger and colleagues' study (2013) on preschoolers' vocabulary acquisition with *Martha Speaks.* Linebarger et al.'s data suggested that children's expressive vocabulary (not receptive) improved after 25 times of target viewing or less, but exposure to more than 25 times resulted in lower vocabulary scores.

Principle 8: Learning should be goal oriented with room for self-exploration.

In Hirsh-Pasek et al. (2005) paper, the importance of having a learning goal was discussed together with scaffolding. Here we consider it a separate principle from Principle 7 to feature the discussion between self-exploration and guided learning experience. Many studies indicate that learning without any instruction may not be as effective as guided learning in learning strategies for problem-solving, conservation or programming concepts (see a literature review by Mayer, 2004). Moreover, with the examination across 164 studies (participants' age range from children to adults) between discovery-based learning and explicit instruction-based learning, Alfieri et al.'s meta-analyses (2011) revealed that there were no benefits from unassisted discovery, and learners learn more with guidance (e.g., worked examples and explicit explanation). However, most of the studies with children in these meta-analyses focused on math and science domains or general problem-solving skills, not many examined children's verbal skills. Nevertheless, among a variety of different domains, verbal and social skills appear to most benefit from explicit instruction.

Goal-oriented learning does not mean there should be no room for joyful self-exploration. Several studies (see a review by Zosh et al., 2018) emphasized the importance of using diverse strategies that embedded structured teaching with playful context—children were more motivated to learn when there was a good balance between direct instruction and entertainment (Diamond, 2012). Most children's TV programs are designed to entertain the audience by having attractive characters doing fun activities that the target audience can relate to in real life. However, there is little guidance on how much instruction is optimal for teaching young children language with screen content, especially for grammar learning that involves more strict rules and patterns than vocabulary learning.

How can the principles be explicitly applied to educational animations?

Animations are widely used in children's educational media not only because they attract more attention from preschool-aged children than real images (Valkenburg & Vroone, 2004), but also because they offer tighter control of pedagogical approaches in content design (Takacs, 2015). For example, to emphasize a vocabulary learning target, *kick*, we can easily zoom in to an actor's foot and a soccer ball for a close look at the action. Exaggerated expressions and reactions from characters can guide children to the target more directly (e.g., showing an exaggerated facial expression when tasting something sour). Moreover, entertaining elements that are impossible to shoot with real life object can be easily emphasized in the animation (e.g., flying superhero dolls). It is also more convenient to manipulate diversities in the content to help children relate to their personal or cultural backgrounds (e.g., showing food from various cultures on the same table).

With such an appealing platform, how do we explicitly apply the principles synthesized from the Science of Learning to produce better educational program that teaches children

language? A number of developmental psychologists, language educators and curriculum experts have tried to do that—developing a series of English learning animations for child L2 learners between three and eight under the Age of Learning, Inc. In particular, the whole curriculum has adapted the scaffolding strategy to spread out common target vocabulary and grammars from easy to difficult. The overall learning path is created in modules of five episodes, culminating in a review episode that has an exciting activity with vocabulary and grammar from the previous episodes in the module (Principle 4—active, Principle 7—scaffolding). For each episode, there is clear learning objectives that cover target vocabulary and grammar at a proper level (Principle 8—goal-oriented). Target words appear at least five times with different objects, subjects, speakers and/or different tenses (Principle 1—computational, Principle 7—scaffolding). Among the main characters, there is a teacher (Ms. Jones) who encourages the students in the animation and the audience to practice learning targets with positive feedback (e.g., *Great job!* Principle 5—engaging). There are five student characters with different racial and ethnic backgrounds, personalities and proficiency levels, and their speech style and actions match their personalities (Principle 6—meaningful). As target terms are introduced in the episode, the characters occasionally talk to the audience, ask related questions and pause for answers (Principle 3—social interaction). Multiple speakers are used for each target word to encourage generalized discrimination of phonemic information and facilitate target word learning (Principle 6—meaningful; Principle 7—scaffolding).

Table 1 lists research studies or papers that discuss empirical evidence of L1 and/or L2 acquisition to support the Science of Learning principles. Also included are studies assessing children's language learning through educational media or APPs for each principle, as well as how the ABCmouse Early Learning Academy adapted or did not adapt these principles in their

curriculum design. Please note that Table 1 does not include an exhaustive list of studies but provides some key examples.

It is important to note, moreover, that although this series of English as a second language animations embedded as many principles as it could in its design, there are still restrictions and limitations to what was accomplished. For example, with pre-made animations, there is no contingent interaction between the audience and the characters, nor is there live feedback from the teacher. However, for the purpose of this dissertation, the aim is not to test a specific principle but to examine whether preschool-aged children can benefit from targeted language learning animations that are equipped with many pedagogical strategies in a short time of exposure.

Table 1 The New Science of Learning principles and how they are applied to research on language learning and educational media

Science of learning principles		**Evidence from empirical studies**		**What's applied/not applied to ABCmouse animation series?**	
		L1 & L2 acquisition	Language learning from TV/APPs	Applied	Not applied
Principle 1 Learning is computational	Infants/young children	Statistical learning based on the frequency and probabilistic distribution of the linguistic input. - Universal listener to language-specific phonetic discrimination (Kuhl et al., 2006) - Word segmentation (Romberg & Saffran, 2010; Saffran et al., 1996) - Object categorization (Fulkerson & Waxman, 2007) - Word learning in general benefits from learning the targets multiple times with a variety of accompanying words and objects (Waxman & Klibanoff, 2000).	N/A	All characters have native English-speaking actors pronouncing words clearly and at a pace appropriate to the level of English abilities. Type and Token frequency of target words are controlled. Target words appear at least five times. Target words appear with different grammatical subjects and objects, subjects, speakers and/or in different tenses (for verbs)	
	Older children/adults	Statistical learning in L2 acquisition - Auditory and visual statistical learning contribute to implicit L2 grammatical knowledge (Godfroid & Kim, 2021). Statistical learning in artificial vocabulary learning - Adult learners were capable of learning artificial words with cross-situational learning strategy (Yu & Smith, 2007)	N/A	Target words are presented with different exemplars and in different contexts to facilitate cross-situational mapping	N/A
Principle 2 Learning is neuro-cognitively linking self and other	Infants/young children	MEG evidence: being exposed to a target language activates multiple language areas in the brain: listening to speech activates the auditory area but also activates Broca's area related to speech production (Imada et al., 2006).	N/A	N/A	N/A
	Older children/adults	Adults showed language related motor neuron activation in Broca's area and primary motor cortex by viewing articulatory gestures of speakers' lip movements (Möttönen et al, 2005).	N/A	N/A	N/A

Principle 3 Learning is/should be social and interactive	Infants/young children	The responsiveness of a caregiver predicts infants' vocalization in general (Dunst et al., 2010). Children (18-24 months) showed vocabulary gain after 10-day with an interactive robot (Movellan et al., 2009) Children (24-36 months) improved word learning from a video with contingent feedback (Kirkorian et al., 2016) Children can learn L2 from social interaction with… - Non-human interactive robots (Tanaka et al., 2007) - Live speakers (Kuhl, 2007) - Speaker on-screen (Troseth et al., 2006)	Social interaction between children and on-screen characters (Elmo and viewers) helps with language learning in general (Calvert & Richards, 2014) Children (3-7 years) showed vocabulary and literacy gains with APPs *Martha Speaks* and *SuperWhy!,* which provided contigent feedback during learning (Chiong & Shuler, 2010).	The characters talk to the audience, ask questions and pause for answers. Ms. Jones and the children in the animations occasionally "break the fourth wall" by prompting a question or offer an opportunity for the audience to respond to the screen. Speakers include both the teacher and five child characters.	No contingent interaction with a character No interaction with real humans
	Older children/adults	Adult learners learned L2 vocabulary the most when they had in-game social interactions with native speakers compared to those who did not (Rankin et al., 2009)	N/A	N/A	N/A
Principle 4 Learning should be mentally active	Infants/young children	Children (3-5 years) resulted in higher vocabulary gains when they had to actively talk about the stories in their reading compared to those who only received silent listening. (Hargrave & Sénéchal, 2000)	N/A	N/A	Although the characters in the animation occasionally interact with the audience by asking them to practice along the way, there is no way to tell if the audience is fully mind-on.
	Older children/adults	College students learning English as a second language improved more on English communication skills by participating in multiple English related activities compared to their counterparts that only studied English in a traditional manner (Hung, 2015).	N/A	N/A	N/A
Principle 5 Learning should be engaging and focused	Infants/young children	Children (18-22 months) learned novel labels less well when the picture books had manipulative features (e.g., animals moving by touching) compared to the traditional picture book (Tare, 2010)	Children learned (3-5 years old) more story understanding and narrative skills from educational TV programs with a narrative (e.g., *Pinky Dinky Doo, Big Red*	Ms. Jones encourages the students (characters) to practice and provides positive feedback when they respond.	There is no live or contingent feedback responding to children who are watching and speaking to the video

		Showing a direct and immediate contrast between learners' error and the correct version by recasting is considered as one of the most effective types of feedback in L2 acquisition studies (Nicholas et al., 2001)	*Dog*) that could guide children's attention to the target elements than those without narrative embedded (e.g., *Zoboomafoo*) (Linebarger & Piotrowski, 2009)	Ms. Jones frequently expands and recasts child characters incomplete or fragmented sentences.	
	Older children/adults	N/A	N/A	N/A	N/A
Principle 6 Learning should be meaningful by linking the new to the old	Infants/young children	Young children's vocabulary learning should include words that appear to their interests and that are common in everyday world (Alexious & Konstantakis, 2009).	L1 Greek L2 English (4-5 years old) children showed significant receptive vocabulary improvement by watching *Dora the Explora* whose target word list include many commonly used English words in real life (Kokla, 2016).	Characters in the animations have different personalities and ethnic backgrounds. Their speech style and actions reflect their personalities and are consistent and predictable across episodes. The target word occurs in settings that involve child characters' real-life activities either in the classroom (e.g., bringing dishes from home) or outside (e.g., planting in the garden)	N/A
	Older children/adults	University-ESL students' reading comprehension skills benefit from reading materials that build on their prior cultural background knowledge (Johnson, 1982)	Children learn target word spelling more effectively when watching the character from *The Electric Company* solving a word spelling problem with a previously learned letter (Fisch, 2004a). Muslim and Buddhist ESL students (university level) outperformed the religion-neutral students in L2 comprehension after watching 20-minute TV broadcast on Islam and Buddhism (Markham, 2001).	N/A	N/A
Principle 7 Learning should be scaffolded	Infants/young children	Children (18-20 months) performed better at word comprehension when they had an adult in the labeling activity (guiding them learn), when compared to those who only listened to the labeling audio (Baldwin et al., 1996). Children (30-42 months) learned novel verbs from a video better when a live adult labeled the target verbs with them (Roseberry et al., 2009)	Children (2-5 years old) showed great improvement in expressive vocabulary after repeatedly reviewing the target content in *Blue's Clues* (Anderson et al., 2000). Children (3-8 years) benefited from using digital e-books with adult scaffolding on language and	The overall curriculum builds on previous episodes that are created in modules of five episodes, culminating in a review episode that has an exciting activity which uses vocabulary and grammar from the previous episodes in the module. Ms. Jones presents new vocabulary and grammar in settings that are familiar to children and scaffolds the complexity of	N/A

			literacy development (Savva, 2020)	each lesson, with capstone lessons at the end of every module of five lessons. Multiple speakers are used for each target word to encourage generalized discrimination of phonemic information and facilitate target word recognition.	
	Older children/adults	Children benefited from repeated readings that helped them gradually understand the same target content than being read a variety of books with the same info (Horst et al., 2011) Reinforcing mastery of content once it has become familiar is good for children's long-term learning. Once they learn a new concept, encountering them in different context may help (Fisch & Truglio, 2001)	Children show improved reading skills by watching *Blue's Clues* with familiar characters and previously mastered goals (Crawley et al., 2002)	The same six characters and teachers are used in all episodes with only occasional visits from other characters (e.g., Ms. Jones' family members, Zoe the zookeeper). Target words are first "sprinkled" in episodes prior to the episode in which they are taught and frequently appear in subsequent episodes as part of the scaffolding for new target words.	N/A
Principle 8 Learning should be goal-oriented with room for self-exploration	Infants/young children	N/A	N/A	Each episode and each module has clear learning objectives that cover target vocabulary and grammar at a proper level.	N/A
	Older children/adults	Children (4-6 graders) with low performance in reading benefited from guided instruction in story reading (Carnie & Kinder, 1985).	L1 Japanese L2 English learners (13 years) showed improvement in extensive reading when their self-directed learning behavior was assisted with goal-oriented reading application (Li et al., 2021)	N/A	N/A

Note. The seven principles listed are adapted from Meltzoff et al. (2009) and Hirsh-Pasek et al. (2015) papers. The empirical studies selected from the literature are roughly divided into "infants/young children" and "older children/adults" based on whether the participants were literate, namely able to read and write.

Chapter III What Do We Know About L1 and L2 Acquisition?

Comprehension vs. Production

It is natural to assume that children must understand a word before they can use it in a conventional manner (Bloom, 1974). Numerous empirical studies have also provided evidence to show that infants are able to identify the meaning of most words before being able to say them (e.g., Fernald et al., 2008). Moreover, recent studies (Bergelson, 2020; Bergelson & Swingley, 2015) have demonstrated that infants start to understand words in their native language around six months of age, suggesting a few months gap between early word comprehension and production. Similarly, in the literature on children's L2 vocabulary development, comprehension is generally achieved earlier than production, and receptive knowledge tends to improve faster than expressive knowledge (Fan, 2000; Kokla, 2016).

It is worth noting, however, that multiple studies in first language learning through audiovisual and digital media provide the opposite results. For example, Linebarger et al. (2003) found that preschool-aged children improved their expressive vocabulary, not receptive vocabulary, after viewing 16 target episodes in *Martha Speaks.* When assisted with various co-viewing techniques from parents, Strouse et al. (2013) showed that three-year-old children significantly improved their expressive vocabulary but not their general receptive vocabulary. As Neuman et al. (2020) suggested in their paper, these mixed findings could reflect different learning processes between comprehension and production. Thus, researchers should assess both aspects of learning when examining the efficacy of certain pedagogical approaches used in educational media.

Vocabulary vs. Grammar

Early vocabulary learning or lexical development involves fast-mapping of visual and acoustic stimuli. Essentially, the big challenge for young children to learn the meaning of a word in spoken language, which is often referred to as Quine's problem (1960), is to identify the correct association between the phonological information (i.e., what they hear) and the visual representation (i.e., what they see). For instance, when an adult holds an apple and says to an infant, *this is an apple* or [ðɪs ɪz ən ˈæpəl], ideal fast-mapping occurs when the infant maps the sound of "*apple"* or [ˈæpəl] to the fruit apple in the adult's hand, not the color of the apple, and not other phonetic information [ðɪs] or extraneous information in the environment. Numerous studies have shed light on the multiple factors that facilitate children's early vocabulary learning from a social and input perspective. For example, child-directed speech or motherese (Ferguson, 1964) features an exaggerated acoustic pattern with high pitch, simplified sentence structure, pause and stress on the target element(s) in the speech, as well as eye contact, which guides children's attention to the learning target (Dominey & Dodane, 2004; Falk, 2003). Moreover, social gestures like pointing direct children to the target visual reference without overwhelming them with complex surroundings (Goldin-Meadow, 2007).

Nouns vs. Verbs

Despite the assistance children receive from their social environment, lexical development itself encounters distinct theoretical challenges. There has been a long-standing debate on early noun and verb learning among children (Waxman et al., 2013). Many researchers consider nouns learning to be more stable and concrete in visual representation than verbs (Hollich et al., 2000). Therefore, children benefit from this noun advantage in fast-mapping and can learn nouns more easily than verbs (e.g., Gentner, 2006). Although this claim was challenged

in verb-rich languages, such as Korean and Chinese (Choi & Gopnik, 1995; Tardif et al., 1999), where children could learn verbs equally as well as nouns, it is robustly supported by empirical evidence among children speaking noun-friendly languages, including English (Bornstein et al., 2004; Kim et al., 2000). Nonetheless, novel verb-learning studies with preschool-aged children suggest that repetition of focused examples performed by fewer actors facilitates children's acquisition of verbs, even in English (Maguire et al., 2008). This finding also aligns with learning theory on children's relational category formation—children learn better with fewer variabilities during the exposure phase when it comes to identifying target relations in an event (Casasola, 2005). At the same time, visual variability is considered an important pedagogical tool for aiding both early noun- and verb-learning among preschoolers (Twomey et al., 2013; Twomey et al., 2014). To date, it is still unclear as to whether nouns and verbs are learned through similar mechanisms or whether there are differences in the cues that benefit nouns, verbs, and other parts of speech.

Grammatical Development

Compared to lexical development, grammatical acquisition begins relatively late around the second year of life or not until children are able to combine words (Brown, 1973; Hagan et al., 2008). According to L1 and L2 empirical data on grammatical acquisition, native speakers need five to seven years to fully master the common rules governing the use of morphemes, and second language learners may need much more practice, if not in a lifetime, to achieve native-like proficiency level (Jia & Fuse, 2007; Marinis & Chondrogianni, 2010). Further, cross-linguistic data demonstrated an interesting learning pattern that children do not acquire morphemes in their native language all at once but follow a consistent order within the language or within a group of similar languages (Baron et al., 2018; Fenson et al., 2007; Huang et al.,

2022; Xanthos et al., 2011). The early vs. late order of acquisition could be explained by a combination of linguistic features possessed by a grammatical morpheme, such as whether or not it is obligatory, the clearness of form-meaning mapping or regularity of its form across different grammatical or semantic variations (see Slobin,1982 for a more detailed discussion). To take English as an example, present progressive -*ing* and possessive -*'s* were reportedly acquired earlier than other morphemes such as regular past tense -*ed* (Brown 1973; Lahey et al., 1992; Fenson et al., 2007). That could be due to the fact that -*ing* and –*'s* are obligatory for marking an ongoing action and possession in English. They both have a straightforward form-meaning association and are regular, with no exceptions for specific words. On the contrary, the past tense -*ed* is less obligatory when certain syntactic rules are applied (e.g., *cry* in *he did not cry*), and many common verbs are marked in irregular past tense form (e.g., *flew* for fly, *went* for *go*). Similarly, in Mandarin Chinese, grammatical morphemes that carry more salient linguistic features appear to be acquired earlier than those that do not (e.g., possessive marker *de* vs. progressive aspect marker *zheng4*; Huang et al., 2022).

Whereas early vocabulary learning in a spoken language involves mapping between sounds and visual representations of an object or action, grammar learning requires the capacity to understand the relation between elements in a dynamic event, in which the mapping is more ambiguous and challenging (Golinkoff & Hirsh-Pasek, 2008). However, the current literature has not provided enough age-appropriate pedagogical guidance for teaching grammar, especially for child second-language learners. Adolescents and adults can benefit more from explicit analytic strategies, whereas children learning a second language draw more on implicit learning with many exposures (DeKeyser, 2018). Thus, the traditional approach to teaching adult learners (e.g., direct translation, grammatical explanations in L1, subtitles) may not be effective for children.

Instead, one approach is to adapt how children learn their first language to teach L2 grammar. In order to do this, child-directed speech would be important (Fernald et al., 1989) as would the addition of more linguistic cues and perceptual salience to the target relation (e.g., showing the result of a target action to emphasize the target verb; Brandone et al., 2007).

As pointed out by Gola and colleagues (2012), studies examining children's grammatical acquisition and educational media input are scarce and face several limitations, some of which include not having a direct assessment of the viewing content and examining screen content too generally. Moreover, most existing research involves correlational studies that do not suggest a positive outcome of learning grammar from media (Selnow & Bettinghaus, 1982). It is prudent, therefore, to withhold judgment about whether grammar can or cannot be learned from media until future experimental research has been conducted.

L1 Transfer in L2 Morphosyntax Acquisition

The role of L1 influence on L2 grammatical acquisition has been studied extensively, with numerous theoretical perspectives and empirical data. One line of research explores why some L2 learners of English across different language backgrounds have persistent difficulties in the use of certain grammatical morphemes such as past tense and articles (Zdorenko & Paradis, 2008). One possible explanation is that it is more likely for a learner to experience difficulty in learning a target grammatical form in L2 if there is no equivalent construction in their native language (Luk & Shirai, 2009). For example, Japanese and Korean speakers can acquire the English possessive marker -'s (as in *mom's apple*) rather easily (Izumi & Isahara, 2004; Pak, 1987), but it is very difficult for Spanish speakers to master this same grammatical feature (Andersen, 1978; Dulay & Burt, 1974). It is possible that marking possession in Japanese with a morpheme *no* and in Korean with a morpheme *uy* is similar to the use of -*'s* in English (e.g.,

mom-no apple, *mom-uy apple*) and therefore facilitates the acquisition, but because Spanish lacks an equivalent grammatical form, the acquisition is hindered.

Mandarin Chinese and Cantonese Chinese both have a straightforward possession marking system similar to Japanese and Korean. One would expect the native speakers of Mandarin and Cantonese to acquire the possessive marker -*'s* early in learning English. However, the expected phenomenon was only observed in Cantonese speakers (Mace-Matluck, 1979) but not in Mandarin speakers (Dulay & Burt, 1974).

Overall, L1 transfer on L2 grammar learning has variations across both morpheme types and learners with the same L1 background. Acquisition of the article system in English appears to be the most affected by L1 (e.g., Hawkins et al., 2006; Ionin, 2004), whereas the effect of L1 transfer on other types of grammatical construction is inconsistent. Moreover, Zdorenko and Paradis (2008) argued that the L1 transfer that exists in adults' learning might not be as strong in young children. Nevertheless, past tense learning in the above studies was consistently reported to be relatively late across Japanese, Korean, Spanish, Mandarin and Cantonese speakers, regardless of whether a similar concept exists in learners' native language. Finally, we are unable to explain why the findings of L1 transfer have discrepancies among similar languages (e.g., possessive learning in Mandarin and Cantonese). In order to do this, more longitudinal and more experimental studies are needed to reduce the effect of individual differences in cross-sectional correlational studies.

Despite the mixed evidence for L1 transfer on L2 grammatical acquisition, for the purpose of the current dissertation, I am interested in whether Mandarin-speaking preschoolers who have mastered possessive marking in Mandarin could learn the equivalent in English. Similarly, I am interested in whether the same children, who have no past tense marking system

in Mandarin, could learn this new grammatical construction and thereby benefit from targeted language-learning animations.

Specific Research Questions and Predictions

Research question 1: Are Children Able to Learn Language from "Educational" Media?

The first aim of the current study is to investigate whether preschool-aged children at a very early stage of learning English as a foreign language can learn the target language (i.e.., specific vocabulary and grammar at comprehension or production level) from minimal exposures (a single viewing of two 4-minute episodes at Time 1, further home viewing, and a second viewing session with another two 4-minute episodes one week later) to immersive English-language animations. In addition, we will examine whether exposure to these animations provides more generalized gains in English morphosyntax over the one-week video-viewing intervention. The current study is distinguished from studies investigating the relationship between second language acquisition and general media use because the current study has a particular focus on targeted language-teaching materials with clear teaching and learning objectives (e.g., learning verbs and possessives) that were developed following general principles of the Science of Learning (Hirsh-Pasek et al., 2015) and language-learning research.

Research question 2: What Exactly Can Children Learn from Targeted English Learning Animations?

The second aim of the current study is to examine whether the manipulation of specific target items (verbs vs. possessives) corresponds to performance gains in these same items, versus more general gains. We also examine both comprehension and production measures for both vocabulary and grammar.

Based on what the literature has suggested (e.g., Paradis, 2011), we hypothesize that children will perform better on vocabulary than grammar. We also expect that they will perform better on comprehension measures than on production measures. Past research on children learning English as both a first and a second language (e.g., Gentner, 1982; Jones et al., 1991; Schmidt, R. 1992) has also found that children tend to perform better on object words (common nouns) than action words (verbs) and on grammatical marking of possessives and the present progressive than either prepositions or past tense marking (Jia & Fuse, 2007; Paradis, 2005).

To ensure that we were able to test the effects of both word type and specific grammatical features, we introduced an experimental manipulation in which half the children viewed episodes which focused on Verbs and their tenses during the first phase of the experiment, and the other half viewed Possessive- (and Noun-) focused episodes during the first phase. Each group then received the opposite condition in the follow-up phase one week later so that we were able to examine immediate exposure effects between groups, with order counterbalanced, and conduct an additional within-subjects comparison which also had order counterbalanced for each word type.

Thus, in addition to the target-item assessment, we are also interested in the effects of our manipulation (exposure to Verbs vs. Possessives at Time 1) on general morphosyntax measures of English (e.g., whether having been exposed to tense markings with kick also helps children gain performance improvement in general English verb tense marking, possessives, and other measures of English grammar).

As we assessed children's language performance immediately after they watched a target set of animations, we expected children to have better performance on the target language items that appeared in the animation than those that were not yet introduced at Time 1. For example,

for children in the Verb condition that was exposed to verb-focused animations at Time 1, we expected them to outperform children in the Possessive condition on the verb items but to not do as well as children in the Possessive condition on the possessive items. In addition, we expected to see the children in the Verb condition to have higher scores on the target verb items than on the possessive items. Conversely, we expected children in the Possessive condition to do better on possessive items than on verb items. Finally, with extra condition-related target exposure at home, we expected children in the verb condition would have improved performance on verb items from Time 1 to Time 2.

Exploratory Research question 3: Does Exposure Frequency Affect Children's Learning from Targeted English Learning Animations?

As has been discussed in L1 and L2 acquisition studies (Bohman et al., 2010; Jia & Fuse, 2007; Paradis, 2011; Scheele et al., 2010), the frequency of linguistic input contributes to children's language development, even if it does not solely predict it. Therefore, the third aim of the current study is to examine the role of target viewing frequency on target language-learning outcomes. Specifically, we were interested in examining whether children who had extra target exposure at home would perform better on all target items when compared to those who had no extra exposure outside of the laboratory. We expected children with home viewing to outperform non-viewers on the target items at Time 2. For example, with more time watching the verb-focused animation at home, home-viewers in the verb condition should have more improvement on the verb items than the non-viewers in the same condition.

Chapter IV Methodologies

Participants

Children

71 children (aged between 4.78 and 6.73 years, M_{age} = 5.76, SD_{age} = .59; male = 37, female = 34) participated in the current study and were recruited at Tai Ping Xiang kindergarten in Nanjing, China, by five Nanjing Normal University graduate research assistants who had rich experience teaching and working with children on a regular basis. All children were fluent monolingual Chinese speakers and were at a very early stage of learning English as a second language. According to caregivers' reports, most children had some type of exposure to English through media or classes, but their overall English skills were rather minimal.

37 children (M_{age} = 5.76, SD_{age} = .53; male = 21, female = 16) were randomly assigned to the possessive-focused condition and 34 to the verb-focused condition (M_{age} = 5.75, SD_{age} = .66; male = 16, female = 18). These two experimental conditions decided the animation viewing sequence in the study. Specifically, children in verb-focused condition would watch verb-focused animations first in Session 1 and possessive-focused animations later in Session 2. Children in the possessive-focused condition would watch the same animations but in a reversed order.

Caregivers

71 caregivers were asked to fill out two questionnaires on both media use in Chinese and English, as well as demographic information in the family before children's animation

intervention began. It was reported that the education level for both fathers and mothers was from technical high school to Ph.D. degree (with the median for both a 4-year college degree).

Animated English-learning materials

There were seven episodes of four-minute-long animations chosen from ABCmouse.com Early Learning Academy (ABCmouse ELL), which is a digital education program for children aged 3-8 created by Age of Learning, Inc. According to Age of Learning's website, their English learning curriculum provides "a developmentally appropriate immersive English language environment" and "teaches English as their second language in the same natural sequence as children learn their native language" (https://www.ageoflearning.com/products). In the current study, we carefully chose two episodes for warm-up activities, two episodes for verb vocabulary and verb grammar training, two episodes for possessive vocabulary and possessive grammar training and one episode for watching at home only. Appendix A, B and C document the target vocabulary and grammar in each training and take-home episode, including type and token frequency of each target word and grammatical forms (Appendix A), detailed time stamps for each target in all episodes (Appendix B), as well as the target learning objectives from the ABCmouse ELL curriculum design (Appendix C).

Warm-up animations

The *Hello* video (Episode 1 of 100) and *Robot* (Episode 2 of 100) video were used as two warm-up videos that did not include any target vocabulary or grammar. They were played at the beginning of each lab session before children were asked to do any English assessments. In the *Hello* video, a female teacher, Ms. Jones, and six international children introduce themselves to each other by asking "what's your name?" and answering "my name is…" After the *Hello* video,

the examiners introduced themselves in English and asked for the participating child's name. In the *Robot* video, a robot brought in by one of the children joins Ms. Jones in teaching the same children how to say different parts of face, such as *nose, mouth, eyes*. After this video, the examiners interacted with children by asking them how to say parts on their face. Regardless of children's condition, the *Hello* video was only played in the first lab session and the *Robot* video was only played in the second lab session.

Verb-focused animations

Episode 17 and Episode 41 from the ABCmouse ELL animation series were chosen as the verb-focused videos. Appendix A and B give lists of action words in different verb forms. For example, in Episode 17, Ms. Jones teaches the children how to draw and she uses the verb *draw* multiple times—in future tense (e.g., *today we will draw*), present tense (e.g., *yes, you can draw*), past tense (e.g., *you drew very beautiful pictures*) and infinitive form (e.g., *what would you like to draw?*). In Episode 41, Ms. Jones teaches the children how to say *kick, throw, hit, catch* -- in future tense (e.g., *I will catch the ball*), present tense (e.g., *I can kick the soccer ball*), past tense (e.g., *I kicked the soccer ball*) and present progressive -*ing* (e.g., *I am throwing the baseball*). Given that there are no subtitles but only phonetic information in the linguistic input, we added how many times a child would hear a target verb in its bare sound form (e.g., /drɔ/ for draw) and in its -*ing* /ɪŋ/ form (e.g., /ˈdrɔɪŋ/ for drawing). Note that in both verb-focused videos, possessive pronouns—*my* and *our*—also occur a few times, and we counted each as possessive exposure in the final analysis.

Possessive-focused animations

Episode 12 and Episode 15 from the ABCmouse ELL animation series were chosen as the possessive-focused videos. These two episodes emphasize the use of possessive grammar -*'s*, such as in *brother's, David's,* as well as possessive pronouns such as *my, your, his, her.* In Episode 12, Ms. Jones introduces her family to her students (*e.g., This is my father*) and in Episode 15, the students show each other their drawings (*e.g., his dog, my dog, her flowers*). Appendix A demonstrates the token frequency for each possessive pronoun and grammar form. Note that the verb *draw* occurs twice in its present tense in Episode 15, and we counted each as verb exposure in the final analysis.

Take-home animations

Episode 52 was chosen as a "swing episode" for children to watch at home after their first lab visit. The plot of this episode is about Ms. Jones taking the whole class out to plant seeds in a garden. There are multiple possessive pronouns (e.g., *my, our, your*) and verbs in different grammatical forms: future tense (e.g., *We will plant our seeds*), present tense (e.g., *Can you dig a hole?*), past tense (e.g., *We planted the seeds*), present progressive (e.g., *We are planting the seeds*), infinitive (e.g., *We need to dig a hole for your seeds*) and gerund form (e.g., *This is digging*). The amount of exposure to possessives and verbs is 17 and 20 times, respectively.

Child language measures

Multiple language measures in Chinese and English were used to test children's language proficiency under both the Possessive and Verb conditions. Prior to the experiment, we measured children's general language proficiency on vocabulary and morphosyntax in both Chinese and English. Table 1 shows the overall performance of these pre-tests separated by possessive and

Verb condition. In addition, Welch's t tests suggest that at the base level (i.e., prior to our study), there were no significant differences in the children's performance between the two conditions.

General vocabulary measure.

A reduced version of the 5th edition of Peabody Picture Vocabulary Test (Dunn, 2019) was used for testing children's vocabulary in English with 52 items in total. For each item, children were shown four pictures, three of which were distractors, and were asked to point to the picture they think we are asking for (e.g., *show me ball*). A Chinese version of the PPVT vocabulary test was also given to test children's vocabulary comprehension in Chinese (Lu & Liu, 1998) with 125 items in total. For both the Chinese and the English PPVT, assessment ended as soon as a child made five consecutive errors. Cronbach's alpha shows the reliability of this measure is $\alpha = .81$. Note that we only measured children's general English and Chinese vocabulary once (at Time 1) to get the baseline of children's vocabulary comprehension in both languages.

General Morphosyntax measure.

There were 8 types of grammatical markings that we measured in English morphosyntax: possessive *'s,* simple present, copula, negation, present progressive *-ing*, past progressive *-ing*, past tense, and passive form. Before testing each grammatical form, we provided two demonstration items to familiarize children with the prompting format (e.g., Examiner: *Look, this is a doctor, and this is the doctor's watch… This is a lady, and this is the…* Child: *lady's watch*). Colored pictures were used for each testing item and the examiners were trained to point to the exact object or part when and where they were supposed to (e.g., pointing to the lady's watch when saying *lady's watch*). There were 2 testing items for each grammatical form, except for

possessive and past tense (the main foci of the study), where we added a third item with made-up words (e.g., *dap's scarf, the boy gorbed*). Finally, we tested children's general English morphosyntax twice—once at the beginning of the first lab session and after the English vocabulary measure, and once during the second lab session after the warm-up video *Robot*. To balance the difficulty level for items, we created two alternative forms of the English morphosyntax test: half of the children were tested with Form A in the first lab session and Form B in the second; the other half were tested in reversed order. As with the vocabulary measures, assessment ended if/when a child made five consecutive errors. Cronbach's alpha shows the reliability of this measure is $\alpha = .70$. Please refer to Table 2 for the total scores and detailed scores for each grammar type.

The Chinese morphosyntax measure was made in the same format as the English version. Based on the common grammar categories in Chinese, there were six types of grammar measured in the study: possessive *de* (4 items), negation *mei2/bie2/bu4* (6 items), present progressive *zai4/zheng4zai4* (3 items), aspect marker *le/guo4* (4 items), classifiers (6 items), active *ba3* and passive *bei4* form (4 items). Again, assessment ended if/when a child made five consecutive errors. Cronbach's alpha shows the reliability of this measure is $\alpha = .75$. According to the longitudinal study in Huang et al.'s work (2022), Chinese-speaking children developed most of their grammatical markers before 6 years old. Hence, we expected these children to achieve a high level of proficiency in the Chinese morphosyntax test, and only measured their Chinese morphosyntax once. Please refer to Table 1 for descriptive statistics on this measure.

Target-Item test.

As soon as children finished watching the assigned pair of animations, a Target-Item test was given to test the immediate learning effect. The items were divided into vocabulary (18

items) and grammar sections (6 items for past tense grammar, 6 items for possessive grammar). For each item, there were pictures (i.e., cropped screenshots) from the four training animations to help children remember the scene or object. For comprehension questions, the child was asked to point to the correct picture (e.g., *show me kicking*) from a set of four pictures. For production items, children were prompted to answer a question (*e.g., what did Sara do to the baseball?*). To balance the difficulty level of items in the grammar section, we created two alternative forms (Form A and Form B) for this Target-Item test. Half of the children who were given Form A on the English morphosyntax test were also tested with Form A version of Target-Item test in the first session and Form B in the second session. The other half were tested in reverse order.

The scoring for Target-Item test was not based on how many items the children got correct but based on the following variables of interest: (1) receptive items (i.e., possessive pronouns and verbs). (2) productive items (i.e., verbs, verb past tense grammar, possessive pronouns, possessive grammar –*'s*). All receptive items were given with an instruction: *now we are going to look at some pictures together, and you are going to help me find the pictures I am looking for*. All productive items were asked with a prompt in the same format as the general English morphosyntax test (*Ms. Jones was drawing a picture. Now she is done. What did she do? She...*). Cronbach's alpha shows the reliability of this measure is $\alpha = .79$. Table 3 shows the means and standard deviations of the overall and detailed scores for Target-Item test by time and condition.

Procedure

All 71 children were invited to a child-friendly and quiet activity room at the preschool they went to for two one-hour-long lab sessions. In the first session, after a short interaction in Mandarin with an examiner who was also a native speaker of Mandarin, each child was given

vocabulary and grammar tests in Mandarin. After the Mandarin tests, the examiner played the warm-up video—*Hello*, in which Ms. Jones and her six students introduce each other with their names. This episode ended with all six student characters looking at the camera and asking for the audience's name. Children who participated in this study were able to respond with their name to the examiner's question "What is your name?". After the *Hello* video, the examiner measured the child's proficiency in English vocabulary and morphosyntax.

Following the language assessments in English, the animation intervention began: each child was randomly assigned to either the Verb condition (N = 34) or the Possessive condition (N = 37). The children in the Verb condition watched two verb-focused animations that emphasized both target verbs and past tense grammar while the children in the Possessive condition watched two possessive-focused animations that emphasized both possessive pronouns and possessive grammar –'s. Finally, the Target-Item test was given to examine whether there was an immediate learning effect on both target vocabulary and grammar items. The first session ended with all children taking three animations home with them—two were the exact animations they watched with the examiner (e.g., Verb condition got verb-focused animations) and the other one was a new video (Episode 52—the swing episode) that included both past tense and possessive grammar. We asked the caregivers to record and report how many times their children watched each animation at home before the second session.

There was one-week interval between the two lab sessions. At the beginning of the second session, each child was shown the second animated episode of the series, which introduced a comic character "Robot" and parts of the face (e.g., *nose, mouth*, etc.). For warm-up purposes, the robot video did not contain any of the target vocabulary or grammar. Immediately after this, we re-assessed children's general English morphosyntax. This post-test allowed us to

understand whether there were any long-term effects on English morphosyntax from the video exposure as well as to control for any external exposure that may have improved performance for children who did not receive the target animation. After this general post-test, each child was given two more target animations to watch. However, this time, the children in the Verb condition watched the two possessive animations that the children in the Possessive condition watched in the first session, and vice versa. In other words, the difference between the Verb condition and possessive condition in the study was the watching sequence among verb and possessive animations. The Target-Item test was given again to examine the immediate learning effect (Appendix D demonstrates an overall design of the current study).

In summary, the current study first measured children's general language performance in vocabulary and grammar in both English and Chinese prior to the animation intervention. We then gave two experimental conditions 4 animations to watch with reversed viewing sequence between verb focused and possessive focused animations and tested their immediate learning effect each time. Children from each condition were encouraged to watch their assigned animations as many times as they wanted at home with a new take-home animation before their second lab session. Figure 1 demonstrates the overall design of the study.

Analytical plan

The first research question—*Are children able to learn language from "educational" media?* is the most general question of this study. The answer to this question focuses on children's learning of target English vocabulary and grammar, as well as their improvement in general English language knowledge, as measured by the morphosyntax test, after a week-long animation intervention. In other words, would preschool-aged Mandarin-speaking children who initially had very minimal knowledge of English show improved performance at Time 2, relative

to Time 1, after viewing the animations, without considering the order of target exposure condition or the amount of exposure children had to the videos?

The second research question—*What exactly can children learn from targeted English language-learning animations?* puts the focus on *what* English language knowledge children can learn from the animations. Based on the vocabulary and grammar emphasis in the training videos, we divided the target language learning into six main categories: (1) verb vocabulary comprehension, (2) verb vocabulary production, (3) verb past tense grammar production, (4) possessive grammar -*'s* comprehension, (5) possessive pronoun production, and (6) possessive grammar -'s production. We first examined this with Time 1 data, where we expected that children in the Verb condition (exposed to verbs, but not possessives at Time 1) would do better (a) on verb items than children in the Possessive condition would do on the same items (between-subjects comparison); and (b) on verb items than they would on possessive items (within-subjects comparison); and). We expected the opposite to hold true for children in the Possessive condition. Namely, children in the Possessive condition (exposed to possessives, but not verbs at Time 1) would do better (a) on possessive items than children in the Verb condition would do on the same items (between-subjects comparison); and (b) on possessive items than they would on verb items (within-subjects comparison).

Because we ultimately exposed children to *both* types of items over the course of the one-week intervention, we also examined their performance on each target type as a function of their exposure condition over the two training sessions. In order to do this, we computed a set of linear mixed-effects models (LMMs) with random subjects and random testing forms in R with lme4 (Bates et al., 2015) and emmeans (Lenth, 2023) package to examine whether condition (Verb vs. Possessive condition), test time (first lab session vs second lab session) and target language

category contributed to any testing score improvement from Time 1 to Time 2 while controlling for children's age and initial English vocabulary score. Furthermore, because the training videos included many more verb tense forms (e.g., present progressive, past progressive, etc.) than verb simple past tense, we also examined whether and how much children improved for all verb tense items, in addition to possessive items, on the general test of English morphosyntax. We expected that the children in the Verb condition would show greater improvement on the verb tense items, whereas the children in the Possessive condition would improve more on the possessive items.

The last research question—*Does target exposure frequency matter?* takes home viewing data into consideration. We divided children from each condition into two groups—viewers, with extra target exposure at home between the two lab sessions, and non-viewers, with no parentally-reported extra target exposure and compared their learning results at Time 2. We again computed a set of linear mixed-effects models with random subjects and random testing forms in R to examine the exposure effect on both the Target-Item test and the general English morphosyntax test. Although these analyses are only exploratory, we also used the number of episodes parents reported as a predictor and discuss the potential and limitations of this approach.

Chapter V Results

In this section, we will tackle each of the three research questions with the above-mentioned analytical plans. In these analyses, we used linear mixed-effects models (LMM) instead of traditional linear models for multiple reasons. The main benefit for computing a LMM is that this type of modeling can handle inter-dependent datasets (i.e., in our study, each participant provided multiple data points from different language tests and was assessed by two alternative forms for each test) to account for by-subject and by-item variation in the analyses (Cnaan et al., 1997). This allows us to control random effects due to individual differences (the children varied widely in both their initial Mandarin- and English-language abilities) and unintentional design bias from alternative forms or versions of a test.

Demographic factors and performance on language measures

Children's age, maternal education, and general language scores in English and Chinese are presented in Table 2 for the whole sample as well as the subsamples for each viewing condition. As can be seen in Table 2, children did not differ on the initial language measures in the two conditions. Children's age was strongly and positively correlated with both English and Chinese vocabulary (Pearson's r = .37, $p < .001$ and r = .46, $p < .001$, respectively). Age was also moderately correlated with general English morphosyntax pre-test (Time 1) total scores (r = .29, $p < .05$) and Target-Item test total pre- (Time 1) and post- (Time 2) scores (r = .27, $p < .05$ with pre-test, r = .25, $p < .05$ with post-test). Mother's education level was positively correlated with children's general English morphosyntax Time 1 score (r = .32, $p <$. 04). In addition, children's English vocabulary at Time 1 (prior to the intervention) was significantly correlated with all

language tests at both Time 1 and Time 2, including Chinese vocabulary and morphosyntax (r = .39, $p < .001$ and r = .34, $p < .001$, respectively), English morphosyntax (r = .55, $p < .001$ and r = .47, $p < .001$, for Time 1 and Time 2, respectively), and Target-Item test (r = .59, $p < .001$ and r = .44, $p < .001$, for Time 1 and Time 2, respectively). Therefore, further analyses control for children's age and general English vocabulary when examining the efficacy of the animation intervention. In addition, considering that the children's English proficiency in this sample was very low, and their score improvement from Time 1 to Time 2 could be minimal, we use less conservative Tukey's HSD instead of the Bonferroni correction method for multiple pairwise comparisons to be more sensitive to and better capture significant differences in performance (Saville, 2015).

Table 2 Demographics and pre-test results for children's English and Chinese vocabulary and grammar

Variables	All Children	Possessive Condition	Verb Condition
Sample (N)	71	37	34
Females (N)	34	16	18
Father Education (median) (Technical high school - PhD)	4-year college	4-year college	4-year college
Mother Education (median) (Technical high school - PhD)	4-year college	4-year college	4-year college
Age (4.78 – 6.73 years)	5.76 (.59)	5.76 (.53)	5.75 (.66)
English Morphosyntax (Max = 18, % correct)	.08 (.12)	.09 (.13)	.07 (.11)
Chinese Morphosyntax (Max = 27, % correct)	.90 (.07)	.91 (.06)	.88 (.07)
English Vocabulary (max = 52, % correct)	.36 (.18)	.34 (.18)	.37 (.19)
Chinese Vocabulary (max = 125, % correct)	.65 (.15)	.67 (.13)	.63 (.17)

Note. Welch's t-test was computed to compare between verb and possessive conditions for each of the scores listed above prior to the animation intervention. There were no significant differences between the two conditions.

Research question 1: Are children able to learn language from "educational" media?

As can be seen in Table 3, children showed significantly better performance on all comprehension items but not all production items after viewing the animations. In order to examine these effects more closely, we ran separate LMMs to analyze all children's (regardless of their viewing condition) comprehension and production scores for both target vocabulary and

grammar. As mentioned above, age and initial English vocabulary at Time 1 were controlled as fixed effects whereas participants and testing forms (Form A vs. Form B) were controlled as random effects (see Table 3 LMM test statistics for Time 1 and Time 2 comparisons).

Table 3 Target-Item Test raw mean (SD) percent correct scores across all children (N = 71)

Variables	**Time 1**	**Time 2**	**LMM test statistics**	***p* values**
Comprehension Items				
Object nouns	.51 (.28)	.61 (.25)	$F_{(1, 69)} = 7.57$	$p < .01$
Prepositions	.35 (.38)	.46 (.40)	$F_{(1, 69)} = 5.31$	$p < .05$
Possessive grammar -*'s*	.41 (.33)	.56 (.30)	$F_{(1, 69)} = 8.99$	$p < .01$
Verbs vocabulary	.25 (.23)	.37 (.26)	$F_{(1, 69)} = 13.49$	$p < .001$
Production Items				
Possessive pronouns	.05 (.05)	.10 (.10)	$F_{(1, 69)} = 5.31$	$p < .05$
Possessive grammar -'s	.15 (.29)	.25 (.36)	$F_{(1, 69)} = 5.90$	$p < .05$
Verbs vocabulary	.06 (.14)	.13 (.24)	$F_{(1, 69)} = 11.01$	$p < .01$
Verb past tense grammar	.02 (.06)	.03 (.08)	$F_{(1, 69)} = 1.72$	$p = .19$

Note. Linear mixed-effects models with random subjects and random test forms were computed for examining testing time effect for each section while controlling for children's age, initial general English vocabulary.

Comprehension of target vocabulary and grammar

There was a significant performance improvement in comprehension of both target object noun and target verb vocabulary from Time 1 to Time 2, as can be seen in Table 3 and Figure 1. In addition, the average accuracy for object nouns (M = .51, SD = .28 at Time 1; M = .61, SD = .26 at Time 2), was higher than for verbs (M = .26, SD = .24 at Time 1; M = .38, SD = .27 at

Time 2) at both Time 1 (Δb = .25, $t_{(1064)}$ = 6.36, p < .001). and Time 2 (Δb = .23, $t_{(1064)}$ = 5.95, p < .001).

Among grammar items, possessive grammar -*'s* and prepositions (*in* and *on*) both showed significant improvement from Time 1 to Time 2 ($F_{(1, 69)}$ = 8.99, p < .01 for possessive grammar -*'s*; $F_{(1, 69)}$ = 5.31, p < .05 for prepositions). In addition, the comprehension of possessive -*'s* (M = .41, SD = .33 at Time 1; M = .56, SD = .30 at Time 2) was not significantly different from that of prepositions (M = .35, SD = .38 at Time 1; M = .46, SD = .40 at Time 2) at both time points.

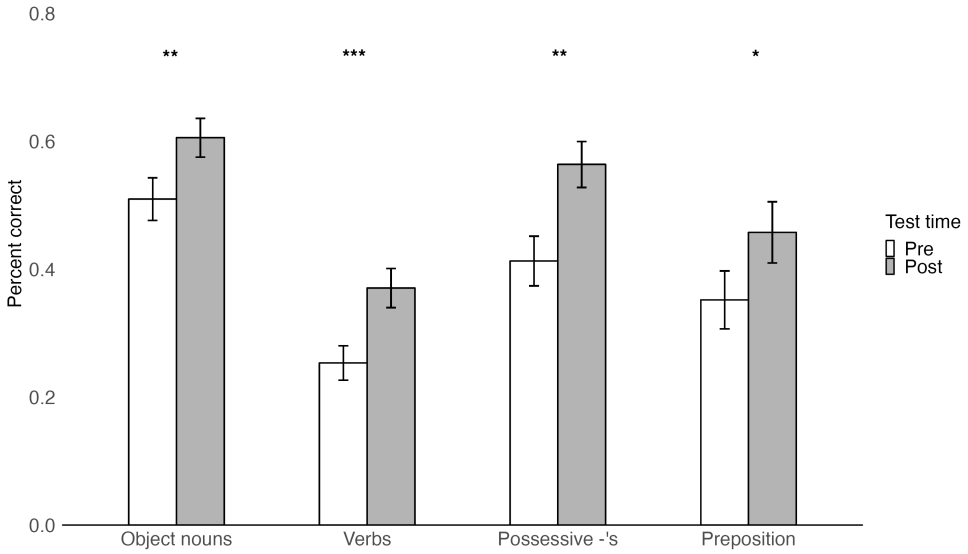


Figure 1 Children's comprehension scores of target vocabulary and grammar at Time 1 and Time 2

Production of target vocabulary and grammar

As can be seen in Table 3 and Figure 2, summing across conditions, children's production of possessive pronouns (M = .05, SD = .05 at Time 1; M = .10, SD = .10 at Time 2) and verb vocabulary (M = .06, SD = .14 at Time 1; M = .13, SD = .24 at Time 2), both showed significantly higher scores on the Target-Item tests from Time 1 to Time 2 ($F_{(1, 69)}$ = 11.01, p < .01 for verb vocabulary; $F_{(1, 69)}$ = 5.31, p < .05 for possessive pronouns). In addition, children's

performance on verb vocabulary items did not show any statistical differences from that on the possessive pronoun items at both time points.

The production of possessive grammar -'s also improved from Time 1 to Time 2 (M = .15, SD = .29 at Time 1; M = .25, SD = .36 at Time 2), $F_{(1, 70)} = 5.90$, $p < .05$, but the .01 accuracy score improvement on the production of verb past tense was not statistically significant (M = .02, SD = .06 at Time 1; M = .03, SD = .08). In fact, production of verb past tense appeared to be the most difficult target for children to learn among all target items.

Despite the significant improvement for possessive grammar -'s from Time 1 to Time 2 on the Target-Item test, children did not show improvement on the *general* English morphosyntax test (M = .15, SD = .29 at Time 1; M = .21, SD = .34 at Time 2), nor on the subset of verb past tense items (M = .01, SD = .04 at Time 1; M = .01, SD = .05 at Time 2), as can be seen in Table 2a. Similar to the Target-Item test, children's verb past tense grammar production on the general English morphosyntax test received the lowest score among all general English grammar items, suggesting it was more difficult for Mandarin-speaking children to learn than the possessive grammar -'s. Nonetheless, we also examined improvement on verb tense marking more generally, since the verbs that children were exposed to occurred in multiple tenses (Appendix B) and there were more exposures in *other* tenses, combined, than past tense per se, even though past tense grammar was the target of the Verb condition videos. When looking at all verb tenses combined, we found a Time effect ($F_{(1, 69)} = 7.35$, $p < .01$; marginal $R^2 = .24$, conditional $R^2 = .64$), indicating there was a performance improvement on verb tense marking among all children from Time 1 (M = .04, SD = .10) to Time 2, (M = .07, SD = .11).

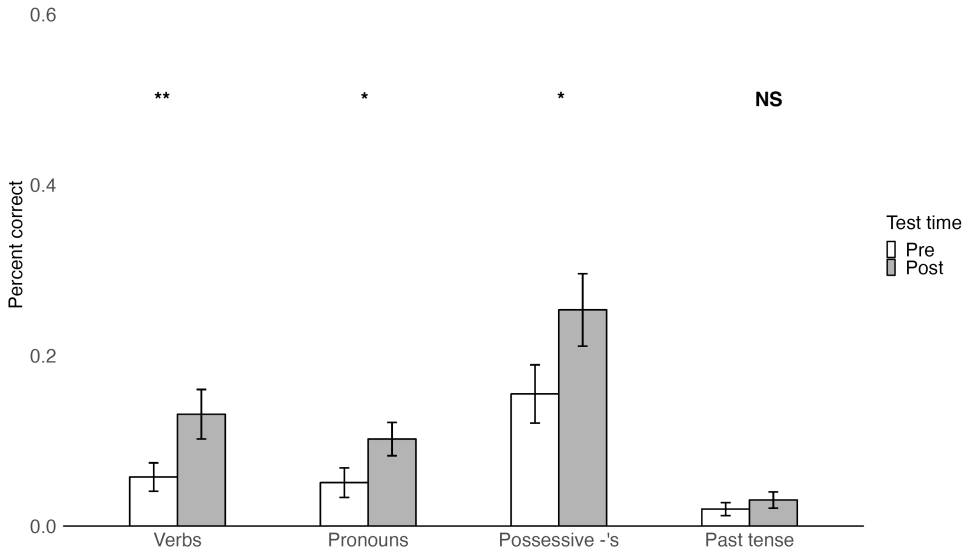


Figure 2 Children's production scores of target vocabulary and grammar at Time 1 and Time 2

Research question 2: What *exactly* can children learn from targeted English-learning animations?

In the analyses for this question, we looked at the results separated by the two viewing conditions to examine three things: (1) the specific effects on items that received the most input in each condition, and (2) the overall effect of the intervention by condition (Time 1 Target-Item test vs. Time 2 Target-Item test). In addition, we restrict our focus to six target types: verb comprehension, verb production, verb past tense grammar production, possessive -'s comprehension, possessive pronoun production, and possessive grammar -'s production. Detailed results can be found in Table 5 and Figure 3.

Verb condition vs. Possessive condition (at Time 1) – between-subjects comparisons

For this analysis, we controlled the fixed effects of age and children's initial English vocabulary, initial general possessive and verb past tense grammar, and the random effects of

subjects and testing forms. With the *emmeans* package (Lenth, 2023) in R, we were able to compare estimated beta values for multiple pair-wise comparisons between target language categories.

Verb measures. At Time 1, when comparing between viewing conditions, we found a significant effect of the type of measure ($F_{(2, 138)} = 61.27$, $p < .001$), but not the effect of viewing condition ($F_{(1, 64)} = .15$, $p = .67$). The marginal model effect size $R^2 = .42$ and the conditional effect size $R^2 = .50$. According to the LMMs, children's performance on verb items had the same pattern across conditions. For example, in both conditions, children had significantly higher scores for verb comprehension than verb production and verb past tense production (Verb condition: $b_{(\text{verb comprehension})} - b_{(\text{verb production})} = .19$, $t_{(138)} = 5.89$, $p < .001$; $b_{(\text{verb comprehension})} - b_{(\text{verb past tense production})} = .23$, $t_{(138)} = 7.17$, $p < .001$. Possessive condition: $b_{(\text{verb comprehension})} - b_{(\text{verb production})} = .20$, $t_{(138)} = 6.35$, $p < .001$; $b_{(\text{verb comprehension})} - b_{(\text{verb past tense production})} = .23$, $t_{(138)} = 7.42$, $p < .001$). Furthermore, children in both conditions performed statistically the same in verb production and verb past tense production.

Possessive measures. For children's performance on the possessive items, we again found a significant effect of the target possessive measure ($F_{(2, 138)} = 43.33$, $p < .001$), but there was no condition effect ($F_{(1, 64)} = 2.36$, $p = .13$). The marginal model effect size $R^2 = .38$ and the conditional effect size $R^2 = .42$. Possessive comprehension received the highest scores in both conditions (Verb condition: $b_{(\text{possessive comprehension})} - b_{(\text{possessive pronoun production})} = .29$, $t_{(138)} = 5.09$, $p < .001$; $b_{(\text{possessive comprehension})} - b_{(\text{possessive -'s production})} = .18$, $t_{(138)} = 3.16$, $p < .01$. Possessive condition: $b_{(\text{possessive comprehension})} - b_{(\text{possessive pronoun production})} = .43$, $t_{(138)} = 7.77$, $p < .001$; $b_{(\text{possessive comprehension})} - b_{(\text{possessive -'s production})} = .33$, $t_{(138)} = 5.98$, $p < .001$). There was no significant difference between

children's performance in possessive pronoun production and possessive grammar -'s production in both conditions.

Zero exposure vs. two-episode exposure (at Time 1) – within-subjects comparisons

For this analysis, we controlled the fixed effects of age and children's initial English vocabulary, initial general possessive and verb past tense grammar, and the random effects of subjects and testing forms. With the *lmerTest* package (Kuznetsova et al., 2017) in R, we were able to compare estimated beta values for multiple pair-wise comparisons between target language categories for each Condition. Means and Standard Deviations for each measure, as well as within-subject comparisons across measures, are presented in Table 4 for each condition.

Verb condition. At Time 1, within the Verb condition, there was a significant main effect of the target language category ($F_{(5, 165)} = 13.86$, $p < .001$) with a marginal effect size $R^2 = .32$ (include fixed effects only) and conditional effect size $R^2 = .34$ (including both fixed and random effects). Verb comprehension and possessive -'s comprehension had statistically the highest scores among all six target types, as shown in Table 4. However, verb comprehension and possessive comprehension were not statistically different from each other. Moreover, there were no differences amongst Verb or Possessive Production, or Verb or Possessive grammar, and this was contrary to expectations. However, as mentioned above, verbs received the lowest scores on the pretest and the English morphosyntax test. Thus, it is not clear whether the animation intervention was not successful for verbs from these data or whether the overall low levels of performance for verbs made it impossible to see improvement from looking at Time 1 and this condition alone.

Possessive condition. At Time 1, within the possessive condition, we again found a significant main effect of the target language category ($F_{(5, 180)} = 32.54$, $p < .001$) with a marginal

effect size R^2 = .48 (include fixed effects only) and conditional effect size R^2 = .54 (including both fixed and random effects). Children in the possessive condition showed the best performance for possessive -'s comprehension. Verb comprehension and possessive grammar -'s production received the second-highest scores and were not significantly different from each other. The lowest scores were obtained for possessive pronoun production, verb production and verb past tense grammar, none of which were significantly different from each other.

In summary, our analyses suggested that compared to production, comprehension items are relatively easier to learn from targeted language learning animation, which mirrors the overall learning results in the current and previous studies. For example, the possessive condition showed improvement for possessive comprehension, relative to verb comprehension, and possessive grammar, relative to verb grammar. Scores in the Verb condition were less clear, with no significant differences in the verb items vs. possessive items, even though verb categories received lower scores on the overall English morphosyntax test given before the animations were shown.

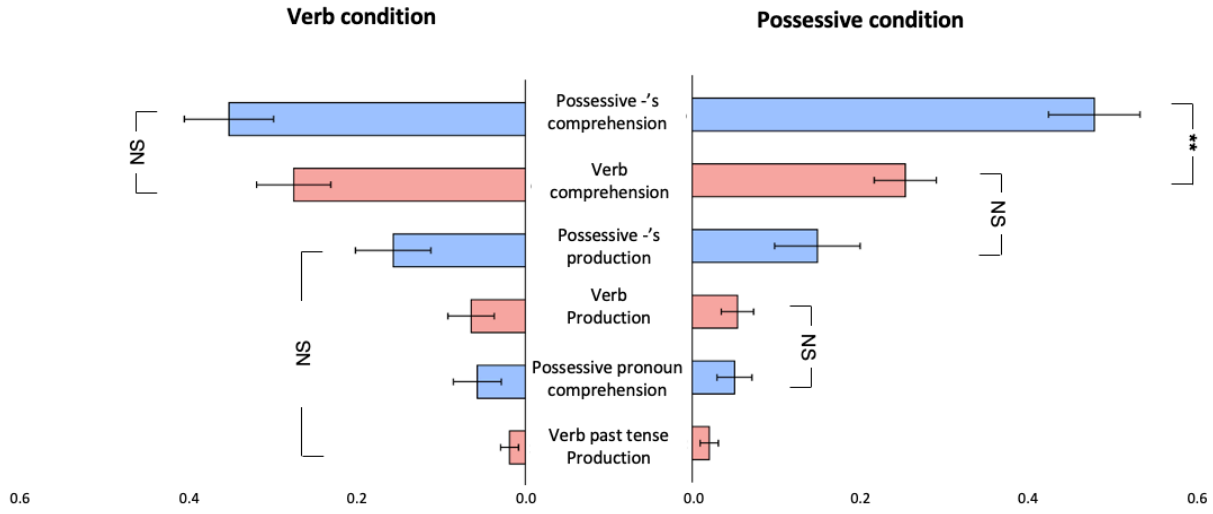


Figure 3 Children's performance scores of target vocabulary (red) and grammar (blue) at Time 1 by conditions

Table 4 Multiple pairwise comparisons of correct percentage scores among target language categories in verb and possessive condition

Time 1 Verb condition (N = 34)		**Time 1 Possessive condition (N = 37)**	
Target category	**Mean (SD)**	**Target category**	**Mean (SD)**
Possessive - *'s* comprehension	.34 (.31)[a]	Possessive - *'s* comprehension	.48 (.33)[A]
Verb vocabulary comprehension	.25 (.23)[a,b]		
		Verb vocabulary comprehension	.25 (.22)[B]
Possessive - *'s* production	.16 (.27)[b,c]	Possessive - *'s* production	.15 (.31)[B,C]
Verb vocabulary production	.06 (.16)[c]		
Possessive pronoun production	.05 (.05)[c]	Verb vocabulary production	.05 (.12)[C,D]
Verb past tense Production	.02 (.06)[c]	Possessive pronoun production	.05 (.05)[C,D,E]
		Verb past tense production	.02 (.07)[E,F]

Note. Means of target categories that are statistically different have different superscripts. Items that share a superscript with at least one common letter are not statistically different using the Tukey HSD correction method for multiple pairwise comparisons. Comparisons for the Verb condition are indicated with lower case letters and comparisons for the Possessive condition are indicated with upper case letters.

Time 1 vs. Time 2 Comparisons by Condition

In addition to looking at the immediate exposure conditions, we also looked at the difference in children's performance at Time 1 and Time 2. According to our model, there was a significant time effect ($F_{(1, 758)} = 36.35$, $p < .001$), target language category effect ($F_{(5, 758)} = 101.16$, $p < .001$) and a significant interaction between time and condition ($F_{(1, 758)} = 5.21$, $p <$

.05). The marginal model effect size R^2 = .38 (fixed effects only) and the conditional effect size R^2 = .50 (fixed and random effects combined).

Note that Table 5 provides basic LMM statistics with which we only examined the time effect for specific morphosyntactic categories on the general test of English morphosyntax given before the videos were shown at Time 1 and Time 2. Figures 4 and 5 show children's performance (correct percentage scores) for Time 1 and Time 2 by condition for the target items on the Target-Item test.

Table 5 General English morphosyntax raw mean (SD) percent correct scores by condition

Variables	**Pre-test**	**Post-test**	**LMM test statistics**	***p* values**
		Possessive Condition (N = 37)		
Possessive -*'s*	.13 (.27)	.24 (.36)	$F_{(1, 35)} = 4.87$	$p < .05$
Simple present[1]	.11 (.24)	.20 (.32)	$F_{(1, 35)} = 3.43$	$p = .07$
Copula	.10 (.22)	.03 (.11)	$F_{(1, 35)} = 5.29$	$p < .05$
Negatives	.09 (.25)	.07 (.24)	$F_{(1, 35)} = .26$	$p = .61$
Present progressive[1]	.05 (.14)	.09 (.19)	$F_{(1, 35)} = 2.59$	$p = .12$
Past progressive[1]	.03 (.12)	.05 (.13)	$F_{(1, 35)} = .62$	$p = .44$
Past[1]	.01 (.05)	.01 (.05)	$F_{(1, 69)} = .00$	$p = 1.00$
Passive	.05 (.16)	.09 (.20)	$F_{(1, 70)} = 1.05$	$p = .31$
Verb tenses	.05 (.10)	.09 (.12)	$F_{(1, 35)} = 5.71$	$p < .05$
Total	.07 (.13)	.10 (.13)	$F_{(1, 34)} = 3.47$	$p = .07$
		Verb condition (N = 34)		
Possessive -*'s*	.18 (.32)	.17 (.31)	$F_{(1, 32)} = .02$	$p = .88$
Simple present[1]	.09 (.23)	.15 (.23)	$F_{(1, 32)} = 2.29$	$p = .14$
Copula	.08 (.16)	.04 (.11)	$F_{(1, 63)} = .1.78$	$p = .19$
Negatives	.10 (.21)	.07 (.17)	$F_{(1, 32)} = .86$	$p = .36$
Present progressive[1]	.04 (.19)	.06 (.20)	$F_{(1, 32)} = .32$	$p = .58$
Past progressive[1]	.01 (.06)	.01 (.06)	$F_{(1, 63)} = .00$	$p = 1.00$
Past[1]	.00 (.03)	.01 (.06)	$F_{(1, 63)} = .19$	$p = .66$
Passive	.03 (.12)	.04 (.13)	$F_{(1, 32)} = .11$	$p = .74$
Verb tenses	.04 (.10)	.06 (.10)	$F_{(1, 32)} = 2.17$	$p = .15$

Total	.07 (.12)	.07 (.10)	$F_{(1, 33)} = .12$	$p = .73$

Note. 1. Welch's t-test was computed to compare pre-tests between verb and possessive conditions for each of the scores listed above. There were no significant differences between the two conditions prior to the animation intervention. 2. Linear mixed-effects models with random subjects and random test forms were computed for examining testing time effect for each grammar section while controlling for children's age, initial general English vocabulary.

Possessive vs. Verb condition on Target-Item test

Among six target language categories on the Target-Item test, children in the possessive condition (Figure 4) had statistically significant performance gain for verb comprehension items only ($t_{(758)} = -.2.90$, $p < .01$) after the intervention. Although possessive comprehension items did not appear to improve from the intervention, children's scores on these items remained the highest at both Time 1 (M = .48, SD = .33) and Time 2 (M = .53, SD = .32) and were significantly higher than the verb comprehension items at Time 1 (Table 4) followed by verb comprehension scores (M = .25, SD = .22 at Time 1; M = .39, SD = .23 at Time 2). Possessive

pronoun production, verb production and verb past tense production were reported to have the lowest scores at both Time 1 and Time 2 and were not statistically different from each other.

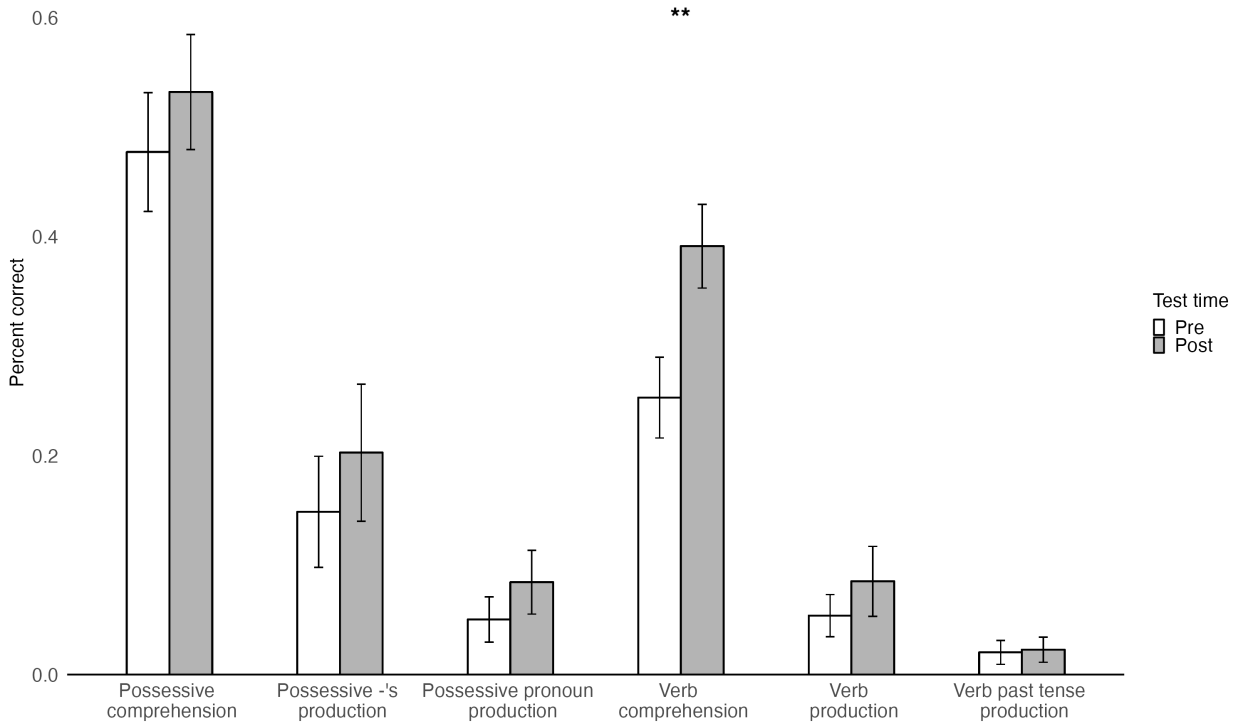


Figure 4 Possessive condition performance scores of six target language categories at Time 1 and Time 2

On the other hand, children in the Verb condition (Figure 5) had increased scores on different target language categories compared to the possessive condition. Higher scores were observed for possessive comprehension items ($t_{(758)} = -5.19$, $p < .001$), possessive grammar *-'s* production items ($t_{(758)} = -2.94$, $p < .01$) as well as verb production items ($t_{(758)} = -2.39$, $p < .05$). Similar to the possessive condition, possessive comprehension items and verb comprehension items received the higher scores than other target items at both Time 1 (M = .34, SD = .31 for possessive comprehension; M = .25, SD = .23 for verb comprehension) and Time 2 (M = .60 SD = .28 for possessive comprehension; M = .35, SD = .28 for verb comprehension). Verb past tense

production and possessive pronoun production were statistically no different from each other receiving the lowest scores at both Time 1 and Time 2.

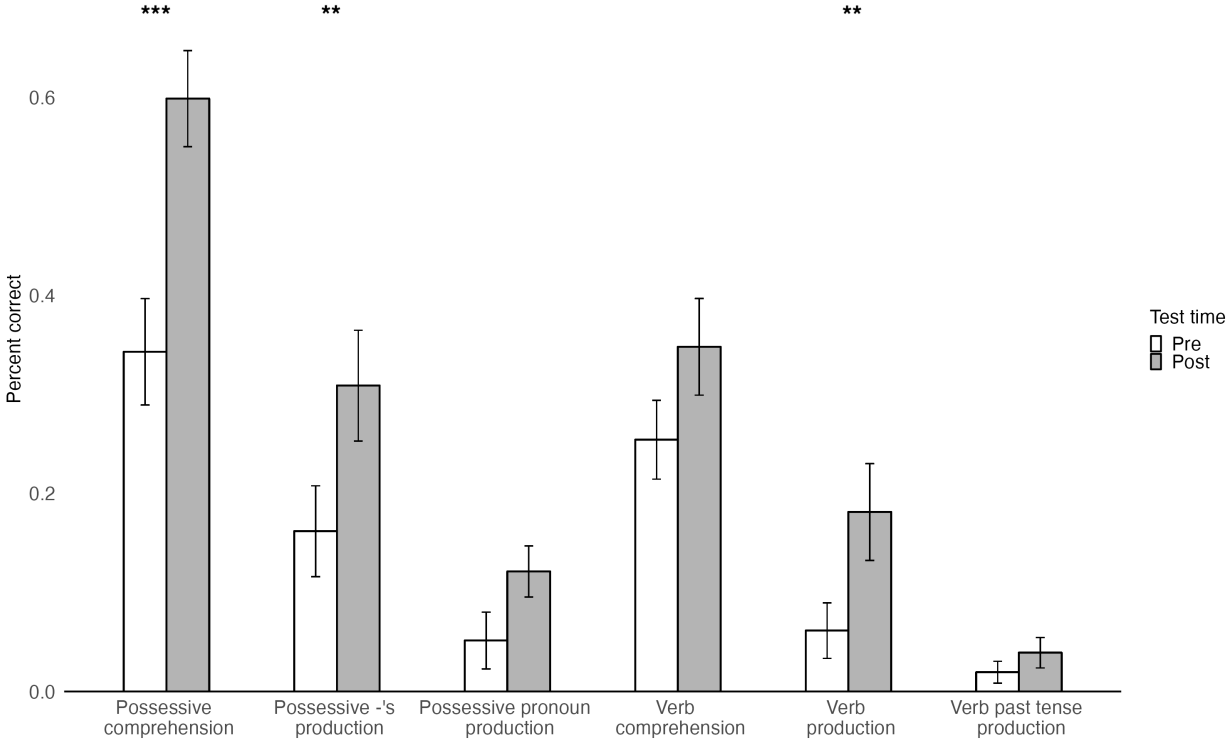


Figure 5 Verb condition performance scores of six target language categories at Time 1 and Time 2

Possessive vs. Verb condition on general English morphosyntax test

As English past tense grammar had the lowest score on the Target-Item test, the score for this grammar category on the general English morphosyntax test was also very low – essentially at floor. Consistent with the performance on the Target-Item test, both conditions had no significant improvement on verb past tense grammar production. In addition, we have not found any significant learning effect from Time 1 to Time 2 for overall verb tense marking (all verb tense items combined) or possessive grammar -'s.

Research question 3: Does relative exposure frequency affect children's learning from English language learning animations?

There was a week-long period between the children's two lab visits, in which parents were asked to show children three episodes of condition-related animations at home (i.e., Verb condition should watch two verb-focused videos and a neutral video. Possessive condition should watch two possessive-focused videos and a neutral video) and log how many times each animation episode was viewed. Among 71 children, only 44 home-viewing reports were collected: 68 % of the children in the Verb condition watched an average of 12.83 condition-related episodes (SD = 10.48) and 57% in the possessive condition watched an average of 13.76 episodes (SD = 10.35). It was not clear whether the rest of the children did not have any exposure to the target episodes, or whether their parents simply forgot to monitor and report.

Next, we computed another set of LMMs to compare children's performance in target vocabulary and grammar measures at Time 2. We were particularly interested in whether children in each condition with extra target exposure at home would perform better on the target language categories than those who had no extra exposure at all. In these LMMs, we controlled for age and initial English vocabulary as fixed effects, and participants and alternative testing forms as random effects. Our aim was to test the effects of home exposure and target language categories and their interaction.

Viewers vs. non-viewers in the Verb condition at Time 2

Overall, the results from the LMM for the Verb condition showed there was a significant effect of target language category ($F_{(2, 64)} = 19.93$, $p < .001$) and a significant interaction between the target category and home-viewing exposure ($F_{(2, 64)} = 3.78$, $p < .05$). However, the main effect of home-viewing exposure was not significant ($F_{(1, 30)} = 3.91$, $p = .06$). The marginal

model effect size R^2 was .35 (fixed effects only) and the conditional effect size R^2 was .65 (fixed and random effects combined).

Target verb vocabulary and grammar measures (within-subjects). It is interesting to note that within the viewer group in the Verb condition, we observed a clear pattern of significant learning differences—children performed the best on the verb comprehension items, followed by verb production ($b_{(\text{verb comprehension})} - b_{(\text{verb production})} = .20$, $t_{(64)} = 4.20$, $p < .001$) with verb past tense grammar showing the lowest performance ($b_{(\text{verb production})} - b_{(\text{verb past tense})} = .18$, $t_{(64)} = 3.75$, $p < .01$). In the non-viewer group, the Verb condition children showed no performance differences across the three verb measures.

Viewers vs. non-viewers in the possessive condition at Time 2

According to the results of LMM for the Possessive condition, there was only a significant effect of the target language category only ($F_{(2, 70)} = 28.37$, $p < .001$). The marginal model effect size was $R^2 = .38$ (fixed effects only) and the conditional effect size was $R^2 = .50$. This suggests the children in the Possessive condition did not benefit from extra target exposure at home for possessive -*'s* comprehension, possessive pronoun production, or possessive -*'s* production.

Target possessive vocabulary and grammar measures (within subjects). Overall, children with and without home viewing performed the best on possessive comprehension items (e.g., $b_{(\text{possessive -'s comprehension})} - b_{(\text{possessive pronoun production})} = .45$, $t_{(70)} = 5.52$, $p < .001$ for viewers; $b_{(\text{possessive -'s comprehension})} - b_{(\text{possessive pronoun production})} = .45$, $t_{(70)} = 4.84$, $p < .001$ for non-viewers). In both groups, possessive pronouns and possessive -*'s* production were not statistically different from each other.

Chapter VI General Discussion

As decades of research in Science of Learning (Benassi et al., 2014; Hirsh-Pasek et al., 2015) inspires and guides educators and curriculum experts to develop language learning programs with the support of modern technology, children around the world have the potential to access to more high-quality second language learning resources outside of school. The current study examined one English-learning program that fall into this design category and carried out a one-week-long intervention experiment with 71 preschool-aged Mandarin-speaking children in China. Our overarching question was whether young learners are able to learn English as a second language from targeted language-learning animations. In this section, we will summarize our findings for each research question, as well as discuss the contributions and limitations of the current study.

RQ1: Are children able to learn language from "educational" media?

By comparing children's Time 1 and Time 2 performances, we found that children improved on multiple target language categories, including all the target comprehension items (e.g., verb vocabulary and possessive grammar -'s) and most target production items (e.g., verb vocabulary, possessive pronoun vocabulary and possessive grammar -'s). Verb past tense production was the only language category that did not show significant improvement from Time 1 to Time 2 (see Table 3). This suggests, for both vocabulary and grammar, young Mandarin-speaking children are able to improve comprehension skills from targeted English language-learning animations. This does not only apply to object-noun vocabulary but also to verb vocabulary, which has been argued to be extremely difficult to learn from passive exposure (e.g.,

Imai et al., 2008). Moreover, targeted grammar exposure also appears to be effective for children's grammar comprehension (e.g., possessive grammar). As far as production skills are concerned, the current study showed children's capability of learning both vocabulary and grammar, although past tense grammar, one of the most difficult grammatical categories for both first- and second-language learners of English (as discussed in Dulay & Burt, 1974; Krok & Leonard, 2015) did not improve from Time 1 to Time 2.

RQ2: *What* exactly can children learn from targeted English-learning animations?

In the analyses for the second research question, we examined children's target learning by the two viewing conditions. Contrary to our predictions, neither the Verb nor the Possessive items showed condition-specific effects. Instead, both conditions showed very similar patterns for verb and possessive test items. More specifically, both conditions showed better performance on verb comprehension than verb production and verb past tense production. In addition, both conditions showed better performance on possessive comprehension than possessive pronoun and possessive grammar production. These patterns are consistent with the findings in the within-subjects analysis for Time 1 exposures and performance.

As can be seen in Table 4, children in the Verb condition did not align with what we predicted. Instead of showing improved performance for verbs, children in this condition did equally well on both possessive -'s comprehension and verb comprehension at Time 1). Children in the possessive condition, however, showed a relatively clearer learning pattern that fit our prediction. Namely, children in this condition had the best scores on the possessive -'s comprehension score, with scores for the rest of the categories following a similar order to those in the verb condition.

These findings suggest two main things. First, overall, comprehension measures show more effective learning from language-learning animations than production measures. Second, for these Mandarin-speaking children, possessive grammar -'s appears to be less challenging to learn, even when compared to verb vocabulary learning. This could also be explained by the similarities between possessive grammar -'s in English and its equivalent grammatical morpheme *de* in Mandarin-Chinese (Huang et al., 2022). The possessive morpheme *de* in Mandarin was also found to be one of the earliest markers to be acquired among Mandarin speakers (Chao, 1968; Erbaugh, 1992). Note also that -'s in English is also one of the first morphemes to be learned by children learning English as a first language (Brown, 1971; Fenson et al., 2007).

In addition to looking at the immediate learning results, we also compared children's performance from Time 1 to Time 2. In the next set of analyses, we looked at three factors: (1) viewing condition, (2) testing time and (3) target language categories. Although there were no main effects for viewing condition, our model suggests that the children who participated in the one-week-long intervention improved over time (testing time effect), that the improvement varied by target category, and that the time effects differed between conditions (time x condition effects), which can be seen in the comparison between Figures 1 and 2. Although there was no condition effect at Time 1, children in the Verb condition appeared to have benefited more from this one-week-long intervention as we observed improved performance on three different target categories (possessive comprehension, possessive grammar -'s production and verb production) for children in the Verb condition. Children in the possessive condition, in contrast, only showed a time effect on one target category—verb comprehension, which reflected their immediate exposure to verb target items from the videos they were shown at the Time 2 laboratory visit.

One possible explanation for such contrast might be related to the interaction between how difficult a target is to learn and *when* we measure a target. From the previous analyses, we argued that Mandarin-speaking children might feel more intuitive or easier to learn possessive grammar -'s because a similar form for possessive marking exists in Mandarin, whereas verb tense marking in English is relatively more complex and there is little transfer from the relatively simpler marking of verb aspect in Mandarin (Paradis et al., 2016; Yang & Lyster, 2010). Thus, when the Verb condition watched possessive-focused animations at Time 2, they showed learning effects on target possessive items, especially for possessive -'s comprehension and possessive -'s production as well as sustained learning for verb vocabulary items. Possessive pronouns are a relatively new concept for Mandarin-speaking children (these are marked with personal pronouns and the general possessive marker "de") and thus may take more exposures or a longer duration of exposure to show improvement than was possible after the two possessive-focused episodes shown at Time 2. Nonetheless, this immediate learning pattern was aligned with what we observed in the Possessive condition at Time 1—after watching the possessive-focused videos, children in the Possessive condition performed better on possessive -'s comprehension and production than possessive pronoun production. Furthermore, it is not clear to us why the Verb condition only showed improvement on verb production items but not on verb comprehension and verb past tense production. It's possible that the amount of target verb exposure at home facilitated children's learning of verb vocabulary production, but was not enough to support effective verb past tense grammar. Mandarin-speaking children have consistently been reported to struggle to learn past tense in English (e.g., Jia & Fuse, 2007; Cameron & Lee, 1999). In addition, although the immediate exposure to target verbs at Time 1

helped children with verb comprehension, the delayed measure could have impacted children's performance even when they had extra target exposure at home.

The same argument can be used to explain why the Possessive condition only showed significant improvement in the verb comprehension items from Time 1 to Time 2. Consistent with the immediate learning results observed in the Verb condition at Time 1, the Possessive condition also showed an immediate learning effect on verb comprehension items after watching the verb-focused videos at Time 2, indicating that verb comprehension learning from targeted animation was easier than verb production items. However, it is not clear again why there was no improvement for all three possessive categories. It could be due to it was a delayed assessment, or the children did not have enough target exposure at home between the two lab visits. To examine this possibility, we must look at children's home-viewing data for further analysis.

Nevertheless, for children's grammar learning, it is difficult to conclude whether the Verb condition truly learned the possessive -'s production because the general English morphosyntax test tells a different story showing no significant improvement in possessive -'s production (see Figure 3). This further highlights the importance of assessing children's learning in multiple ways to avoid misleading or biased intrastation from one test. When we compared all 71 children's performance on the general test of English morphosyntax where the testing pictures and the testing targets were not from the animation, there were no significant learning results for either possessive grammar -'s or verb past tense production (Table 2a). However, when we examined the effects of home viewing, we did find improvement for children who watched the Verb condition videos at home, as described below.

RQ3: Does media exposure frequency affect children's learning?

To better answer the role of target exposure frequency, we further examined children's home-viewing data. First, we compared the children with extra target exposure at home (home-viewers) with those who did not (non-viewers) within each condition at Time 2. We expected children who viewed more target animations at home would perform better on the target language categories (e.g., in the verb condition, home-viewers would do better on verb items than the non-viewers). Consistent with our prediction, the viewers in the verb condition showed better performance, but on the verb vocabulary comprehension items only. In contrast, home-viewer and non-viewer children in the Possessive condition did not show any performance differences. These findings suggest that children might benefit from more target exposure, but the effect varies across target language categories. In our study, the more verb exposure children had, the better they performed on verb vocabulary comprehension, but possessive exposure did not appear to help with their possessive learning.

We then took a step further to analyze the effect of target exposure frequency by counting the exact number of episodes home-viewers had at home. Within a sub-sample of 45 children whose parents submitted viewing data, we found that the more verb exposure children had overall, the better they learned verb tense marking. This relation was even more salient for general possessive -'s marking (see Figure 4a). However, for the Possessive condition, the total amount of target possessive exposure children had did not appear to have an effect on the general English morphosyntax measures of possessive -'s marking or verb tense marking (see Figure 4b). Nonetheless, it is important to interpret these findings with caution because they may suffer from reduced-power analysis—the number of children we had for each condition was below 25 (N = 24 for the Verb condition and N = 21 for the Possessive condition).

Furthermore, this sub-sample also provided slightly different learning results for the Verb condition if we compared it with the full-sample (home-viewer and non-viewer combined) analysis done for RQ 2. From the analysis of RQ2, children in the verb condition showed improvement for verb production only among all three verb items. However, the same analysis with home-viewers only showed the verb condition improved on not only verb production but also verb comprehension. This difference might be due to the possibility that children who had home-viewing data were generally more motivated to learn English hence they showed more learning on the target categories from the training videos. Nevertheless, we did not observe a different learning pattern of possessive items between the home-viewer-only possessive condition and the all-children possessive condition.

Interestingly, when we compare all 71 children's performance on the general test of English morphosyntax where the testing pictures and the testing targets are not from the animation, there were no significant learning results for either possessive grammar -'s or verb past tense production (Table 2a). This finding could either suggest that the general test English morphosyntax itself did not capture children's learning of possessive grammar -'s production or that the children were not able to apply the possessive grammar knowledge to the more general English test.

Future directions

Examining whether an APP has educational value is essentially examining whether the pedagogical instruction utilized in the APP is effective. This type of examination requires bridging two lines of research—learning material content analysis and empirical data analysis. The former analyzes the pedagogical elements in the target educational content and the latter

gathers real-world data from learners. However, to the best of our knowledge, there is a lack of communication between the two in the current literature.

On the one hand, several studies have conducted detailed content analyses of some popular bilingual shows teaching English as L2 (e.g., Dora the Explorer, Go, Diego, Go!, Maya & Miguel, Handy Manny, Nina's World) by examining implicit and explicit vocabulary-learning cues as well as pedagogical strategies of L2 instruction (e.g., noun tokens, common daily rote phrases, L1-L2 language order, audio-video synchronization, target word repetition, and translation equivalent). These content analyses demonstrated a wide range of differences in pedagogical instruction of L2 vocabulary learning. For example, it was reported that Dora the Explorer and Go, Diego, Go! provided more receptive and expressive participation opportunities for children compared to the others, and Handy Manny was reported to have the most verbal definitions of target words (e.g., Rivera Pérez et al., 2021; Wong & Neuman, 2019). With respect to more general learning principles, an APP is most likely to promote learning from children if the educational program includes cognitively active learning experiences and opportunities to engage with meaningful content. Social interaction and goal-oriented learning guidelines are also important (Hirsh-Pasek et al., 2015). A recent study by Meyer et al. (2021) developed a coding scheme based on these principles to evaluate APP quality. Surprisingly, the study revealed that more than half of the top educational APPs (free and paid) in the Google Play, Apple APP stores as well as the most popular APPs used by preschoolers in their study did not meet the requirements of what the Science of Learning suggests.

On the other hand, very few studies examining the efficacy of language learning media provide empirical evidence at the same detailed level as those content analysis studies. Most of the empirical studies were focused on rather general language gains or did not measure content-

related learning outcomes. Without fine-grained analyses matching assessment to specific curricular goals, little is known about whether any pedagogical approach plays an effective role in facilitating children's L2 acquisition.

Fortunately, we have seen increasing attention paid to educational APPs over the years, among which language learning is reported to be one of the most popular learning topics (Madigan et al., 2020). This line of research, however, has suffered several critical weaknesses in the study design and analysis method and thus is not able to provide consistent and reliable results. We suggest that future research should focus on the following.

Be more specific in testing target-item learning

Learning is a rather broad, abstract, and complex concept. An educational APP can have educational value but provide no effective target learning at the same time. For example, a child may learn color words in an APP designed to teach them how to count but fails at counting after using it. Likewise, a child may incidentally learn some numerical knowledge in a video targeted at animal words but is not able to match target animal pictures with the right names after watching the video. Therefore, examining the efficacy of a language learning program is not just a question of *whether, how much, how fast* a child can learn, but most importantly *what* can be learned under the provided pedagogical instruction.

Social-cognitive skills (e.g., emotion regulation, theory of mind, risky behaviors) can have a wide variety of definitions and be tested in various forms and situations, measuring basic language skills is relatively more straightforward. For instance, to test children's understanding of the word *apple*, we need to represent an apple instead of a pear as a testing target. Similarly, a child who does not mark third person singular in a required linguistic context would not be considered to have mastered such grammar. Unfortunately, most language programs or APPs

research with learning data have not achieved a decent level of specificity in terms of matching learning stimuli and testing items. For instance, many studies consider "effective language learning" as improved performance on the total score of standardized measures only. This type of analysis does not offer in-depth explanation of what aspects of the learning material is effective or useful beyond a rough correlation.

Our study results have shown how sensitive children could be to testing items with and without familiar context (e.g., familiar characters in the animation viewing test vs. random people characters in the general English morphosyntax test). If the current study were to test possessive grammar -'s in the general English morphosyntax test only, we would not discover that children in the Verb condition could understand and produce possessive -'s when children were given familiar scenes from the animation, even though they were not able to apply such rule to general language use.

More advanced statistical modeling

As important as experimental design, advanced data analysis methods can help us with uncontrollable random factors in the design that might impact learning results, especially when the amount of learning is limited. In the current study, all the children were at a very low proficiency level of English. The learning outcomes were positive, but in some cases quite minimal. Proper advanced modeling (e.g., mixed-effects modeling) allows us to detect even small improvements among children with low proficiency by controlling both random and fixed effects at the same time. The results could be very different in traditional linear models.

In the current literature, studies rarely take advantage of the power of mixed-effects modeling. Paired samples t-tests are commonly used to report post-pre learning outcomes even in complex longitudinal studies (e.g., Bang & Collins, 2020; Kokla, 2016). More surprisingly, not

many studies even utilize multiple linear regression to control for important cofounding variables such as children's language proficiency prior to any intervention and target language exposure outside of lab. As our study suggests, after a short period of target intervention, young L2 learners are capable of learning both vocabulary and grammar, researchers should not underestimate the impact of daily media exposure in the target language.

Conclusions and Limitations

In summary, our study suggests that preschool-aged Mandarin-speaking L2 learners can benefit from targeted English learning materials that are embedded with good pedagogical instruction for both vocabulary and grammar acquisition for both comprehension and production. Consistent with findings in both first and second language acquisition (e.g., Acarlar & Johnston, 2011; Fan 2000; Kokla, 2016; Slobin, 1982), some target language categories appear to be easier to acquire than others. For instance, possessive –'s grammar appeared to be easier for Mandarin-speaking children to acquire, whereas verb past tense grammar received no performance improvement during our one-week-long intervention. In addition, comprehension measures for both vocabulary and grammar appear to be easier for preschool-aged children to demonstrate their learning, at least when exposed to animations that did not involve interaction or further practice. More generally, standardized English tests that reflect children's general language knowledge might not be sensitive enough to capture children's learning from targeted language learning materials.

There are a few limitations in this study that also need to be mentioned. The first limitation comes from the disadvantage of using an existing commercial product. While polished commercial APPs offer great user experience and pretty animations, we have very little control of the learning stimuli. For example, amongst the verbs that were taught, there were more

irregular (*draw*, *throw, catch, put, hit, dig*) than regular past tense verbs (*kick*, plant) in the verb-focused episodes and verb-possessive-balanced (take-home) episodes. Morpho-phonologically speaking, there is quite a variation—irregular past tense verbs change differently in sounds (*draw* vs. *hit*) and regular past tense verbs seem to have rules for how *-ed* should sound (*kicked vs. planted*). Not only that, the number of times a target verb is presented in past tense form varies. Even though the frequency could be treated as fixed effects in a mixed-effects model, it would have been ideal to control the frequency difference between the target verbs—for instance, manipulating the frequency to reflect real life language use frequency. This is and will continue to be a big disadvantage for researchers when examining the efficacy of a real-life learning product by using the actual product. A possible solution could be involving research experts in the curriculum development and animation design process, allowing them to conduct studies and collect data along the way. Another potential solution requires researchers to work closely with statisticians in order to tweak statistical models in a way that supports fine-grained data analyses. With advanced statistical knowledge, we might be able to tear complex features apart and pinpoint what matters the most in early second language learning.

The second limitation in the current study is the low rate of parental home-viewing reports. By losing nearly 38% of the data in a sample of 71 children with two conditions, we lost statistical power in the third analysis. In fact, it is common to have low quality parental report. Linebarger and Piotrowski (2009) discussed in their longitudinal TV intervention study that the response rate from parents to report children's at-home TV viewing habits was still quite low after multiple attempts, and they eliminated these data from the final analysis. Although many language learning APPs have built-in progress tracking feature that automatically records learners' login times, learning time as well as what has been viewed, we rarely see studies take

advantage of such technology for research purposes. Nevertheless, having access to these data not only allows us to have more reliable and accurate learning monitoring outside of labs, but also helps with the pressure to rely on parents for at-home exposure data.

Lastly, it was unfortunate that we did not manage to collect data for the general English vocabulary test at Time 2, and we eventually only used the pre-test for controlling children's initial English vocabulary. Ideally, it would be interesting to see how learning target vocabulary and grammar from the animation would impact children's general English vocabulary. Nonetheless, we hope this study will shed light on the complexity of second/foreign language learning through media in everyday settings, the challenge of monitoring young learners' target exposure during the targeted intervention and the sensitivity of children's learning outcomes to *how* and *when* we measure the targeted/general English knowledge. We would further encourage researchers to manipulate learning conditions as well as potential influential factors more rigorously during intervention experiments in the future, while also providing better control of children's language exposure outside of the laboratory.

Appendix: Tables and Figures

Appendix Table 1 Summary table of target vocabulary and grammar

Target	Type	Type frequency	Token frequency	Episode
Possessive *'s*		6	14	P1, P2
	Brother**'s**		6	P1
	David**'s**		1	P1
	Raj**'s**		4	P2
	Carmon**'s**		1	P2
	Ken**'s**		1	P2
	Lina**'s**		1	P2
Possessive pronouns		4	61	P1, P2, V1, V2, Home
	My		36	P1, P2, V2, Home
	Our		11	V1, Home
	Your		6	P2, Home
	Her		4	P2
	His		4	P2
Past tense (irregular)		4	11	V1, V2, Home
	Drew		6	V1
	Threw		1	V2
	Hit		2	V2
	Dug		2	Home
Past tense (regular)		2	4	V2, Home
	kicked		3	V2
	planted		1	Home
Verb stems		7	34	V1, V2, Home
	draw		11	V1
	Put		4	Home
	Kick		4	V2
	throw		4	V2
	Hit		3	V2
	plant		3	Home
	Dig		5	Home
Verb *-ing* form		6	9	V1, V2, Home
	drawing		1	V1
	kicking		1	V2
	throwing		2	V2
	digging		3	Home
	putting		1	Home
	planting		1	Home

Appendix Table 2 Table of time stamps for each target

Target	Grammar category	Type	Episode
My **brother's** children (1:03)	Possessive 's	Brother's	P1
I am my **brother's**…? (2:42)	Possessive 's	Brother's	P1
Jenny is my **brother's**…? (3:10)	Possessive 's	Brother's	P1
My **brother's** daughter (3:17)	Possessive 's	Brother's	P1
David's son (3:35)	Possessive 's	David's	P1
My **brother's** (3:41)	Possessive 's	Brother's	P1
My **brother's** son (3:45)	Possessive 's	Brother's	P1
Raj's picture (1:54)	Possessive 's	Raj's	P2
Raj's dog (2:08)	Possessive 's	Raj's	P2
Raj's dog (2:19)	Possessive 's	Raj's	P2
Raj's dog (2:48)	Possessive 's	Raj's	P2
Carmon's flowers (2:50)	Possessive 's	Carmon's	P2
Ken's robot (3:09)	Possessive 's	Ken's	P2
Lina's robot (3:16)	Possessive 's	Lina's	P2
Bird's nest (0:55)	Possessive 's	Bird's	54
Bird's home (0:58)	Possessive 's	Bird's	54
Bird's nest (2:20)	Possessive 's	Bird's	54
My family (0:54)	Possessive pronouns	My	P1
My father (0:56)	Possessive pronouns	My	P1
My mother (0:58)	Possessive pronouns	My	P1
My brother (1:01)	Possessive pronouns	My	P1
My brother's children (1:03)	Possessive pronouns	My	P1
My father (1:11)	Possessive pronouns	My	P1
My mother (1:15)	Possessive pronouns	My	P1
My father (1:29)	Possessive pronouns	My	P1
My father (1:33)	Possessive pronouns	My	P1
Mr. Jones is **my**…? -Father (1:39)	Possessive pronouns	My	P1
My father (1:42)	Possessive pronouns	My	P1
My mother (1:49)	Possessive pronouns	My	P1
My mother (1:51)	Possessive pronouns	My	P1
This is **my**…? – mother (2:01)	Possessive pronouns	My	P1
My brother (2:12)	Possessive pronouns	My	P1
My name is David (2:19)	Possessive pronouns	My	P1
He is **my**…? – brother (2:25)	Possessive pronouns	My	P1
My brother (2:31)	Possessive pronouns	My	P1
My sister (2:34)	Possessive pronouns	My	P1
My sister (2:36)	Possessive pronouns	My	P1
My brother's (2:42)	Possessive pronouns	My	P1
My children (2:49)	Possessive pronouns	My	P1
My daughter (3:03)	Possessive pronouns	My	P1
My brother's (3:10)	Possessive pronouns	My	P1
My brother's (3:16)	Possessive pronouns	My	P1
My son (3:25)	Possessive pronouns	My	P1
My brother's (3:41)	Possessive pronouns	My	P1

My brother's (3:45)	Possessive pronouns	My	P1
My class (3:50)	Possessive pronouns	My	P1
His dog (2:10)	Possessive pronouns	His	P2
My dog (2:15)	Possessive pronouns	My	P2
Your picture (2:35)	Possessive pronouns	Your	P2
Her flowers (2:45)	Possessive pronouns	Her	P2
His dog (2:53)	Possessive pronouns	His	P2
Her flowers (2:54)	Possessive pronouns	Her	P2
His robot (3:13)	Possessive pronouns	His	P2
Her robot (3:18)	Possessive pronouns	Her	P2
My girl (3:25)	Possessive pronouns	My	P2
Her princess (3:28)	Possessive pronouns	Her	P2
His dinosaur (3:29)	Possessive pronouns	His	P2
My butterfly (2:24)	Possessive pronouns	My	V1
Our butterfly pictures (2:36)	Possessive pronouns	Our	V1
Our class (3:26)	Possessive pronouns	Our	V1
My foot (1:25)	Possessive pronouns	My	V2
Our garden (0:24)	Possessive pronouns	Our	Home
Our garden (0:29)	Possessive pronouns	Our	Home
My garden (0:33)	Possessive pronouns	My	Home
Your garden (0:37)	Possessive pronouns	Your	Home
My garden (00:39)	Possessive pronouns	My	Home
Our garden (00:41)	Possessive pronouns	Our	Home
Our seeds (1:21)	Possessive pronouns	Our	Home
Our seeds (1:26)	Possessive pronouns	Our	Home
Your seeds (1:39)	Possessive pronouns	Your	Home
Your seeds (1:49)	Possessive pronouns	Your	Home
Our seeds (1:57)	Possessive pronouns	Our	Home
Our seeds (2:19)	Possessive pronouns	Our	Home
My seeds (2:25)	Possessive pronouns	My	Home
Your seeds (2:30)	Possessive pronouns	Your	Home
Your seeds (2:39)	Possessive pronouns	Your	Home
Our seeds (2:53)	Possessive pronouns	Our	Home
Our garden (3:26)	Possessive pronouns	Our	Home
I **kicked** the soccer ball (0:58)	Past tense (regular)	Kicked	V2
You **kicked** the ball (1:12)	Past tense (regular)	Kicked	V2
I **kicked** the ball with my foot (1:24)	Past tense (regular)	Kicked	V2
We **planted** the seeds (2:49)	Past tense (regular)	Planted	Home
I **drew** a picture (0:41)	Past tense (irregular)	Drew	V1
I **drew** a butterfly (2:21)	Past tense (irregular)	Drew	V1
Ken **drew** a caterpillar (2:53)	Past tense (irregular)	Drew	V1
You **drew** very beautiful pictures (3:04)	Past tense (irregular)	Drew	V1
I really like the pictures that you **drew** (3:10)	Past tense (irregular)	Drew	V1
We **drew** beautiful butterflies (3:23)	Past tense (irregular)	Drew	V1
You **threw** the ball (2:11)	Past tense (irregular)	Threw	V2
You **hit** the baseball very far (3:23)	Past tense (irregular)	hit	V2
I **hit** the baseball (3:28)	Past tense (irregular)	hit	V2

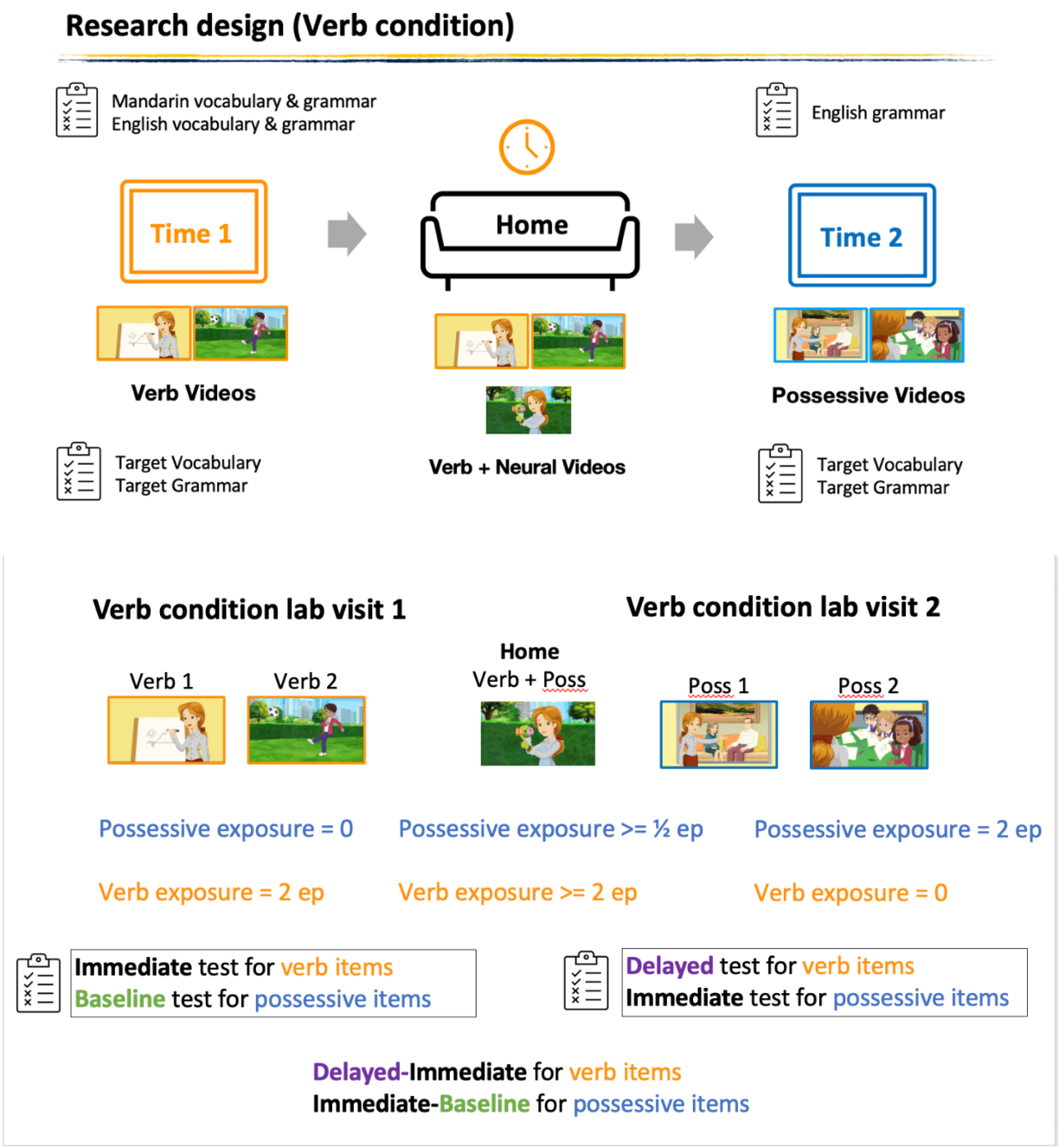
We can put the seeds in the holes we **dug** (2:23)	Past tense (irregular)	dug	Home
Can you put your seeds in the hole you **dug** (2:31)	Past tense (irregular)	dug	Home
Can you **draw** something for… (2:18)	Present tense	Draw	P2
You can **draw**, too (2:23)	Present tense	Draw	P2
Today we will **draw** (0:22)	Future tense	Draw	V1
Draw (0:24)	Present tense	Draw	V1
I'm going to **draw** a picture (0:27)	Future tense	Draw	V1
What would you like to **draw**? (0:48)	infinitive	Draw	V1
What would you like to **draw**? (1:04)	infinitive	Draw	V1
Today we will **draw** butterflies (1:13)	Future tense	Draw	V1
To **draw** pictures… (1:20)	infinitive	Draw	V1
What can you do with… **-draw** (2:08)	Present tense	Draw	V1
Yes, you can **draw** (2:12)	Present tense	Draw	V1
Let's **draw** a butterfly (2:14)	Present tense	Draw	V1
We've learned the words for **draw** (3:15)	Present tense	Draw	V1
Can you **kick** the soccer ball (1:04)	Present tense	kick	V2
I can **kick** the soccer ball (1:06)	Present tense	kick	V2
I can **throw** the ball (2:00)	Present tense	throw	V2
Can you **throw** the ball (2:15)	Present tense	throw	V2
I can **throw** the baseball (2:33)	Present tense	throw	V2
Hit the ball (2:52)	Present tense	Hit	V2
You can **hit** the ball (2:56)	Present tense	Hit	V2
I can **hit** the baseball (3:04)	Present tense	Hit	V2
I can **throw** the baseball (3:12)	Present tense	throw	V2
Kick the ball to me (3:32)	Present tense	kick	V2
I can **kick** the soccer ball (3:39)	Present tense	kick	V2
We will **plant** our seeds (1:21)	Future tense	plant	Home
Where will we **plant** our seeds (1:26)	Future tense	plant	Home
We need to **dig** a hole for your seeds (1:38)	infinitive	dig	Home
Can you **dig** a hole (1:47)	Present tense	dig	Home
I can **dig** a hole (1:52)	Present tense	dig	Home
Let's all **dig** holes (1:56)	Present tense	dig	Home
I like to dig (2:04)	infinitive	dig	Home
Now we can **plant** our seeds (2:19)	Present tense	plant	Home
We can **put** the seeds in the holes we dug (2:23)	Present tense	put	Home
Can you **put** your seeds (2:30)	Present tense	put	Home
Please **put** your seeds in the dirt (2:39)	Present tense	put	Home
I can **put** it in the dirt (3:13)	Present tense	put	Home
Now, I am **drawing** (0:31)	Progressive *-ing*	drawing	V1
I am **kicking** (0:56)	Progressive *-ing*	kicking	V2
I am **throwing** the ball (2:02)	Progressive *-ing*	throwing	V2
I am **throwing** the baseball (2:07)	Progressive *-ing*	throwing	V2
This is **digging** (1:40)	gerund	digging	Home
I am **digging** a hole (1:43)	Progressive *-ing*	digging	Home
That is enough **digging** (2:09)	gerund	digging	Home

I am **putting** my seeds (2:25)	Progressive *-ing*	putting	Home
We are **planting** the seeds (2:44)	Progressive *-ing*	planting	Home

Appendix Table 3 Table of curriculum target objectives by ABCmouse ELL

Episode #	Target vocabulary	Word type	Total frequency
P1 (ep 12)	brother	target-people	12
P1 (ep 12)	daughter	target-people	6
P1 (ep 12)	father	target-people	8
P1 (ep 12)	mother	target-people	8
P1 (ep 12)	sister	target-people	4
P1 (ep 12)	son	target-people	7
P1 (ep 12)	very good	target-other	4
P1 (ep 12)	also	bonus words	1
P1 (ep 12)	family	bonus words	1
P1 (ep 12)	fix	bonus words	1
P1 (ep 12)	he	bonus words	2
P1 (ep 12)	help	bonus words	1
P1 (ep 12)	honey	bonus words	1
P1 (ep 12)	introduce	bonus words	1
P1 (ep 12)	meet	bonus words	2
P1 (ep 12)	screen	bonus words	1
P1 (ep 12)	who	bonus words	2
P2 (ep 15)	he	target-pronoun	5
P2 (ep 15)	I	target-pronoun	6
P2 (ep 15)	She	target-pronoun	9
P2 (ep 15)	her	target-possesive adjective	4
P2 (ep 15)	his	target-possesive adjective	4
P2 (ep 15)	my	target-possesive adjective	2
P2 (ep 15)	dog	target-noun	8
P2 (ep 15)	robot	target-noun	4
P2 (ep 15)	dinosaur	bonus words	1
P2 (ep 15)	draw	bonus words	2
P2 (ep 15)	give	bonus words	1
P2 (ep 15)	nice	bonus words	1
P2 (ep 15)	princess	bonus words	1
P2 (ep 15)	too	bonus words	1
V1 (ep 17)	draw	target-verb	16
V1 (ep 17)	butterfly	target-noun	7
V1 (ep 17)	crayon	target-noun	4
V1 (ep 17)	paper	target-noun	7
V1 (ep 17)	pencil	target-noun	10
V1 (ep 17)	picture	target-noun	7
V1 (ep 17)	caterpillar	bonus words	2
V1 (ep 17)	great work	bonus words	1
V1 (ep 17)	our	bonus words	2
V1 (ep 17)	really	bonus words	1
V1 (ep 17)	take	bonus words	1

V2 (ep 41)	catch	target-verb	4
V2 (ep 41)	hit	target-verb	5
V2 (ep 41)	I can [action] the [ball].	target-verb	9
V2 (ep 41)	kick	target-verb	8
V2 (ep 41)	throw	target-verb	7
V2 (ep 41)	ball	target-noun	20
V2 (ep 41)	baseball	target-noun	11
V2 (ep 41)	soccer	target-noun	6
V2 (ep 41)	far	bonus words	1
V2 (ep 41)	outside	bonus words	1



Appendix Figure 1 Overall study design using verb condition as an example.

References

Al Mamun, M. A., Lawrie, G., & Wright, T. (2020). Instructional design of scaffolded online learning modules for self-directed and inquiry-based learning environments. *Computers & Education, 144* (103695), 1-17. https://doi.org/10.1016/j.compedu.2019.103695

Alexiou, T. & N. Konstantakis. 2009. Lexis for young learners: Are we heading for frequency or just common sense?" In Selected Papers from the 18th International Symposium on Theoretical and Applied Linguistics (ISTAL), Thessaloniki: Aristotle University of Thessaloniki, School of English, 59-66. https://doi.org/10.26262/istal.v18i0.5422

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*(1), 1–18. https://doi.org/10.1037/a0021017.supp

Andersen, R. W. (1978). An implicational model for second language research. *Language Learning, 28*(2), 221–282. https://doi.org/10.1111/j.1467-1770.1978.tb00134.x

Anderson, D. R., Bryant, J., Wilder, A., Santomero, A., Williams, M., & Crawley, A. M. (2000). Researching Blue's Clues: Viewing Behavior and Impact. *Media Psychology*, *2*(2), 179–194. https://doi.org/10.1207/S1532785XMEP0202_4

Anderson, D. R., & Pempek, T. A. (2005). Television and Very Young Children. *American Behavioral Scientist*, *48*(5), 505–522. https://doi.org/10.1177/0002764204271506

Baldwin, D. A. (2000). Interpersonal understanding fuels knowledge acquisition. *Current Directions in Psychological Science, 9*, 40–45. https://doi.org/10.1111/1467-8721.00057

Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child Development, 67*, 3135–3153. https://doi.org/10.1111/j.1467-8624.1996.tb01906.x

Baron, A., Bedore, L. M., Peña, E. D., Lovgren-Uribe, S. D., López, A. A., & Villagran, E. (2018). Production of Spanish grammatical forms in U. S. bilingual children. *American Journal of Speech-Language Pathology, 27*(3), 975–987. https://doi.org/10.1044/2018_AJSLP-17-0074

Benassi, V., Overson, C. E., & Hakala, C. (Eds.). (2014). *Applying science of learning in education: Infusing psychological science into the curriculum*. Retrieved from http://teachpsych.org/ebooks/asle2014/index.php

Bergelson, E. (2020). The Comprehension Boost in Early Word Learning: Older Infants Are Better Learners. *Child Development Perspectives*, *14*(3), 142–149. https://doi.org/10.1111/cdep.12373

Bergelson, E., & Swingley, D. (2015). Early word comprehension in infants: Replication and extension. *Language Learning and Development, 11*,369–380. https://doi.org/10.1080/15475441.2014. 979387

Bloom, L. (1974). Talking, understanding, and thinking: Developmental relationship between receptive and expressive language. In R. Schiefelbusch & L. Lloy (Eds.), *Language perspectives: Acquisition, retardation, and intervention* (pp. 285-311). Baltimore MD: University Park Press. https://doi.org/10.7916/D88S4VMC

Bohman, T., Bedore, L., Peña, Mendez-Perez, A., & Gillam, R. (2010). What you hear and what you say: Language performance in Spanish-English bilinguals. *International Journal of*

Bilingual Education and Bilingualism, *13*, 325–344. https://doi.org/10.1080/13670050903342019

Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S. Y., Pascual, L., … & Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child development, 75*(4), 1115-1139. https://doi.org/10.1111/j.1467-8624.2004.00729.x

Bosseler, A., & Massaro, D. W. (2003). Development and Evaluation of a Computer-Animated Tutor for Vocabulary and Language Learning in Children with Autism. *Journal of Autism and Developmental Disorders*, *33*(6), 653–672. https://doi.org/10.1023/B:JADD.0000006002.82367.4f

Brandone, A. C., Pence, K. L., Golinkoff, R. M., & Hirsh-Pasek, K. (2007). Action Speaks Louder Than Words: Young Children Differentially Weight Perceptual, Social, and Linguistic Cues to Learn Verbs. *Child Development*, *78*(4), 1322–1342. https://doi.org/10.1111/j.1467-8624.2007.01068.x

Bransford, J. B., Brown, A. L., & Cocking, R. R. (Eds.). (1999). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

British Council. (2013). The English effect: The impact of English, what it's worth to the UK and why it matters to the world. https://www.britishcouncil.org/sites/default/files/english-effect-report-v2.pdf

Brown, R. (1973). *A first language: The early stages*. London: George Allen & Unwin.

Calvert, S. L., & Richards, M. (2014). Children's parasocial relationships with media characters. In J. Bossert, A. Jordan, & D. Romer (Eds.), *Media and the well being of children and adolescents*. Oxford: Oxford University Press.

Calvert, S. L., Strong, B. L., Jacobs, E. L., & Conger, E. E. (2007). Interaction and Participation for Young Hispanic and Caucasian Girls' and Boys' Learning of Media Content. *Media Psychology*, *9*(2), 431–445. https://doi.org/10.1080/15213260701291379

Carnine, D., & Kinder, D. (1985). Teaching low-performing students to apply generative and schema strategies to narrative and expository material. *Remedial and Special Education, 6*(1), 20-30. https://doi.org/10.1177/074193258500600l0

Casasola, M. (2005). When less is more: how infants learn to form an abstract categorical representation of support. *Child Development, 76* (1), 279 –290. https://doi.org/10.1111/j.1467-8624.2005.00844.x

Celce-Murcia, M. (1991). Grammar pedagogy in second and foreign language teaching. *TESOL quarterly, 25*(3), 459-480. https://doi.org/10.2307/3586980

Chao, Y. (1968). *A grammar of spoken Chinese*. Berkeley: University of California Press.

Chen, J. (2013). *Breaking through the traditional second language learning model—exploring different exposure approaches for learners of different ages*. [Doctoral dissertation, University of Michigan]. Deep Blue database. https://deepblue.lib.umich.edu/handle/2027.42/97887?show=full

Chen, J., Tardif, T., Ip, K., & Hu, X.-S. (April, 2017). Does translation facilitate or interfere with learning words in a second language? Poster presented at Society for Research in Child Development. Austin, TX.

Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Chiong, C., & Shuler C. (2010). *Learning: Is there an app for that? Investigations of young children's usage and learning with mobile devices and apps*. New York: The Joan Ganz Cooney Center at Sesame Workshop. https://joanganzcooneycenter.org/wp-content/uploads/2010/10/learningapps_final_110410.pdf

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language, 22*(3),497–529. https://doi.org/10.1017/S0305000900009934

Comstock, G., & Scharrer, E. (2010). *Media and the American child*. Burlington, MA: Academic Press.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268–294. https://doi.org/10.1037/0096-3445.104.3.268

Crawley, A. M., Anderson, D. R., Santomero, A., Wilder, A., Williams, M., Evans, M. K., & Bryant, J. (2002). Do children learn how to watch television? The impact of extensive experience with Blue's Clues on preschool children's television viewing behavior. *Journal of Communication, 52*(2), 264-280. https://doi.org/10.1111/j.1460-2466.2002.tb02544.x

Crawley, A. M., Anderson, D. R., Wilder, A., Williams, M., & Santomero, A. (1999). Effects of repeated exposures to a single episode of the television program Blue's Clues on the viewing behaviors and comprehension of preschool children. *Journal of Educational Psychology, 91*(4), 630–637. https://doi.org/10.1037/0022-0663.91.4.630

DeKeyser, R. M. (2018). Age in learning and teaching grammar. *The TESOL encyclopedia of English language teaching*, 1-6. https://doi.org/10.1002/9781118784235.eelt0106

Diamond, A. (2012). Activities and programs that improve children's executive functions. *Current Directions in Psychological Science, 21*(5), 335-341. https://doi.org/10.1177/0963721412453722

Dominey, P. F., & Dodane, C. (2004). Indeterminacy in language acquisition: The role of child directed speech and joint attention. *Journal of Neurolinguistics*, *17*(2–3), 121–145. https://doi.org/10.1016/S0911-6044(03)00056-3

Dulay, H. C., & Burt, M. K. (1974). Natural sequences in child second language acquisition. Language Learning, 24, 37–53. https://doi.org/10.1111/j.1467-1770.1974.tb00234.x

Erbaugh, M. S. (1992). The acquisition of Mandarin. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 3, pp.373-455). Hillsdale, NJ: Erlbaum.

Espinosa, L. M., Laffey, J. M., Whittaker, T., & Sheng, Y. (2006). Technology in the Home and the Achievement of Young Children: Findings From the Early Childhood Longitudinal Study. *Early Education & Development*, *17*(3), 421–441. https://doi.org/10.1207/s15566935eed1703_5

Fan, M. (2000). How big is the gap and how to narrow it? an investigation into the active and passive vocabulary knowledge of l2 learners. *RELC Journal*, *31*(2), 105–119. https://doi.org/10.1177/003368820003100205

Falk, D. (2004). Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences*, *27*(4), 491–503. https://doi.org/10.1017/S0140525X04000111

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur–Bates Communicative Development Inventories: User's guide and technical manual* (2nd ed.). Baltimore, MD: Brookes Publishing.

Ferguson, C. J., & Beresin, E. (2017). Social science's curious war with pop culture and how it was lost: The media violence debate and the risks it holds for social science. *Preventive medicine*, 99, 69–76. https://doi.org/10.1016/j.ypmed.2017.02.009

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of child language, 16*(3), 477-501. https://doi.org/10.1017/S0305000900010679

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. *Developmental Psycholinguistics: On-Line Methods in Children's Language Processing, 2008*,97–135. https://doi.org/10.1075/lald.44.06fer

Fiorella, L., & Zhang, Q. (2018). Drawing boundary conditions for learning by drawing. *Educational Psychology Review, 30*(3), 1115–1137. https://doi.org/10.1007/s10648-018-9444-8.

Fisch, S. (2004a). *Children's learning from educational television: Sesame Street and beyond*. Lawrence Erlbaum Associates Publishers. http://choicereviews.org/review/10.5860/CHOICE.42-2935

Fisch, S. (2004b). Characteristics of effective materials for informal education: a cross-media comparison of television, magazines, and interactive media. In M. Rabinowitz, F. C. Blumberg, & H. T. Everson (Eds.) *The Design of Instruction and Evaluation: Affordances*

of Using Media and Technology (1st ed. pp. 17-32). Routledge. https://doi.org/10.4324/9781410611376

Fisch, S. M., & Truglio, R. T. (Eds.). (2001). "G" is for growing: Thirty years of research on children and Sesame Street. Mahwah, NJ: Erlbaum

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Fulkerson, A. L., & Waxman, S. R. (2007). Words (but not Tones) facilitate object categorization: Evidence from 6- and 12-month-olds. *Cognition*, *105*(1), 218–228. https://doi.org/10.1016/j.cognition.2006.09.005

Gentner, D. (2006). Why verbs are hard to learn. In K. Hirsh-Pasek & R. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 544–564). New York: Oxford University Press.

Godfroid, A., & Kim, K. M. (2021). The contributions of implicit-statistical learning aptitude to implicit second-language knowledge. *Studies in Second Language Acquisition*, *43*(3), 606–634. https://doi.org/10.1017/S0272263121000085

Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development, 71*(4), 878–894. https://doi.org/10.1111/1467-8624.00197

Gola, A. A. H., Mayeux, L., & Naigles, L. R. (2012). Electronic media as incidental language teachers. In D. G. Singer & J. L. Singer (Eds.), *Handbook of children and the media* (pp. 139–156). Sage Publications, Inc.

Gola, A. A., Richards, M. N., Lauricella, A. R., & Calvert, S. L. (2013). Building meaningful relationships between toddlers and media characters to teach early mathematical skills. *Media Psychology, 16*, 390–411. https://doi.org/10.1080/15213269.2013.783774.

Goldin-Meadow, S. (2007). Pointing sets the stage for learning language—and creating language. *Child Development*, *78*(3), 741–745. https://doi.org/10.1111/j.1467-8624.2007.01029.x

Goldstone, R. L., & Day, S. B. (2012). Introduction to "new conceptualizations of transfer of learning." *Educational Psychologist, 47*(3), 149-152. https://doi.org/10.1080/00461520.2012.695710

Golinkoff, R. M., & Hirsh-Pasek, K. (2008). How toddlers begin to learn verbs. *Trends in Cognitive Sciences*, *12*(10), 397–403. https://doi.org/10.1016/j.tics.2008.07.003

van den Heuvel, M., Ma, J., Borkhoff, C. M., Koroshegyi, C., Dai, D. W. H., Parkin, P. C., Maguire, J. L., Birken, C. S., & on behalf of the TARGet Kids! Collaboration. (2019). Mobile Media Device Use is Associated with Expressive Language Delay in 18-Month-Old Children. *Journal of Developmental & Behavioral Pediatrics*, *40*(2), 99–104. https://doi.org/10.1097/DBP.0000000000000630

Hagan, J., Shaw, J. S., & Duncan, P. M. (Eds.). (2008). *Bright futures: Guidelines for health supervision of infants, children, and adolescents* (3rd ed.). Elk Grove Village, IL: American Academy of Pediatrics.

Hassinger-Das, B., Brennan, S., Dore, R. A., Golinkoff, R. M., & Hirsh-Pasek, K. (2020). Children and Screens. *Annual Review of Developmental Psychology*, *2*(1), 69–92. https://doi.org/10.1146/annurev-devpsych-060320-095612

Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly, 15*(1), 75-90. https://doi.org/10.1016/S0885-2006(99)00038-1

Hawkins, R., Al-Eid, S., Almahboob, I., Athanasopoulos, P., Chaengchenkit, R., Hu, J., Rezai, M., Jaensch, Carol., Jeon, Y., Jiang, A., Leung, Y.-K.I., Matsunaga, K., Ortega, M., Sarko, G., Snape, N. and VelascoZárate, K. (2006). Accounting for English article interpretation by L2 speakers. In Foster-Cohen, S.H., Medved Krajnovic, M. and Mihaljevic Djigunovic, J.(Eds.), *EUROSLA Yearbook*, *6*(1), pp. 7–25. https://doi.org/10.1075/eurosla.6.04haw

Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting Education in "Educational" Apps: Lessons From the Science of Learning. *Psychological Science in the Public Interest*, *16*(1), 3–34. https://doi.org/10.1177/1529100615569711

Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., Hennon, E., Rocroi, C., & Bloom, L. (2000). Breaking the Language Barrier: An Emergentist Coalition Model for the Origins of Word Learning. *Monographs of the Society for Research in Child Development, 65*(3), i–135. http://www.jstor.org/stable/3181533

Horst, J. S., Parsons, K. L., & Bryan, N. M. (2011). Get the Story Straight: Contextual Repetition Promotes Word Learning from Storybooks. *Frontiers in Psychology, 2*. https://doi.org/10.3389/fpsyg.2011.00017

Huang, R. L., Fletcher, P., Zhang, Z., Liang, W., Marchman, V., & Tardif, T. (2022). Early grammatical marking development in Mandarin-speaking toddlers. *Developmental psychology, 58*(4), 631–645. https://doi.org/10.1037/dev0001316

Hung, H. T. (2015). Flipping the classroom for English language learners to foster active learning. Computer Assisted Language Learning, 28(1), 81-96. https://doi.org/10.1080/09588221.2014.967701

Huston, A. C., Watkins, B. A., & Kunkel, D. (1989). Public policy and children's television. *American Psychologist, 44*(2), 424–433. https://doi.org/10.1037/0003-066X.44.2.424

Huston, A. C., & Wright, J. C. (1983). Children's processing of television: The informative functions of formal features. In J. Bryant & D. R. Anderson (Eds.), *Children's understanding of television: Research on attention and comprehension* (pp. 35-68). New York: Academic Press

Hwang, B. H., & Lee, D. (2022). Association between Motor and Language Skills Development in Children with Autism Spectrum Disorder: A Scoping Review. *International Journal of Disability*, Development and Education, 1-15. https://doi.org/10.1080/1034912X.2022.2092081

Hurwitz, L. B. (2019). Getting a Read on Ready To Learn Media: A Meta-analytic Review of Effects on Literacy. *Child Development*, *90*(5), 1754–1771. https://doi.org/10.1111/cdev.13043

Hyde, J., Kiesler, S., Hodgins, J. K., & Carter, E. J. (2014). Conversing with children: Cartoon and video people elicit similar conversational behaviors. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1787–1796. https://doi.org/10.1145/2556288.2557280

Imada, T., Zhang, Y., Cheour, M., Taulu, S., Ahonen, A., & Kuhl, P. K. (2006). Infant speech perception activates Broca's area: A developmental magnetoencephalography study. *NeuroReport, 17*(10), 957–962. https://doi.org/10.1097/01.wnr.0000223387.51704.89

Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-Pasek, K., Golinkoff, R.M. and Shigematsu, J. (2008), Novel Noun and Verb Learning in Chinese-, English-, and Japanese-Speaking Children. *Child Development*, 79: 979-1000. https://doi.org/10.1111/j.1467-8624.2008.01171.x

Izumi, E., & Isahara, H. (2004). Investigation into language learners' acquisition order based on an error analysis of a learner corpus. In *IWLeL 2004: an Interactive Workshop on Language e-learning* (pp. 63-71). http://dspace.wul.waseda.ac.jp/dspace/bitstream/2065/1396/1/07.pdf

Jia, G., & Fuse, A. (2007). Acquisition of English grammatical morphology by native Mandarin speaking children and adolescents: Age-related differences. *Journal of Speech, Language, and Hearing Research*, 50, 1280–1299.

Jing, M., Ye, T., Kirkorian, H. L., & Mares, L. (n.d.). *Screen media exposure and young children's vocabulary learning and development: A meta-analysis*.

Kannass, K. N., & Colombo, J. (2007). The Effects of Continuous and Intermittent Distractors on Cognitive Performance and Attention in Preschoolers. *Journal of Cognition and Development*, *8*(1), 63–77. https://doi.org/10.1080/15248370709336993

Kilicgun, M. Y. (2015). An examination on the quality of contents of the cartoons that children aged 3-6 years preferred to watch: The cartoon I like most. *Educational Research and Reviews*, *10*(10), 1415–1423. https://doi.org/10.5897/ERR2015.2208

Kim, M., McGregor, K. K., & Thompson, C. K. (2000). Early lexical development in English- and Korean-speaking children: Language-general and language-specific patterns. *Journal of Child Language, 27* (2), 225–254. https://doi.org/10.1017/S0305000900004104

Kirkorian, H. L., Choi, K., & Pempek, T. A. (2016). Toddlers' word learning from contingent and noncontingent video on touch screens. *Child Development*, *87*(2), 405–413. https://doi.org/10.1111/cdev.12508

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41, 75–86. https://doi.org/10.1207/s15326985ep4102_1

Kobayashi, M., Kakigi, R., Kanazawa, S., & Yamaguchi, M. K. (2020). Infants' recognition of their mothers' faces in facial drawings. *Developmental Psychobiology*, *62*(8), 1011–1020. https://doi.org/10.1002/dev.21972

Kokla, N. (2016). Dora the Explorer: A TV character or a preschoolers' foreign language teacher? *Selected Papers on Theoretical and Applied Linguistics*, *Vol 21*, 666-683 Pages. https://doi.org/10.26262/ISTAL.V21I0.5261

Krcmar, M. (2010). Can social meaningfulness and repeat exposure help infants and toddlers overcome the video deficit? *Media Psychology, 13*(1), 31–53. https://doi.org/10.1080/15213260903562917

Krcmar, M., Grela, B., & Lin, K. (2007). Can toddlers learn vocabulary from television? An experimental approach. *Media Psychol, 10*, 41-63.

Krok, W. C., & Leonard, L. B. (2015). Past tense production in children with and without specific language impairment across Germanic languages: A meta-analysis. Journal of Speech, Language, and Hearing Research, 58(4), 1326–1340. https://doi.org/10.1044/2015_JSLHR-L-14-0348

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100*(15), 9096–9101. https://doi.org/10.1073/pnas.1532872100

Li, X., & Brand, M. (2009). Effectiveness of music on vocabulary acquisition, language usage, and meaning for mainland Chinese ESL learners. *Contributions to music education*, 73-84. http://www.jstor.org/stable/24127119

Li, H., Majumdar, R., Chen, M.-R. A., & Ogata, H. (2021). Goal-oriented active learning (GOAL) system to promote reading engagement, self-directed learning behavior, and motivation in extensive reading. *Computers & Education*, *171*, 104239. https://doi.org/10.1016/j.compedu.2021.104239

Lin, L.-Y., Cherng, R.-J., Chen, Y.-J., Chen, Y.-J., & Yang, H.-M. (2015). Effects of television exposure on developmental skills among young children. *Infant Behavior and Development*, *38*, 20–26. https://doi.org/10.1016/j.infbeh.2014.12.005

Linebarger, D., Moses, A., Garrity Liebeskind, K., & McMenamin, K. (2013). Learning vocabulary from television: Does onscreen print have a role? *Journal of Educational Psychology, 105*, 609 – 621. http://dx.doi.org/10.1037/a0032582

Linebarger, D. L., & Piotrowski, J. T. (2009). TV as storyteller: How exposure to television narratives impacts at-risk preschoolers' story knowledge and narrative skills. *British Journal of Developmental Psychology*, *27*(1), 47–69. https://doi.org/10.1348/026151008X400445

Linebarger, D. L., & Walker, D. (2005). Infants' and Toddlers' Television Viewing and Language Outcomes. *American Behavioral Scientist*, *48*(5), 624–645. https://doi.org/10.1177/0002764204271505

Lisosky, J. M. (2001). For all kids' sakes: comparing children's television policy-making in Australia, Canada and the United States. *Media, culture & society, 23*(6), 821-842.

Luk, Z. P., & Shirai, Y. (2009). Is the Acquisition Order of Grammatical Morphemes Impervious to L1 Knowledge? Evidence From the Acquisition of Plural -s, Articles, and Possessive 's. *Language Learning*, *59*(4), 721–754. https://doi.org/10.1111/j.1467-9922.2009.00524.x

Mace-Matluck, B. J. (1979). The order of acquisition of English structures by Spanish speaking children: Some possible determinants. In R. W. Andersen (Eds.), *The acquisition and use of Spanish and English as first and second languages* (pp. 75–89). Washington, DC: TESOL.

Madigan, S., McArthur, B. A., Anhorn, C., Eirich, R., & Christakis, D. A. (2020). Associations Between Screen Use and Child Language Skills: A Systematic Review and Meta-analysis. *JAMA Pediatrics*, *174*(7), 665. https://doi.org/10.1001/jamapediatrics.2020.0327

Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., & Brandone, A. C. (2008). Focusing on the relation: Fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science*, *11*(4), 628–634. https://doi.org/10.1111/j.1467-7687.2008.00707.x

Mares, M. L., & Acosta, E. E. (2010). Teaching inclusiveness via TV narratives in the US: Young viewers need help with the message. *Journal of Children and Media, 4*(3), 231-247. https://doi.org/10.1080/17482798.2010.486127

Martins, N., & Booth, M. (2022). Media and Emotional Development. In D. Lemish, *The Routledge International Handbook of Children, Adolescents, and Media* (2nd ed., pp. 227–234). Routledge. https://doi.org/10.4324/9781003118824-29

Mayer, R. E. (2004). Should There Be a Three-Strikes Rule Against Pure Discovery Learning? *American Psychologist*, *59*(1), 14–19. https://doi.org/10.1037/0003-066X.59.1.14

Meltzoff, A. N., Kuhl, P. K., Movellan, J., & Sejnowski, T. J. (2009). Foundations for a New Science of Learning. *Science*, *325*(5938), 284–288. https://doi.org/10.1126/science.1175626

Meyer, M., Zosh, J. M., McLaren, C., Robb, M., McCaffery, H., Golinkoff, R. M., Hirsh-Pasek, K., & Radesky, J. (2021). How educational are "educational" apps for young children? App store content analysis using the Four Pillars of Learning framework. *Journal of Children and Media*, *15*(4), 526–548. https://doi.org/10.1080/17482798.2021.1882516

Miron, D., Bryant, J., & Zillmann, D. (2001). Creating vigilance for better learning from television. In D. Singer & J. L. Singer (Eds.), Handbook of children and the media (pp. 153–181). Thousand Oaks, CA: Sage.

Möttönen, R., Järveläinen, J., Sams, M., & Hari, R. (2005). Viewing speech modulates activity in the left SI mouth cortex. *NeuroImage*, *24*(3), 731–737. https://doi.org/10.1016/j.neuroimage.2004.10.011

Nicholas, H., Lightbown, P. M., & Spada, N. (2001). Recasts as Feedback to Language Learners. *Language Learning, 51*(4), 719–758. https://doi.org/10.1111/0023-8333.00172

Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic approaches to bilingualism, 1*(3), 213-237. https://doi.org/10.1075/lab.1.3.01par

Pratt, M. W., Green, D., MacVicar, J., & Bountrogianni, M. (1992). The mathematical parent: Parental scaffolding, parenting style, and learning outcomes in long-division mathematics homework. *Journal of Applied Developmental Psychology, 13*, 17–34. https://doi.org/10.1016/0193-3973(92)90003-Z

Puzio, D., Makowska, I., & Rymarczyk, K. (2022). Raising the Child—Do Screen Media Help or Hinder? The Quality over Quantity Hypothesis. *International Journal of Environmental Research and Public Health*, *19*(16), 9880. https://doi.org/10.3390/ijerph19169880

Quine, W. V. O. (1960). *Word and object.* MIT Press.

Rankin, Y. A., Morrison, D., McNeal, M., Gooch, B., & Shute, M. W. (2009, April). Time will tell: In-game social interactions that facilitate second language acquisition. In *Proceedings of the 4th international conference on foundations of digital games* (pp. 161-168). https://doi.org/10.1145/1536513.1536546

Rao, P. S. (2019). The role of English as a global language. *Research Journal of English*, 4(1), 65-79. https://www.rjoe.org.in/Files/vol4issue1/new/OK%20RJOE-Srinu%20sir(65-79).pdf

Rice, M. L., Huston, A. C., Truglio, R., & Wright, J. C. (1990). Words from "Sesame Street": Learning vocabulary while viewing. *Developmental Psychology*, 26, 421–428. https://doi.org/10.1037//0012-1649.26.3.421

Richards, M. N. & Calvert, S. L. (2017). Media characters, parasocial relationships, and the social aspects of children's learning across media platforms. In R. Barr & D. N.

Linebarger, *Media Exposure During Infancy and Early Childhood* (pp. 141-163). Springer International Publishing. https://doi.org/10.1007/978-3-319-45102-2

Richert, R. A., Robb, M. B., Fender, J. G., & Wartella, E. (2010). Word learning from baby videos. *Archives of Pediatrics & Adolescent Medicine*, *164*(5), 432-437. https://doi.org/10.1001/archpediatrics.2010.24

Richert, R. A., Robb, M. B., & Smith, E. I. (2011). Media as Social Partners: The Social Nature of Young Children's Learning From Screen Media: Media as Social Partners. *Child Development*, *82*(1), 82–95. https://doi.org/10.1111/j.1467-8624.2010.01542.x

Rideout, V., & Robb, M. B. (2020). *The Common Sense census: Media use by kids age zero to eight, 2020*. San Francisco, CA: Common Sense Media.

Rivera Pérez, J. F., Hart, R., & Lund, E. (2021). Vocabulary-learning cues used on children's bilingual programming. *Journal of Children and Media, 15*(3), 301–319. https://doi.org/10.1080/17482798.2020.1802315

Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2014). Skype me! Socially contingent interactions help toddlers learn language. *Child development, 85*(3), 956–970. https://doi.org/10.1111/cdev.12166

Roseberry, S., Hirsh-Pasek, K., Parish-Morris, J., & Golinkoff, R. M. (2009). Live action: can young children learn verbs from video? *Child Development*, *80*(5), 1360–1375. https://doi.org/10.1111/j.1467-8624.2009.01338.x

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical Learning by 8-Month-Old Infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Salmon, K., Yao, J., Berntsen, O., & Pipe, M. E. (2007). Does providing props during preparation help children to remember a novel event?. *Journal of Experimental Child Psychology, 97*(2), 99-116. https://doi.org/10.1016/j.jecp.2007.01.001

Savva, M., Higgins, S., & Beckmann, N. (2022). Meta-analysis examining the effects of electronic storybooks on language and literacy outcomes for children in grades Pre-K to grade 2. *Journal of Computer Assisted Learning*, *38*(2), 526–564. https://doi.org/10.1111/jcal.12623

Sawyer, R. K. (2014). Introduction: The New Science of Learning. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (2nd ed., pp. 1–18). Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.002

Scheele, A., Leseman, P., & Mayo, A. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics*, 31, 117-140.

Schmidt, R. (1992). Awareness and second language acquisition. *Annual review of applied linguistics*, 13, 206-226.

Selnow, G. W., & Bettinghaus, E. (1982). Television exposure and language development. *Journal of Broadcasting, 26*(1), 469-479. https://doi.org/10.1080/08838158209364014

Shin, S. J. (2008). Preparing non-native English-speaking ESL teachers. Teacher Development, 12(1), 57-65. https://doi.org/10.1080/13664530701827749

Silverman, R., Crandell, J. D., & Carlis, L. (2013). Read aloud and beyond: The effects of read aloud extension activities on vocabulary in Head Start classrooms. Early Education and Development, 24 (2), 98–122. https://doi.org/10.1080/10409289.2011.649679

Slobin, D. I. (1982). Universal and particular in the acquisition of language. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 128–170). Cambridge University Press.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

Straker, L., Zabatiero, J., Danby, S., Thorpe, K., & Edwards, S. (2018). Conflicting Guidelines on Young Children's Screen Time and Use of Digital Technology Create Policy and Practice Dilemmas. *The Journal of Pediatrics*, *202*, 300–303. https://doi.org/10.1016/j.jpeds.2018.07.019

Strouse, G. A., O'Doherty, K., & Troseth, G. L. (2013). Effective coviewing: Preschoolers' learning from video after a dialogic questioning intervention. *Developmental Psychology, 49*(12), 2368–2382. https://doi.org/10.1037/a0032463

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/S1366718915000243

Sung, Y., Chang, K., & Yang, J. (2015). How effective are mobile devices for language learning? A meta-analysis. *Educational Research Review*, 16, 68–84. https://doi.org/10.1016/j.edurev.2015.09.001.

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

Talbott, M. R., Young, G. S., Munson, J., Estes, A., Vismara, L. A., & Rogers, S. J. (2020). The Developmental Sequence and Relations Between Gesture and Spoken Language in Toddlers

With Autism Spectrum Disorder. *Child Development*, *91*(3), 743–753. https://doi.org/10.1111/cdev.13203

Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the "Noun Bias" in Context: A Comparison of English and Mandarin. *Child Development*, *70*(3), 620–635. https://doi.org/10.1111/1467-8624.00045

Tare, M., Chiong, C., Ganea, P., & DeLoache, J. (2010). Less is more: How manipulative features affect children's learning from picture books. *Journal of Applied Developmental Psychology*, *31*(5), 395–400. https://doi.org/10.1016/j.appdev.2010.06.005

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.

Twomey, K. E., Lush, L., Pearce, R., & Horst, J. S. (2014). Visual variability affects early verb learning. *British Journal of Developmental Psychology*, *32*(3), 359–366. https://doi.org/10.1111/bjdp.12042

Twomey, K. E., Ranson, S. L., & Horst, J. S. (2013). That's more like it: More exemplars facilitate word learning. *Infant and Child Development. 23*(2), 105–122. https://doi.org/10.1002/icd.1824

Valkenburg, P. M., & Vroone, M. (2004). Developmental changes in infants' and toddlers' attention to television entertainment. *Communication Research, 31*(3), 288–311. https://doi.org/10.1177/00936 50204263435

Van Meter, P. (2001). Drawing construction as a strategy for learning from text. *Journal of Educational Psychology, 93*(1), 129–140. https://doi.org/10.1037/0022-0663.93.1.129.

Viner R; Royal College of Pediatrics and Child Health. The health impacts of screen time: a guide for clinicians and parents. Published 2019. Accessed December 5, 2019.

https://www.rcpch.ac.uk/sites/default/files/2018-12/rcpch_screen_time_guide_-_final.pdf

Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard university press.

Waxman, S., Fu, X., Arunachalam, S., Leddon, E., Geraghty, K., & Song, H. (2013). Are Nouns Learned Before Verbs? Infants Provide Insight Into a Long-Standing Debate. *Child Development Perspectives*, *7*(3), 155–159. https://doi.org/10.1111/cdep.12032

Wasik, B. A., & Jacobi-Vessels, J. L. (2017). Word Play: Scaffolding Language Development Through Child-Directed Play. *Early Childhood Education Journal*, *45*(6), 769–776. https://doi.org/10.1007/s10643-016-0827-5

Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental psychology*, 36(5), 571. https://doi.org/10.1037/0012-1649.36.5.571

Xanthos, A., Laaha, S., Gillis, S., Stephany, U., Aksu-Koç, A., Christofidou, A., Gagarina, N., Hrzica, G., Ketrez, F. N., Kilani-Schoch, M., KoreckyKröll, K., Kovacõevic¨, M., Laalo, K., Palmovic¨, M., Pfeiler, B., Voeikova, M. D., & Dressler, W. U. (2011). On the role of morphological richness in the early development of noun and verb inflection. First Language, 31(4), 461–479. https://doi.org/10.1177/0142723711409976

Xu, Z., Chen, Z., Eutsler, L., Geng, Z., & Kogut, A. (2020). A scoping review of digital game-based technology on English language learning. *Educational Technology Research and Development*, *68*(3), 877–904. https://doi.org/10.1007/s11423-019-09702-2

Yamanaka, N., Kanazawa, S., & Yamaguchi, M. K. (2022). Infants' brain activity to cartoon face using functional near-infrared spectroscopy. *PLOS ONE*, *17*(2), e0262679. https://doi.org/10.1371/journal.pone.0262679

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science, 18*(5), 414–420. https://doi.org/10.1111/j.1467-9280.2007.01915.x

Yu, C., Smith, L. B., Klein, K. A., & Shiffrin, R. M. (2007). Hypothesis testing and associative learning in cross-situational word learning: Are they one and the same? In *Proceedings of the Annual Meeting of the Cognitive Science Society,* 29(29), 737-742. https://escholarship.org/content/qt40c145pc/qt40c145pc.pdf

Zdorenko, T., & Paradis, J. (2008). The acquisition of articles in child second language English: Fluctuation, transfer or both? *Second Language Research*, *24*(2), 227–250. https://doi.org/10.1177/0267658307086302

Zosh, J. M., Hirsh-Pasek, K., Hopkins, E. J., Jensen, H., Liu, C., Neale, D., Solis, S. L., & Whitebread, D. (2018). Accessing the Inaccessible: Redefining Play as a Spectrum. *Frontiers in Psychology*, *9*, 1124. https://doi.org/10.3389/fpsyg.2018.01124