Author/s:
Leung, S;Perry, C;Guy, J;Loats, D;Highfield, K;Kaufman, J

Title:
Short-Term Exposure to Second Language Apps Modulates Brain Responses in Preschoolers

Date:
2025-05-01

Citation:
Leung, S., Perry, C., Guy, J., Loats, D., Highfield, K. & Kaufman, J. (2025). Short-Term Exposure to Second Language Apps Modulates Brain Responses in Preschoolers. Infant and Child Development, 34 (3), https://doi.org/10.1002/icd.70015.

Persistent Link:
https://hdl.handle.net/11343/362676

Infant and Child Development

WILEY

REPORT OPEN ACCESS

Short-Term Exposure to Second Language Apps Modulates Brain Responses in Preschoolers

Sumie Leung[1,2] | Conrad Perry[3] | Jessica Guy[1] | Deborah Loats[1] | Kate Highfield[4] | Jordy Kaufman[1]

[1]Department of Psychological Sciences, Swinburne University of Technology, Melbourne, Australia | [2]Melbourne School of Psychological Sciences, University of Melbourne, Melbourne, Australia | [3]Faculty of Psychology, University of Adelaide, Adelaide, Australia | [4]Faculty of Education, University of Canberra, Acton, Australian Capital Territory, Australia

Correspondence: Jordy Kaufman (jkaufman@swin.edu.au)

Received: 24 April 2023 | **Revised:** 6 March 2025 | **Accepted:** 22 March 2025

Funding: This work was supported by the Australian Government Department of Education, Skills and Employment.

Keywords: auditory event-related potentials | EEG | language-immersive apps | preschoolers | second language learning

ABSTRACT

Previous research showed that short-term second language training modulates children's brain responses to language processing. However, little is known about whether short-term training from language-immersion apps would have the same effect on young children's neural processing of a newly learnt language. We examined the auditory event-related potentials generated by two groups of 3- to 5-year-old children (total $N = 32$; 14 male, 18 female; mean age = 49.6 months, SD = 6.0 months), in response to known and unknown non-native language words. The 'known' word stimuli were previously exposed to the children via either the ELLA language immersion applications ('ELLA' group) or flash cards ('FLASH' group). Electroencephalography data were analysed within early (200–300 ms) and late (400–600 and 600–800 ms) time-windows, to determine the main and interaction effects of group (ELLA vs. FLASH) and condition (KNOWN vs. UNKNOWN). We found that the early positive potential (of both groups) for the known words was significantly larger than that for the unknown words. Further, the early negative potential of the apps group was significantly larger than that of the flash card group. Our study showed that short-term training with language-immersion apps modulates language processing in preschool children's brains differently compared to digital flash cards.

1 | Introduction

Preschool children as young as 5 years old can learn from educational touchscreen apps, as shown by numerous studies and a recent meta-analysis (Xie et al. 2018). This benefit accrues in multiple aspects of learning, including science, telling time, word learning, and story comprehension (Aladé et al. 2016; Wang et al. 2016; Piotrowski and Krcmar 2017; Russo-Johnson et al. 2017). Among these aspects, touchscreen language learning has garnered considerable research attention in the area of emergent literacy in the preschool age range (e.g., Neumann 2014, 2016, 2018; O'Toole and Kannass 2018). For example, researchers have found that the multisensory features of some literacy applications (apps) contribute to the enhanced learning of letter names and sounds in 2- to 5-year-old children (Neumann 2018) and facilitate word learning in preschool children (Smeets and Bus 2015). This benefit extends to second language learning, with the interactional and haptic feedback features of touchscreens providing an immersion-style learning experience (Minogue and Jones 2006; Nassaji 2016;).

Measuring the extent to which children benefit from language learning programs can be done a number of ways (Barac et al. 2014; see review by Bialystok 2007). The majority of

Summary

- The aim of the study was to examine how well young children's brains process second language-related information acquired from short-term usage of language-immersion apps.
- Early positive ERP for the known words was larger than that for the unknown words, and early negative ERP for the apps group was larger than that of the flash card group.
- Short-term training from language apps modulates language processing in preschool children's brains differently compared to digital flash cards.

research has assessed children's learning outcomes using simple behavioural tests. Cognitive neuroscience technologies, including electroencephalography (EEG), may provide useful information that behavioural studies cannot measure, and the temporal order effects during speech processing can be directly examined (see Steinhauer and Connolly 2008). Event-related potential (ERP) is a method used to examine EEG activity that is time-locked to individual stimuli. It allows researchers to examine children's online speech processing via differences in the brain electrical activity that different types of stimuli elicit. Focusing on children's native languages, early components of ERPs around 100–200 ms after word onset have been linked to phonetic/phonological processing (e.g., Bonte and Blomert 2004). Following these are intermediate ERP components at around 200–300 ms, which have been linked to early lexical access (e.g., Kast et al. 2010;) or phoneme recognition (Ceponiene et al. 2005). Later on, lexical/semantic processing is reflected by ERP components at around 300–600 ms window (e.g., Benau et al. 2011). These are all important components of normal language processing.

Developmental-neural studies have also examined how young children's brains register different stages of second language acquisition including phonological and lexical processing (e.g., Bovo et al. 2018; Cheour et al. 2002; Jost et al. 2015; Marques et al. 2021; Peltola et al. 2005; Rinker et al. 2010, 2017; Tong et al. 2014; Shestakova et al. 2003). While some of these studies have focussed on children's electrophysiological responses from long-term second-language learning (e.g., Rinker et al. 2010, 2017), others have examined the effects of short-term second-language training (e.g., Shestakova et al. 2003; Moreno et al. 2015). Shestakova et al. (2003) observed larger early and late auditory ERP components (P3a in 200–500 ms and late discriminative negativity in 350–700 ms) in response to French vowels in 3- to 6-year-old Finnish-speaking children after they completed a 2-month French-learning program. These enhanced ERP components were also observed about 12–16 weeks after the learning program, compared to their peers who did not learn any foreign language. These larger ERP components may reflect the process of children successfully learning the phonemes of a second language. Similarly, 4- to 6-year-old English-speaking children showed an enhanced late ERP component (late discriminative negativity in 450–850 ms) after 4 weeks of French language training (Moreno et al. 2015). Further, this training effect persisted even after 1 year of completing the program, indicating that even short-term early childhood training could have lasting benefits on the developing brain. It is also worth noting that there was no fine distinction between 'early' or 'late' time-windows in previous studies focussing on young children's ERP components. Different stimuli and paradigm characteristics, as well as the age ranges of the participants, might have contributed to these variations.

The advantages of learning a second language in language-immersion programs in a traditional classroom setting, as compared to regular monolingual programs, are well established (e.g., Barbu et al. 2019; Bialystok et al. 2014; Hermanto et al. 2012). For example, Bialystok et al. (2014) found that children learning another language after only 2 years of an educational immersion program had a metalinguistic advantage over their peers in a non-immersion program. Barbu et al. (2019) showed that children with only 1 year of immersion education had a cognitive advantage compared to their peers, as indicated by faster responses in a selective auditory task. While the benefits of the language-immersion program are well documented, schools in rural areas lack the resources or infrastructure to implement second language classes, and are even more constrained in establishing language-immersion programs. One might argue that teachers could digitise flash cards (one of the traditional language teaching methods), and teach language remotely, but this method is less engaging for young children.

Well-developed language educational apps give children, especially in rural areas, an opportunity to be exposed to a second language, and they might provide a platform resembling a language-immersion program in the digital space. However, in the digital space, it is unclear how well children can learn from language-immersion apps, as compared to learning from traditional language learning methods, such as flash cards. In addition, little is known about how well young children's brains process second language-related information acquired from short-term usage of language-immersion apps. The current research is the electrophysiological component of the multi-faceted evaluation of the Early Learning Languages Australia (ELLA) program (see Section 2.1 for a description of the ELLA program). The ELLA apps provide an immersion environment for children to learn a second language in the digital space.

We conducted the evaluation study using various methods (e.g., behavioural, eye-tracking, parental reports and electrophysiological measures) to examine how much the digital immersion environment contributes to children's acquisition of new language knowledge. We used digital flash cards as control, as digital flash cards have been shown to be an effective method in teaching second language vocabulary to primary school students (Yowaboot and Sukying 2022). This control allowed us to compare ELLA, an activity-based method with an immersion environment, to flash cards, a more traditional and less interactive word-learning approach. Both methods were delivered via the same technology (touchscreen) to a group of preschool children who had no prior knowledge of the Indonesian language. A final report of this evaluation study was submitted to the Australian Government Department

of Education (Kaufman et al. 2017). In the current paper, we report the electrophysiological component of this evaluation study.

In the EEG study, we compared the auditory ERP waveforms generated in response to Indonesian words learnt via the ELLA apps and digital flash cards, as compared to those generated in response to Indonesian words that the preschool children had not been exposed to. Given that previous research has shown enhanced ERP components after short-term training (e.g., Shestakova et al. 2003; Moreno et al. 2015) and the advantages of language learning from language-immersion programs (e.g., Bialystok et al. 2014; Hermanto et al. 2012; Barbu et al. 2019), we hypothesised that children who learnt from the ELLA apps would exhibit larger ERP components in both early and late time-windows, as compared to the group who learnt from digital flash cards. We also hypothesised that children would exhibit larger ERP components in both early and late time-windows to the words that they were exposed to, as compared to the words that they were not exposed to.

2 | Materials and Methods

2.1 | Learning Materials

The ELLA program is a set of language-learning touchscreen apps designed to enhance interest in language learning among children in preschool services (ELLA website). This Australian Government initiative was established in alignment with their Early Years Learning Framework (EYLF). The apps feature characters called 'polyglots'. Each polyglot character is multilingual and is involved in several language-learning activities within the apps. Each app consists of several language-learning activities of the same theme, for example, App 1 (the Polyglots in the Playroom), App 2 (the Polyglots at the Beach), App 3 (the Polyglots in the Birthday Party) and App 4 (the Polyglots at the Zoo). These activities range from greetings, making a virtual cake to counting during an egg-race (see Figure 1 top row for screenshots of three ELLA app activities). While the program has been implemented in many languages, here only Indonesian (*Bahasa Indonesia*) was examined. Indonesian was chosen as the language, because historically Australia has been a world leader in teaching Indonesian as a foreign language and has a pool of expertise in teaching among non-native learners, as stated by the Australian Curriculum Assessment and Reporting Authority (ACARA 2014). In addition, at the time of this experiment, the Indonesian ELLA apps were one of the languages with the most complete set of ELLA apps that were developed.

The digital flash cards were made and presented on FC Maker on an iPad (see Figure 1 bottom row for screenshots of a digital flash card stimulus). The words used in the flash cards were structured in the same way as the Apps, such that there was not a single set of flash cards, but rather a number of different sets. Each Indonesian word sound file was embedded in each slide and was activated by a finger tap on the slide. The word lists used in the flash card apps were imageable words that could easily be associated with pictures, for example, 'ball' and 'ice-cream' rather than 'Hi, my name is …', as phrases such as greetings could not be easily displayed as a flash card. Thus, the list of words in the flash card apps was not as long as the list of words in the ELLA apps. However, only the words shared by both ELLA apps and flash card apps were used as stimuli in the experiment.

2.2 | Participants

Sixty-five preschool children who had no previous exposure to *Bahasa Indonesia* participated in the main experiment. Among those, 32 participants (14 male, 18 female; mean age = 49.6 months, SD = 6.0 months; range 37–59 months) completed the EEG experiment. Sixteen of them (7 female) were randomly allocated to the ELLA group and 16 (11 female) to the flash card group. There was no significant mean age

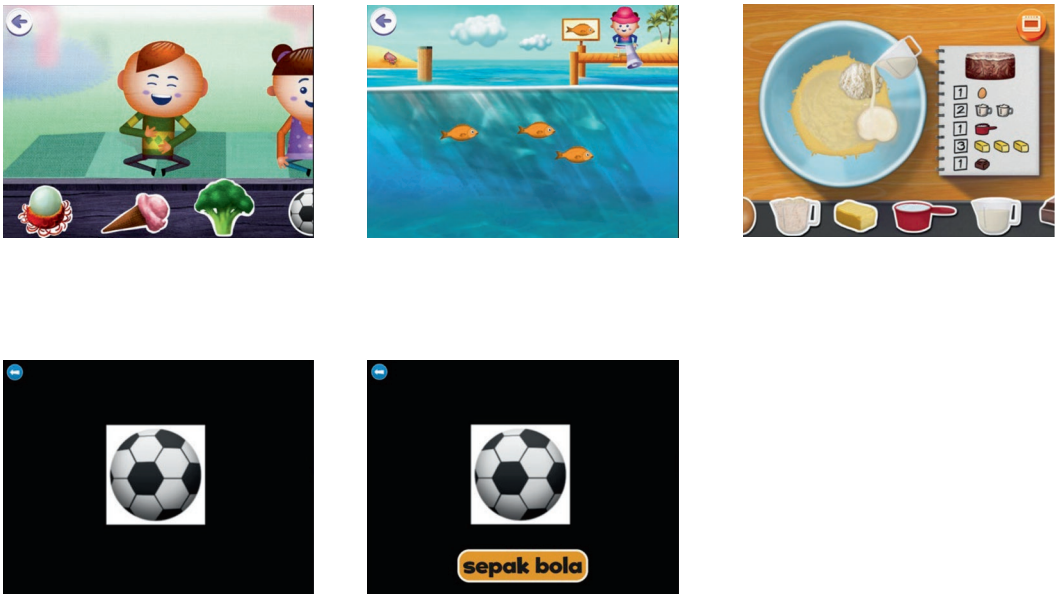

FIGURE 1 | Top row: Example screenshots of three ELLA app activities. Bottom row: Example screenshots of a digital flash card stimulus.

difference between the two groups (ELLA: 49.5 months, flash card: 49.7 months; $p = 0.937$). To control for differences in language aptitude of the two groups, participants' verbal ability was assessed by the Peabody Picture Vocabulary Test IV (PPVT-IV; Dunn and Dunn 2007). Of the 32 EEG participants, 1 participant did not complete the PPVT-IV test as he/she did not co-operate, although researchers were able to informally confirm that his/her verbal ability was very likely to be age normal through conversations; hence, this participant's data were included in the analysis. The group mean percentiles of PPVT scores for the remaining EEG participants were 72.33 (SD = 24.07) for the ELLA group and 74.00 (SD = 24.63) for the flash card group.

2.3 | Experiment Paradigm

During the first lab visit, participants learnt the same set of Indonesian words from either ELLA App 1 (Playroom) or a digital flash card app for 20 min (or less if participants were unwilling to continue). Following the 20-min ELLA or flash card activity, participants were asked to complete a few tasks, including a behavioural word task, an eye-tracking word task, and the PPVT-IV.

After the first lab visit, the ELLA group took an iPad home loaded with two ELLA apps: Apps 1 + 2 (Playroom + Beach) or Apps 1 + 3 (Playroom + Birthday Party) or Apps 1 + 4 (Playroom + Zoo). The flash card group took an iPad home loaded with two 'decks' of digital flash cards containing words and pictures from the two ELLA apps. Parents of both groups were given the same instructions: to ask their child to play on their given apps for 15–20 min per day for approximately 2 weeks. Parents were instructed that children could stop sooner if they expressed the desire not to continue for the whole 15 min. Parents were asked to record the duration of use and rate their child's engagement during each home-use session.

After 2 weeks of home-based learning, participants visited the lab and were asked to complete the behavioural word task and the eye-tracking word task. The tasks were the same as those in the first lab visit, except they contained materials from two apps. During this visit, researchers would communicate with the children and their guardians to assess their willingness to participate in the EEG study (e.g., their willingness of wearing the EEG net and/or their tiredness). For a subset of children who were willing to participate, their EEG data were also collected. The current paper focused on the EEG data and also reported the parent-reported home usage data. Results of the remaining tasks (including the behavioural word task and the eye-tracking word task) and other parent-reported data will be published elsewhere.

Upon completion of the experiment, participants and their guardians were reimbursed AUD 50–100 for their time and travelling costs. The study was approved by the Human Research Ethics Committee at Swinburne University of Technology, and written informed consent was obtained from the parents of all child participants.

2.4 | Stimulus Paradigm

An (audio) word-picture pair paradigm was used. The audio words are Indonesian words that participants had either been exposed to (KNOWN) or had not been exposed to (UNKNOWN). The pictures were either congruent or incongruent to the audio words. Each pair (trial) consisted of (1) an audio word file, (2) a blank screen for 1 s and (3) a picture for 1.5 s. The inter-trial interval between each word-picture pair was 2 s.

All audio word and picture stimuli were extracted from the ELLA apps. Twenty-three known audio words were selected from the two apps that the participants were exposed to, and 23 unknown audio words were selected from two other apps that the participants were not exposed to. The mean loudness, duration and fundamental frequency (F0) of the audio sound stimuli were calculated by the software Praat (Boersma 2001) and are presented in Table 1. The unknown words were slightly louder than the known words for Apps 1 + 2 (Playroom + Beach) and Apps 1 + 3 (Playroom + Birthday Party), with the mean loudness levels as 3.32 and 5.17 dB, respectively. The F0 of unknown words was higher than that of the known words for Apps 1 + 2 (Playroom + Beach) and Apps 1 + 4 (Playroom + Zoo), with the mean frequencies as 35.71 and 47.99 Hz, respectively. The duration was not different between the known words and the unknown words, as shown by the two-tailed *t*-tests in Table 1.

Participants were instructed to inform the researcher verbally by saying 'yes' or 'no' if the audio word matched the picture, after each trial. At the start of the experiment, participants were given three word-picture pairs to practice, which were discarded from the analyses. After this, a total of 95 word-picture pairs were presented to the participants. The total time of the presentation was approximately 9 min. Note that the task was designed

TABLE 1 | The mean loudness, duration and fundamental frequency (F0) of the known and unknown words, for each app group.

Apps 1 and 2 (Playroom + Beach)	Loudness (dB)	Duration (s)	F0 (Hz)
Known	69.12	0.80	315.21
Unknown	74.29	0.71	350.92
t-test	0.01*	0.12	0.00*
Apps 1 and 3 (Playroom + Birthday Party)	**Loudness (dB)**	**Duration (s)**	**F0 (Hz)**
Known	70.59	0.75	339.03
Unknown	73.91	0.76	355.12
t-test	0.02*	0.86	0.19
Apps 1 and 4 (Playroom + Zoo)	**Loudness (dB)**	**Duration (s)**	**F0 (Hz)**
Known	71.77	0.79	307.94
Unknown	74.16	0.74	355.93
t-test	0.15	0.40	0.00*

Note: The *t*-tests represent whether the loudness, duration or F0 of the known words were significantly different from that of the unknown words.
*$p < 0.05$.

to maintain participants' attention to the audio words, hence the ERP responses were not error-rate-corrected.

2.5 | EEG Data Acquisition

During the task, one researcher sat in a dimly lit electrically shielded room with the participant. Participants were seated approximately 60 cm in front of a video monitor on which visual stimuli were presented. The pictures were presented on a black background.

EEG was recorded with Netstation 4.4.2 acquisition software and an NA300 amplifier with a Hydrocel Geodesic Sensor Net comprised of 124 sensors (Electrical Geodesics Inc n.d.). Online EEG data were sampled at 500 Hz and were referenced to the vertex electrode.

2.6 | EEG Data Analysis

EEG data were examined in two phases. The first was designed to identify participants with particularly poor data for exclusion and to interpolate bad channels for the remaining participants. The second was designed to remove trials with artefacts after participants with potentially reasonable data were identified.

In the first data examination stage, 32 participants (ELLA = 15; FLASH = 17) completed the task without removing the cap in the task or similar major mishap, and thus had a full data set. Data from these participants were first filtered between 0.1 and 30 Hz. To examine the data, we first removed all electrodes that were external to the layer the mastoid was on (see Figure 2). These electrodes were not located on the scalp. Epochs for each participant were extracted from −200 to 800 ms at the stimulus onset of the audio words. Data were then re-referenced to the mastoids.

For each trial of each subject at each electrode, we computed a number of different measures: variance, kurtosis, maximum range, mean z values, and maximum z values, using code from the Fieldtrip visual toolbox. The scores for each of these measures for each participant were sorted from highest to lowest. For each electrode, for each measure, and for each participant, if more than 30% of the epochs on any given score were above these criteria (variance: 5000; kurtosis: 16, range: 400 μV, mean Z: 16, maximum Z: 16), a score of 1 was added to a total score. If a score was greater than 20, the subject was discarded. This stage led to five subjects (three from the ELLA group and two from the flash card group) being discarded entirely.

In the second data examination stage, for each subject, any channel which was bad on any measures was recalculated using the triangulation method in Fieldtrip with a Neighbourhood distance of 1. The average number of channels recalculated was 0.55 (SD = 1.09). Independent component analysis (ICA) was then used to eliminate ocular artefacts and other artefacts. This was done using only 20 components to simplify component

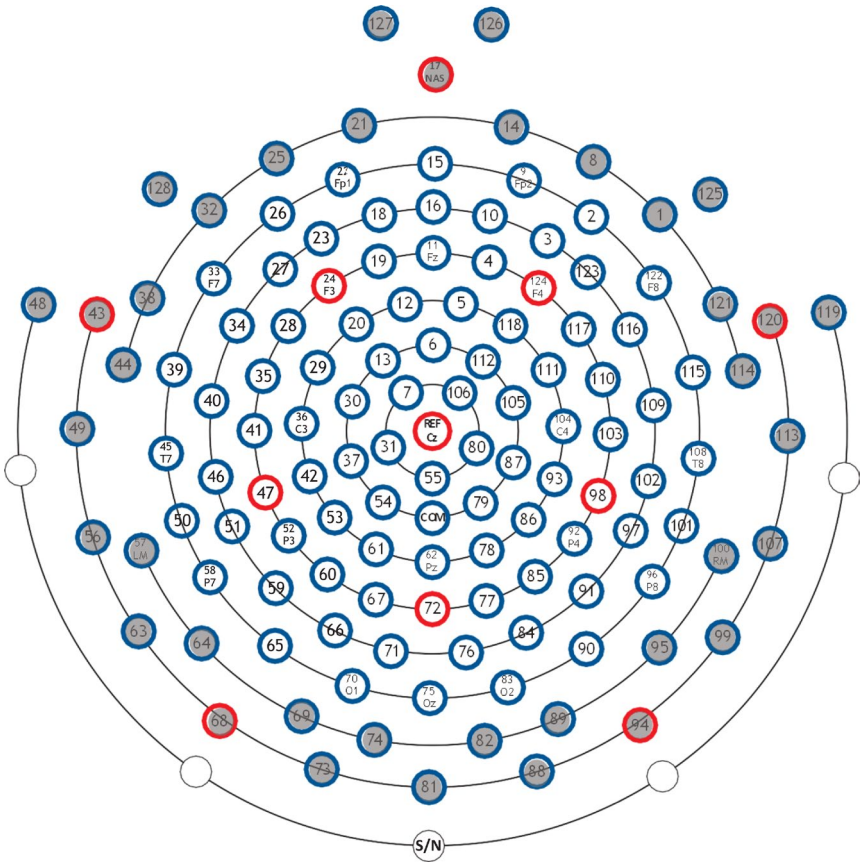

FIGURE 2 | An electrode map of the Hydrocel Geodesic Sensor Net. The greyed-out electrodes were not on the scalp and were not included in the analysis. Figure adapted from Geodesic Sensor Net Technical Manual (Electrical Geodesics Inc n.d.)

identification. Components were removed if they were associated with eye blinks, ocular movements, and likely muscle artefacts. For each subject, individual epochs were then rejected if they were over a new set of cut-offs (variance: 4000; kurtosis: 16; range: 350 μV; maximum *Z*: 16) or if they were 4.5 SDs over the mean of any of those measures. This procedure left one participant with only 21% of their original trials, and the data of this participant (from the ELLA group) were discarded.

After these two data examination stages, EEG data of 12 participants in the ELLA group and 14 in the FLASH group remained. On average, the number of epochs left in each condition of the two groups was: ELLA KNOWN: $M = 30.8$, SD = 8.663; ELLA UNKNOWN: $M = 31.42$, SD = 9.87; FLASH KNOWN: $M = 26.29$, SD = 7.54; FLASH UNKNOWN: $M = 26.92$, SD = 5.81.

2.7 | Statistical Analysis

For the parent-reported home usage data, we used a two-tailed *t*-test to examine the amount of time the ELLA group spent on the apps at home, in comparison to the amount of time the FLASH group spent on the digital flash cards at home.

For ERPs, we examined the results using cluster-based permutation tests (Maris and Oostenveld 2007), as implemented in Fieldtrip. This technique uses a data-driven method to identify significant electrodes (i.e., electrodes that were significantly different in the two groups/conditions) and thus does not require areas to be a priori defined and compared.

To identify the time-windows of interests, given the discrepancies between previous paradigms and ours, we initially conducted post hoc time-course analyses on the data using a 100–800 ms window. We performed cluster-based tests on the ERPs in 100 ms blocks from 100 to 800 ms. Note that when using cluster-based permutation methods, it is not possible to identify exact time-windows when significance occurs based on significant clusters of *p* values; thus, blocks are used to identify time regions of interest. The results of this are shown in Table 2.

Based on these preliminary results and considering that early effects are likely related to relatively short underlying processes, we chose a 200–300 ms window to examine the difference between unknown and known words. Selecting a later window is more challenging, as processing in this window typically represents slower effects such as semantic processing and more conscious strategies, at least in the adult literature. We tested two windows: 400–600 and 600-800 ms. These were designed to examine semantic and later, more complex processing as often reported in adult literature (e.g., Kutas and Federmeier 2011; Smith and Guster 1993) and in results from a similar age group (Schneider et al. 2023). However, these time-windows are more speculative because the typical features used to identify them were not obvious in our data set, likely due to our experimental paradigm and the age of the participants.

For statistical analysis on ERPs, we applied cluster-based permutation tests on each data point of these two time-windows (early: 200–300 ms; late: 400–600 ms; 600–800 ms) and used a two-tailed distribution. With this technique, we examined the main effects of Presentation group (ELLA vs. FLASH), Word Novelty condition (KNOWN vs. UNKNOWN) and the interaction between the Presentation group and Word Novelty condition in the 200–300, 400–600 and 600–800 ms time-windows.

3 | Results

Among the 32 participants who participated in the EEG study, we received 20 parental reports of home usage, with 9 from the ELLA group and 11 from the FLASH group. Results from parental reports showed that the ELLA group spent significantly more time on the apps (mean = 27.7 min, SD = 17.5 min) than their peers on the digital flash cards (mean = 10.7 min, SD = 5.8 min) at home ($p = 0.020$).

The EEG results showed that within the 200–300 ms window, there was a significant main effect of the Presentation group in certain frontal right electrodes ($p = 0.045$), and a significant main effect of the Word Novelty condition in other electrodes ($p = 0.028$), but no interaction. Specifically, as shown in the topographic difference map between the ELLA and the FLASH groups, the negative potential of the ELLA group is significantly larger than that of the FLASH group in the frontal and right lateralised electrodes (as indicated by the crosses on the map of Figure 3 and '*EvF' on Electrodes 16, 123, 109 and 97 in Figure 5). As shown in the topographic difference map between the KNOWN and the UNKNOWN conditions, the positive potential of the KNOWN condition is significantly larger than that of the UNKNOWN condition in the left central electrodes (as indicated by the crosses on the map of Figure 4 and '*KvU' on

TABLE 2 | *p* values of the most significant cluster in 100 ms time-windows.

	Time window (ms)						
Comparison	**100–200**	**200–300**	**300–400**	**400–500**	**500–600**	**600–700**	**700–800**
Known vs. Unknown (Main effect)	0.27	0.024	0.070	0.11	0.59		
Ella vs. Flash (Main effect)	0.14	0.015	0.17				
Ella vs. Flash (Interaction)	0.093		0.56	0.28	0.63	0.040	0.12
Known vs. Unknown (Ella)						0.010	
Known vs. Unknown (Flash)							

Note: Blank values are when there are no significant clusters or when tests are not conducted due to the interaction not being significant.

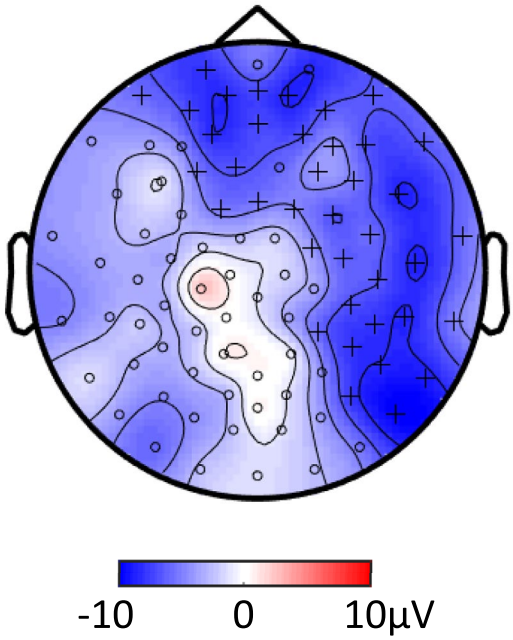

FIGURE 3 | Topographic difference map between the ELLA and the Flash Card groups. The crosses on the map indicate the electrodes where the negative potential at 200–300 ms of the ELLA group is significantly larger than that of the Flash Card group.

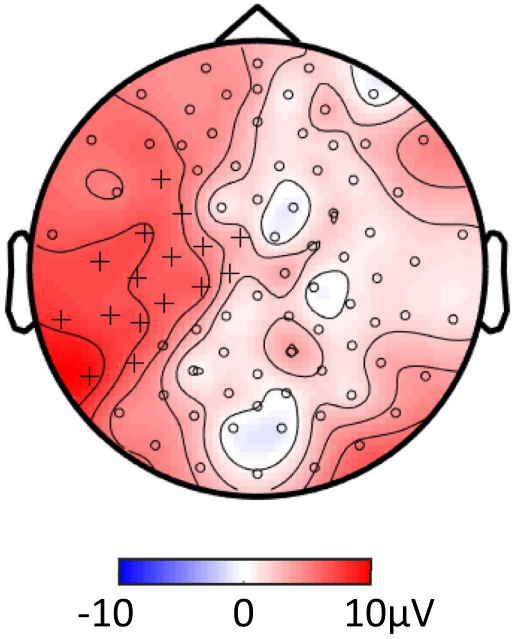

FIGURE 4 | Topographic difference map between the Known Word and the Unknown Word conditions. The crosses on the map indicate the electrodes where the positive potential at 200–300 ms of the Known Word condition is significantly larger than that of the Unknown Word condition.

Electrode 40 on Figure 5). To account for timing uncertainties, we also tested 150–300, 200–350 and 150–350 ms time-windows. All of these windows showed significant differences between unknown and known words ($p = 0.024$, $p = 0.028$ and $p = 0.036$, respectively). Similar results were found for the main effect comparing Flash vs. Ella ($p = 0.018$, $p = 0.023$ and $p = 0.023$, respectively).

There were no significant main effects and no significant interaction in the 400–600 ms window. In the 600–800 ms window, the interaction showed a marginal p value ($p = 0.062$).

4 | Discussion

In this study, we compared the auditory ERPs generated in response to Indonesian words learnt via the ELLA apps and digital flash cards, as compared to Indonesian words that the preschool children had not been exposed to. We hypothesised that children who learned from the ELLA apps would exhibit larger ERP components in both early and late time-windows, as compared to the group who learned from digital flash cards. We also hypothesised that children would exhibit larger ERP components in both early and late time-windows to the words that they were exposed to, as compared to the words that they were not exposed to. The early positive (200–300 ms) potential (of the whole group) to the known words was significantly larger than that of the unknown words in the left central electrodes, but there was no difference between the known and unknown words for the late potential (400–800 ms). Based on comparisons of ERPs generated by children who learnt via the ELLA apps vs. via the digital flash cards, the early negative potential (200–300 ms) of the ELLA group was significantly larger than that of the flash card group in the frontal and right regions, but there was no difference observed in the late time-windows (400–800 ms).

Our enhanced positive response to the known words was similar to the enhanced P3a exhibited by 3- to 6-year-old children after learning French for several weeks (Shestakova et al. 2003). Positive responses at around 200–350 ms have been associated with involuntary attention shifts to novel sounds (e.g., Escera et al. 1998, 2000) and to spoke stimuli (Shtyrov et al. 2012). Further to this, our centrally dominant response was similar to the topography and the latency of the early portion of P3a (Escera et al. 1998). The left-hemisphere dominance of our response provided further support that the response is of linguistic nature, especially when neural modulations in the left hemisphere, such as the left middle frontal gyrus and inferior frontal gyrus, have been associated with increases in second language proficiency (Mårtensson et al. 2012; Stein et al. 2012; Yang et al. 2015). Taken together, the observed enhanced positive response to known foreign words in the current study could be an indication of children switching their attention to the known foreign words due to their increased salience compared to the unknown words.

The children who learnt from the ELLA apps exhibited a larger early negative response to all stimuli (regardless of whether the stimuli were known or unknown) than the children who learnt from the digital flash cards. Negative responses with latency circa 250 ms have been associated with various domains of early cognitive processing, such as phonological (Connolly and Phillips 1994), lexical selection (van den Brink et al. 2001), sound content feature processing (Ceponiene et al. 2005) and

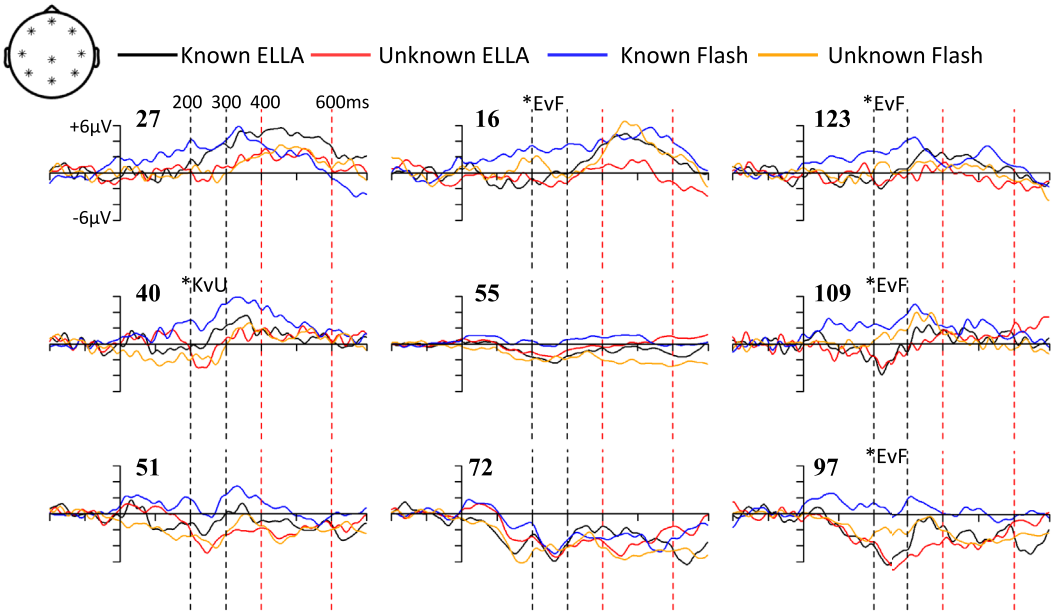

FIGURE 5 | Event-related potentials of the two Word conditions of the two groups: Known ELLA, Unknown ELLA, Known Flash, and Unknown Flash. The numbers in bold represent the electrode number. *KvU represents a significant difference between Known words and Unknown words. *EvF represents a significant difference between the ELLA and the Flash Card groups.

word form priming (Friedrich and Friederici 2008). Children who learnt from the ELLA apps might be able to better detect some of these features of the foreign language than the children who learnt from the digital flash cards.

Another possible explanation for this enhancement is related to word familiarity. Children as young as 11 months old have shown larger negative responses within 250 ms when they heard familiar compared to unfamiliar words (Thierry et al. 2003), and 20-month-old children have shown larger early negative responses to known words than to nonsense words (Mills et al. 2004). Similar to our observed topography, the responses circa 250 ms in previous research were more negative on the electrodes on the right hemisphere (Thierry et al. 2003) and in the right frontal sites (Mills et al. 2004). The similarity in topography as previous studies suggested that children who learnt from the ELLA apps were more familiar with the foreign words; hence, they might be more able to detect certain features of the foreign words better than the children who learnt from the digital flash cards.

The notion of word familiarity could be related to the ELLA group having spent more time on the apps than their peers on the digital flash cards at home. It is also possible that children could have learnt less from the Apps than from the digital flash cards in the same amount of time. This is because digital flash cards only present words to be learnt, whereas the Apps require a reasonable amount of non-linguistic actions to use (e.g., moving a polyglot around a screen or performing moves in a puzzle). However, without the presence of an experimenter, children in both groups could have behaved differently as expected and might have skipped or missed some words. To clarify this point, future research should include time-logs when each audio file was played.

There are a number of methodological issues that warrant discussion. First, the loudness of our Unknown Word stimuli was slightly louder than those of the known words, and the fundamental frequency of our Unknown Word stimuli was slightly higher than those of the known words in two sets of our apps. These loudness and frequency differences might affect the ERPs, and hence may be a possible explanation for the observed larger positive potential of the known words. However, it should be noted that louder and higher frequency stimuli usually produce larger auditory event-related components in both adults and children (e.g., Davies et al. 2010; Picton et al. 1974; Putkinen et al. 2012). Despite these differences in the stimuli, our softer and lower frequency (known) word stimuli actually produced a larger positive potential than the louder and higher frequency (unknown) words. It is, therefore, unlikely that the observed difference was due to the physical differences of the stimuli. Another reason that the loudness difference would be unlikely to confound the task is that the differences were less than 6 dB. Such a small difference should have a negligible effect on our results, especially when previous research showed a much larger loudness difference (i.e., 20 dB) led to a relatively small ERP amplitude modulation (Davies et al. 2010).

A second methodological issue with this study is that the number of accepted epochs in our EEG data did not allow us to further divide the Known Word condition according to children's accuracy of the task or of the behavioural word test results, as that would have severely impacted the signal-to-noise ratio of the ERP responses. In a behavioural task (separate from the EEG experiment), both groups of children were shown a grid of four pictures and were instructed to point out the correct picture corresponding to the known word the researcher read out. For the EEG cohort, both groups of children performed above chance level (25%), with the ELLA group at a mean accuracy

of 61.83% (SD: 17.94%) and the FLASH group significantly better at a mean accuracy of 76.54% (SD: 16.76%) ($p = 0.026$). The apparent better performance from the flash card group should be interpreted with caution. This is because there are comparatively fewer words and phrases in the digital flash cards than in the ELLA apps. This means that it is likely children from the flash card cohort would have had a greater number of exposures to the test words used than children in the ELLA cohort in the behavioural word test. While these data have shown that both groups of children learnt from both the apps and the digital flash cards, future studies will be necessary to link the behavioural word accuracy with the ERP responses, to examine if children's semantics and understanding of the foreign language contributed to the modulation of the electrophysiological response rather than other factors such as familiarity.

Two important points of this research warrant further clarification. The first point is to clarify the underlying mechanism of the enhanced response observed in the ELLA group: specifically, whether it was a representation of familiarity with the words or because of the children's enhanced ability to detect certain distinct features in the new language. The second is to investigate whether well-designed foreign language apps could lead to long-term neural modulation in children, especially when training-related enhancement in children's auditory ERP components has been observed months to 1 year post-training (Shestakova et al. 2003; Moreno et al. 2015).

To summarise, our study showed that (a) short-term training from either language-immersion apps or digital flash cards generated larger neural responses to known words than to unknown words; and (b) short-term training of a non-native language using language-immersion apps elicited larger brain responses in preschool children, compared to training using traditional flash cards. Although further research is needed to clarify the underlying mechanisms and the long-term effect, our study showed that short-term training with language-immersion apps modulates language processing in preschool children's brains differently compared to digital flash cards.

Author Contributions

Sumie Leung: formal analysis, investigation, methodology, project administration, supervision, writing – original draft, writing – review and editing. **Conrad Perry:** formal analysis, funding acquisition, investigation, methodology, software, supervision, validation, visualization, writing – review and editing. **Jessica Guy:** data curation, project administration, supervision, writing – review and editing. **Deborah Loats:** data curation, investigation, writing – review and editing. **Kate Highfield:** funding acquisition, investigation, methodology, supervision, writing – review and editing. **Jordy Kaufman:** funding acquisition, investigation, methodology, project administration, resources, software, supervision, writing – review and editing.

Acknowledgements

The authors would like to thank the participants and their families, as well as the Swinburne Babylab, for providing support during the data collection for this project. Open access publishing facilitated by Swinburne University of Technology, as part of the Wiley - Swinburne University of Technology agreement via the Council of Australian University Librarians.

Ethics Statement

This study was approved by the Human Research Ethics Committee at Swinburne University of Technology, and written informed consent was obtained from the parents of all child participants.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data are available upon reasonable request from the corresponding author, J.K., and with permission of the Department of Education, Skills and Employment.

Peer Review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/icd.70015.

References

ACARA. 2014. *Australian Curriculum: Languages Indonesian (Revised)*. Australian Curriculum Assessment and Reporting Authority. https://docs.acara.edu.au/resources/F-10_Australian_Languages_Indonesian_(revised)_for_public_viewing_-__February_2014.pdf.

Aladé, F., A. R. Lauricella, L. Beaudoin-Ryan, and E. Wartella. 2016. "Measuring With Murray: Touchscreen Technology and Preschoolers' STEM Learning." *Computers in Human Behavior* 62: 433–441. https://doi.org/10.1016/j.chb.2016.03.080.

Barac, R., E. Bialystok, D. C. Castro, and M. Sanchez. 2014. "The Cognitive Development of Young Dual Language Learners: A Critical Review." *Early Childhood Research Quarterly* 29, no. 4: 699–714. https://doi.org/10.1016/j.ecresq.2014.02.003.

Barbu, C., A. Gonzalez, S. Gillet, and M. Poncelet. 2019. "Cognitive Advantage in Children Enrolled in a Second-Language Immersion Elementary School Program for One Year." *Psychologica Belgica* 59, no. 1: 416–435. https://doi.org/10.5334/pb.469.

Benau, E. M., J. Morris, and J. W. Couperus. 2011. "Semantic Processing in Children and Adults: Incongruity and the N400." *Journal of Psycholinguistic Research* 40, no. 3: 225–239. https://doi.org/10.1007/s10936-011-9167-1.

Bialystok, E. 2007. "Acquisition of Literacy in Bilingual Children: A Framework for Research." *Language Learning* 57, no. Suppl 1: S45–S77. https://doi.org/10.1111/j.1467-9922.2007.00412.x.

Bialystok, E., K. F. Peets, and S. Moreno. 2014. "Producing Bilinguals Through Immersion Education: Development of Metalinguistic Awareness." *Applied PsychoLinguistics* 35, no. 1: 177–191. https://doi.org/10.1017/S0142716412000288.

Boersma, P. 2001. "Praat, a System for Doing Phonetics by Computer." *Glot International* 5, no. 9/10: 341–345.

Bonte, M., and L. Blomert. 2004. "Developmental Changes in ERP Correlates of Spoken Word Recognition During Early School Years: A Phonological Priming Study." *Clinical Neurophysiology* 115, no. 2: 409–423. https://doi.org/10.1016/s1388-2457(03)00361-4.

Bovo, R., E. Lovo, L. Astolfi, et al. 2018. "Speech Perception in Noise by Young Sequential Bilingual Children." *Acta Otorhinolaryngologica Italica* 38, no. 6: 536–543. https://doi.org/10.14639/0392-100X-1846.

Ceponiene, R., P. Alku, M. Westerfield, M. Torki, and J. Townsend. 2005. "ERPs Differentiate Syllable and Nonphonetic Sound Processing in Children and Adults." *Psychophysiology* 42, no. 4: 391–406. https://doi.org/10.1111/j.1469-8986.2005.00305.x.

Cheour, M., A. Shestakova, P. Alku, R. Ceponiene, and R. Näätänen. 2002. "Mismatch Negativity Shows That 3-6-Year-Old Children Can Learn to Discriminate Non-Native Speech Sounds Within Two Months." *Neuroscience Letters* 325, no. 3: 187–190. https://doi.org/10.1016/s0304-3940(02)00269-0.

Connolly, J. F., and N. A. Phillips. 1994. "Event-Related Potential Components Reflect Phonological and Semantic Processing of the Terminal Word of Spoken Sentences." *Journal of Cognitive Neuroscience* 6, no. 3: 256–266. https://doi.org/10.1162/jocn.1994.6.3.256.

Davies, P. L., W. P. Chang, and W. J. Gavin. 2010. "Middle and Late Latency ERP Components Discriminate Between Adults, Typical Children, and Children With Sensory Processing Disorders." *Frontiers in Integrative Neuroscience* 4: 16. https://doi.org/10.3389/fnint.2010.00016.

Dunn, L. M., and D. M. Dunn. 2007. *Peabody Picture Vocabulary Test (PPVT™-4)*. 4th ed. Pearson Education.

Electrical Geodesics Inc. n.d. www.ella.edu.au.

Escera, C., K. Alho, E. Schröger, and I. Winkler. 2000. "Involuntary Attention and Distractibility as Evaluated With Event-Related Brain Potentials." *Audiology and Neuro-Otology* 5, no. 3–4: 151–166. https://doi.org/10.1159/000013877.

Escera, C., K. Alho, I. Winkler, and R. Näätänen. 1998. "Neural Mechanisms of Involuntary Attention to Acoustic Novelty and Change." *Journal of Cognitive Neuroscience* 10, no. 5: 590–604. https://doi.org/10.1162/089892998562997.

Friedrich, M., and A. D. Friederici. 2008. "Neurophysiological Correlates of Online Word Learning in 14-Month-Old Infants." *NeuroReport* 19, no. 18: 1757–1761. https://doi.org/10.1097/WNR.0b013e328318f014.

Hermanto, N., S. Moreno, and E. Bialystok. 2012. "Linguistic and Metalinguistic Outcomes of Intense Immersion Education: How Bilingual?" *International Journal of Bilingual Education and Bilingualism* 15, no. 2: 131–145. https://doi.org/10.1080/13670050.2011.652591.

Jost, L. B., A. K. Eberhard-Moscicka, G. Pleisch, et al. 2015. "Native and Non-Native Speech Sound Processing and the Neural Mismatch Responses: A Longitudinal Study on Classroom-Based Foreign Language Learning." *Neuropsychologia* 72: 94–104. https://doi.org/10.1016/j.neuropsychologia.2015.04.029.

Kast, M., S. Elmer, L. Jancke, and M. Meyer. 2010. "ERP Differences of Pre-Lexical Processing Between Dyslexic and Non-Dyslexic Children." *International Journal of Psychophysiology* 77, no. 1: 59–69. https://doi.org/10.1016/j.ijpsycho.2010.04.003.

Kaufman, J., S. Leung, C. Perry, et al. 2017. *Evaluation of the Early Learning Languages Australia Apps: Final Report to the Australian Government Department of Education and Training*. Swinburne University of Technology.

Kutas, M., and K. D. Federmeier. 2011. "Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP)." *Annual Review of Psychology* 62: 621–647. https://doi.org/10.1146/annurev.psych.093008.131123.

Maris, E., and R. Oostenveld. 2007. "Nonparametric Statistical Testing of EEG- and MEG-Data." *Journal of Neuroscience Methods* 164, no. 1: 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

Marques, P. M., A. L. Mattiazzi, L. Ferreira, S. J. Opptiz, and E. P. V. Biaggio. 2021. "The Effect of Learning English on P300 in Children." *International Archives of Otorhinolaryngology* 25, no. 2: e284–e288. https://doi.org/10.1055/s-0040-1710304.

Mårtensson, J., J. Eriksson, N. C. Bodammer, et al. 2012. "Growth of Language-Related Brain Areas After Foreign Language Learning." *NeuroImage* 63, no. 1: 240–244. https://doi.org/10.1016/j.neuroimage.2012.06.043.

Mills, D. L., C. Prat, R. Zangl, C. L. Stager, H. J. Neville, and J. F. Werker. 2004. "Language Experience and the Organization of Brain Activity to Phonetically Similar Words: ERP Evidence From 14- and 20-Month-Olds." *Journal of Cognitive Neuroscience* 16, no. 8: 1452–1464. https://doi.org/10.1162/0898929042304697.

Minogue, J., and M. G. Jones. 2006. "Haptics in Education: Exploring an Untapped Sensory Modality." *Review of Educational Research* 76, no. 3: 317–348. https://doi.org/10.3102/00346543076003317.

Moreno, S., Y. Lee, M. Janus, and E. Bialystok. 2015. "Short-Term Second Language and Music Training Induces Lasting Functional Brain Changes in Early Childhood." *Child Development* 86, no. 2: 394–406. https://doi.org/10.1111/cdev.12297.

Nassaji, H. 2016. "Anniversary Article Interactional Feedback in Second Language Teaching and Learning: A Synthesis and Analysis of Current Research." *Language Teaching Research* 20, no. 4: 535–562. https://doi.org/10.1177/1362168816644940.

Neumann, M. M. 2014. "An Examination of Touch Screen Tablets and Emergent Literacy in Australian Pre-School Children." *Australian Journal of Education* 58, no. 2: 109–122. https://doi.org/10.1177/0004944114523368.

Neumann, M. M. 2016. "Young Children's Use of Touch Screen Tablets for Writing and Reading at Home: Relationships With Emergent Literacy." *Computers & Education* 97: 61–68. https://doi.org/10.1016/j.compedu.2016.02.013.

Neumann, M. M. 2018. "Using Tablets and Apps to Enhance Emergent Literacy Skills in Young Children." *Early Childhood Research Quarterly* 42: 239–246. https://doi.org/10.1016/j.ecresq.2017.10.006.

O'Toole, K. J., and K. N. Kannass. 2018. "Emergent Literacy in Print and Electronic Contexts: The Influence of Book Type, Narration Source, and Attention." *Journal of Experimental Child Psychology* 173: 100–115. https://doi.org/10.1016/j.jecp.2018.03.013.

Peltola, M. S., M. Kuntola, H. Tamminen, H. Hämäläinen, and O. Aaltonen. 2005. "Early Exposure to Non-Native Language Alters Pre-Attentive Vowel Discrimination." *Neuroscience Letters* 388, no. 3: 121–125. https://doi.org/10.1016/j.neulet.2005.06.037.

Picton, T. W., S. A. Hillyard, H. I. Krausz, and R. Galambos. 1974. "Human Auditory Evoked Potentials. I. Evaluation of Components." *Electroencephalography and Clinical Neurophysiology* 36, no. 2: 179–190. https://doi.org/10.1016/0013-4694(74)90155-2.

Piotrowski, J. T., and M. Krcmar. 2017. "Reading With Hotspots: Young Children's Responses to Touchscreen Stories." *Computers in Human Behavior* 70: 328–334. https://doi.org/10.1016/j.chb.2017.01.010.

Putkinen, V., R. Niinikuru, J. Lipsanen, M. Tervaniemi, and M. Huotilainen. 2012. "Fast Measurement of Auditory Event-Related Potential Profiles in 2-3-Year-Olds." *Developmental Neuropsychology* 37, no. 1: 51–75. https://doi.org/10.1080/87565641.2011.615873.

Rinker, T., P. Alku, S. Brosch, and M. Kiefer. 2010. "Discrimination of Native and Non-Native Vowel Contrasts in Bilingual Turkish-German and Monolingual German Children: Insight From the Mismatch Negativity ERP Component." *Brain and Language* 113, no. 2: 90–95. https://doi.org/10.1016/j.bandl.2010.01.007.

Rinker, T., V. L. Shafer, M. Kiefer, N. Vidal, and Y. H. Yu. 2017. "T-Complex Measures in Bilingual Spanish-English and Turkish-German Children and Monolingual Peers." *PLoS One* 12, no. 3: e0171992. https://doi.org/10.1371/journal.pone.0171992.

Russo-Johnson, C., G. Troseth, C. Duncan, and A. Mesghina. 2017. "All Tapped out: Touchscreen Interactivity and Young Children's Word Learning." *Frontiers in Psychology* 8: 578. https://doi.org/10.3389/fpsyg.2017.00578.

Schneider, J. M., S. Poudel, A. D. Abel, and M. J. Maguire. 2023. "Age and Vocabulary Knowledge Differentially Influence the N400 and

Theta Responses During Semantic Retrieval." *Developmental Cognitive Neuroscience* 61: 101251. https://doi.org/10.1016/j.dcn.2023.101251.

Shestakova, A., M. Huotilainen, R. Ceponiene, and M. Cheour. 2003. "Event-Related Potentials Associated With Second Language Learning in Children." *Clinical Neurophysiology* 114, no. 8: 1507–1512. https://doi.org/10.1016/s1388-2457(03)00134-2.

Shtyrov, Y., M. L. Smith, A. J. Horner, et al. 2012. "Attention to Language: Novel MEG Paradigm for Registering Involuntary Language Processing in the Brain." *Neuropsychologia* 50, no. 11: 2605–2616. https://doi.org/10.1016/j.neuropsychologia.2012.07.012.

Smeets, D. J. H., and A. G. Bus. 2015. "The Interactive Animated e-Book as a Word Learning Device for Kindergartners." *Applied PsychoLinguistics* 36, no. 4: 899–920. https://doi.org/10.1017/S0142716413000556.

Smith, M. E., and K. Guster. 1993. "Decomposition of Recognition Memory Event-Related Potentials Yields Target, Repetition, and Retrieval Effects." *Electroencephalography and Clinical Neurophysiology* 86, no. 5: 335–343. https://doi.org/10.1016/0013-4694(93)90046-x.

Stein, M., A. Federspiel, T. Koenig, et al. 2012. "Structural Plasticity in the Language System Related to Increased Second Language Proficiency." *Cortex* 48, no. 4: 458–465. https://doi.org/10.1016/j.cortex.2010.10.007.

Steinhauer, K., and J. F. Connolly. 2008. "Event-Related Potentials in the Study of Language." In *Handbook of the Neuroscience of Language*, edited by B. Stemmer and H. Whitaker. Elsevier. https://doi.org/10.1016/B978-0-08-045352-1.00009-4.

Thierry, G., M. Vihman, and M. Roberts. 2003. "Familiar Words Capture the Attention of 11-Month-Olds in Less Than 250 ms." *NeuroReport* 14, no. 18: 2307–2310. https://doi.org/10.1097/00001756-200312190-00004.

Tong, X., C. McBride, J. Zhang, et al. 2014. "Neural Correlates of Acoustic Cues of English Lexical Stress in Cantonese-Speaking Children." *Brain and Language* 138: 61–70. https://doi.org/10.1016/j.bandl.2014.09.004.

van den Brink, D., C. M. Brown, and P. Hagoort. 2001. "Electrophysiological Evidence for Early Contextual Influences During Spoken-Word Recognition: N200 Versus N400 Effects." *Journal of Cognitive Neuroscience* 13, no. 7: 967–985. https://doi.org/10.1162/089892901753165872.

Wang, F., H. Xie, Y. Wang, Y. Hao, J. An, and J. Chen. 2016. "Using Touchscreen Tablets to Help Young Children Learn to Tell Time." *Frontiers in Psychology* 7: 1800. https://doi.org/10.3389/fpsyg.2016.01800.

Xie, H., J. Peng, M. Qin, X. Huang, F. Tian, and Z. Zhou. 2018. "Can Touchscreen Devices Be Used to Facilitate Young Children's Learning? A Meta-Analysis of Touchscreen Learning Effect." *Frontiers in Psychology* 9: 2580. https://doi.org/10.3389/fpsyg.2018.02580.

Yang, J., K. M. Gates, P. Molenaar, and P. Li. 2015. "Neural Changes Underlying Successful Second Language Word Learning: An fMRI Study." *Journal of Neurolinguistics* 33: 29–49. https://doi.org/10.1016/j.jneuroling.2014.09.004.

Yowaboot, C., and A. Sukying. 2022. "Using Digital Flashcards to Enhance Thai EFL Primary School Students' Vocabulary Knowledge." *English Language Teaching* 15, no. 7: 61–74. https://eric.ed.gov/?id=EJ1353241.