Integration of verbal and nonverbal information from picture storybooks: An Event-Related Potential study with adults.

Master Thesis

Thijs M.J. Nielen
Leiden University, The Netherlands

Developmental Psychopathology in Education and Child Studies

Specialization: Learning Problems and Impairments

Supervisor: Dr. M. T. de Jong

Second reader: Prof. Dr. A. G. Bus

Abstract

In an experimental Event-Related Potential (ERP) study amongst 7 highly-educated Dutch speaking females (aged 21-36), the neurological processing of congruent and incongruent picture-word combinations was investigated. Electroencephalogram (EEG) recordings were obtained prior to and after a single reading of the picture storybooks. Behavioral results showed that participants were better able to decide whether a picture-word combination was congruent or incongruent after they knew the story context from which the words and pictures were taken. Furthermore, ERP data analysis revealed the occurrence of an N400 (peak related to semantic incongruencies) following the presentation of incongruent picture-word combinations. Implications and directions for future research are discussed.

One of the important predictors of children's academic achievement is their vocabulary knowledge when entering Grade 1 (Hemphill & Tivnan, 2008). Therefore, research regarding the development of vocabulary knowledge and interventions to enhance vocabulary knowledge in preschoolers are important to prevent that children fall behind early in their school career. One way to enhance children's vocabulary knowledge is through shared book reading (Bus, Van IJzendoorn & Pellegrini, 1995; Mol, Bus & De Jong, 2009). Furthermore, encounters with both digital picture storybooks (storybooks with additional animations, sounds and special effects such as zooming) and static picture storybooks result in receptive and expressive vocabulary learning and digital picture storybooks are especially facilitative for expressive vocabulary learning (Verhallen & Bus, 2010).

Processing of Verbal and Nonverbal Information

A possible explanation for the beneficial effects of picture storybooks can be derived from Paivio's dual-coding theory (DCT, 2007). Paivio states that a combination of verbal and nonverbal information adds up to a level of comprehension that is larger than merely the sum of the separate verbal and nonverbal parts. According to Paivio (2007), verbal and nonverbal information (referred to as logogens and imagens) are processed in separate cognitive systems in the brain that form referential connections to support each other and thereby construct meaning. In picture storybooks the text (logogens) is accompanied by illustrations (imagens) resulting in enhanced comprehension.

When applying the DCT to a situation in which a child reads a picture storybook, the following processes can be distinguished:

- The verbal information of the story text is processed in the verbal symbolic system in the child's mind.
- The nonverbal information (picture) is processed in the nonverbal symbolic system in the child's mind.
- Referential connections are formed between the mental representation of the text and the picture.
- Through the referential connections the child is better able to understand both the text and the picture.

The processing steps are non-sequential in nature, verbal and nonverbal information is processed simultaneously in the two symbolic systems. Furthermore, the referential connections are continuously updated during reading in light of new information. Therefore, the child's story comprehension is continuously developing following new verbal and nonverbal information.

In line with the DCT it was found that illustrations play a crucial role in bridging inferences (Pike, Barnes & Barron, 2010). Bridging inferences and vocabulary are both important for reading comprehension. Furthermore, vocabulary knowledge supports the making of correct bridging inferences (Pike et al., 2010; Singer, Andrusiak, Reisdorf & Black, 1992). However, in the study of Pike and colleagues (2010) only young children (Grade 2-4) needed the nonverbal support to make a correct inference. For the older children in the study (Grade 5 & 6) the nonverbal support was not necessary. This finding suggests that nonverbal information is especially beneficial if the verbal information is difficult to comprehend, for instance due to limited vocabulary knowledge, as is often the case for young children when reading a picture storybook.

According to Stein and colleagues (2010), the simultaneous processing of information obtained in different sensory modalities, such as seeing a picture and hearing text at the same time, results in a neural product that resembles a combination of the information obtained from different modalities. Stein and colleagues (2010) refer to the process of combining information from multiple sensory modalities as multisensory integration. In line with Paivio's DCT (2007), they argue that the multisensory integration results in a product that differs from the sum of the separate parts. Whether this product reinforces or undermines comprehension depends on the fit of the information derived from different sensory modalities. However, studies regarding the multisensory integration of verbal and nonverbal information are scarce and focus on the behavioral level (e.g. Mayer & Sims, 1994).

Evidence for the proposed neurological mechanisms underlying picture-text integration is limited. Integration of verbal and nonverbal information is a very fast and often unconscious process. For example in a Stroop task a certain color (e.g. red) must be named as quickly as possible. If however, interfering verbal information is presented (e.g. the color red is displayed as letters forming the word 'blue') the participant is slower in naming the color compared to a situation without interfering information (the color red is displayed in a red square) (Stroop, 1935). This finding shows that verbal and nonverbal information are processed simultaneously and can support (e.g. Paivio, 2007; Stein et al., 2010) or interfere (Stroop, 1935) with each other. The support or interference is generally a very quick process, for instance a Stroop effect resulted in an average increase in naming time of 470 milliseconds (ms) per stimulus (Stroop, 1935). It is therefore necessary to use a technique with a high temporal resolution to reveal the neural processes underlying the integration of verbal and nonverbal information. Furthermore, this technique should provide insight in the neurological activity while pictures and text are presented to the participant. An electroencephalogram (EEG), and more specifically the event-related potential (ERP) is a technique through which this goal can be obtained.

Event-Related Potential (ERP)

Event-related potentials (ERPs) are part of an electroencephalogram (EEG) and reflect the electrical activity associated with neural responses following an event. The main advantage of ERP recordings is the high temporal resolution (Luck, 2005). ERPs are observed following many different experimental manipulations, such as oddball tasks (e.g. Itagaki, Yabe, Mori, Ishikawa, Takanashi & Nawi, 2011), semantic and syntactic violations in sentences (e.g. Silva-Pereyra, Conboy, Klarman & Kuhl, 2007) and the presentation of two pictures that do not match (e.g. Barret & Rugg, 1990).

One ERP component that is particularly interesting in the context of the present study is the N400. The N400 is a negative deflection in the waveform that starts after 250 ms and peaks approximately 400 ms after the onset of a word or picture that does not matches the context. Typically, the N400 is observed in the central-parietal regions and is slightly stronger in the right hemisphere (Kutas & Federmeier, 2011). This component is sensitive to the perceived deviance; words or pictures that are considered more deviant in the context result in larger negative peaks (Kutas & Hillyard, 1980). Previous research has shown that the N400 effect occurs after incongruencies in verbal tasks (e.g. reading sentences with an incongruent ending; Van Berkum, Hagoort & Brown, 1999) as well as in nonverbal tasks (e.g. comparing line drawings; Barret & Rugg, 1990), and when a sentence is presented with an incongruent visual stimulus (e.g. Ganis, Kutas & Sereno, 1996).

It is hypothesized that if the N400 is stronger in the processing of words that do not match a picture storybook illustration (incongruent) compared to words that do match a picture storybook illustration (congruent), it can be concluded that the processing of the two stimuli is integrated on a neural level. According to Paivio’s DCT (2007) and the concept of multisensory integration as proposed by Stein and colleagues (2010) integrated processing of the two stimuli will support comprehension of both the verbal and nonverbal information.

Present Study

The aim of the present pilot study is to examine the integration of verbal and nonverbal information from picture storybooks on a neurological level. Whereas previous studies have shown that picture storybooks are effective for (expressive) vocabulary learning and the recall of implied story elements (Verhallen & Bus, 2010; Verhallen, Bus & De Jong, 2006), these studies rely on behavioral measures. The goal of the present study is to add to these findings by examining verbal and nonverbal information processing on a neuropsychological level. Measures at the neuropsychological level can provide more insight in the processes underlying word learning and story comprehension in picture storybooks.

Previous ERP studies have used simple nonverbal stimuli (e.g. line drawings of common objects; Barret & Rugg, 1990), verbal stimuli (e.g. Van Berkum et al., 1999), or a combination of simple visual and verbal stimuli (e.g. Ganis et al., 1996). In the present study complex pictures from picture storybooks are combined with auditorily presented target words. Participants are tested prior to and after the reading of a picture storybooks from which the words and pictures were taken and have to indicate whether the combination of picture and word is congruent or incongruent while an

EEG measure is conducted. The following research questions will be addressed in this study:

- *Is there a difference in ERP signal between the picture-word combinations in the congruent condition and the picture-word combinations in the incongruent condition?*
- *Is there a difference in ERP signal between the pre- and posttest for the picture-word combinations from the storybooks that the participants read as compared to the pre- and posttest ERP signal for the storybooks that the participants have not read?*

Hypotheses. It is expected that, in line with previous studies regarding semantic anomalies (e.g. Van Berkum et al., 1999; Kutas & Hillyard, 1980), there will be a difference between the target words that match with the storybook picture (congruent) and target words that do not match with the storybook picture (incongruent). More specifically, the waveform in the incongruent picture-word condition will be more negative than the waveform in the congruent condition after at least 250 ms with a peak at approximately 400 ms after the onset of the target word.

Whereas adults who are expected to be familiar with the target words participate in the present pilot study, a difference in ERP signal is expected in both the pre- and the posttest. However, due to the sensitivity of the N400 to differences in congruency (Kutas & Hillyard, 1980), it is expected that the N400 for incongruent picture-word combinations will be stronger in the posttest than in the pretest due to the picture storybook reading. This hypothesis is supported by previous studies using specific ERP components to evaluate intervention effectiveness on a neuropsychological level (Huotilainen, Lovio, Kujala, Tommiska, Karma & Fellman, 2011; Fukuda et al., 1997; Woltering, Granic, Lamm & Lewis, 2011) and by promising findings regarding interventions with picture storybooks (Verhallen, Bus & De Jong, 2006; Verhallen & Bus, 2010). For instance, Huotilainen and colleagues (2011) evaluated a dyslexia remediation programme by examining the changes in the mismatch negativity (MMN, an ERP signal reflecting preattentive auditory discrimination). They found a change both in duration and frequency when comparing the MMN in the pretest to the MMN in the posttest indicating that the programme had an effect on the neuropsychological level.

If a significantly larger negativity is found in the incongruent than in the congruent condition with a peak at approximately 400 ms after the presentation of the target word, it can be concluded that the N400 is also present in a more ecologically valid setting regarding information processing from picture storybooks than in previous studies regarding this ERP peak. Furthermore, an increase in the signal difference between the picture word combinations from the books in the intervention in the posttest as compared to the pretest is an indication that the N400 is indeed sensitive to small differences in perceived congruency. Finding the N400 in the present study supports the hypothesis that the processing of pictures and words is integrated on the neurological level, in line with Paivio's DCT (2007) and the multisensory integration hypothesis of Stein and colleagues (2010).

Method

Participants

Seven right-handed females (aged 21-36, *M* = 24.43, *SD* = 5.19) participated in the present study. All participants were highly educated (Bachelor up till PhD degree), spoke Dutch as their first language and had normal or corrected-to-normal vision.

Participants were recruited from the environment of the researcher (acquaintances and colleagues). No medication use was known by the experimenter except for the use of a respiratory and cerebral stimulant (dexamphetamine) reported by one participant. Informed consent was obtained prior to the study. The study was approved by the ethics committee of the Leiden Institute of Education and Child Studies.

Research Design

Storybook reading. A within-subject experimental design with two conditions was used. Four picture storybooks were selected for the intervention condition and four books for the control condition. The storybook reading consisted of a single session in which four books were read twice to the participants while they watched the dynamic pictures. The stories were told in the same pace and intonation with the same animation-text combinations for all participants. The remaining four books formed the control condition and were not presented in the intervention (see Appendix A for an overview of the picture storybooks used in each condition).

Pre- and posttest. To examine the differences between the EEG signal in congruent and incongruent picture-word combinations, pre- and posttest EEG sessions were conducted. A description of the target word selection and content of the EEG sessions is provided in the following section.

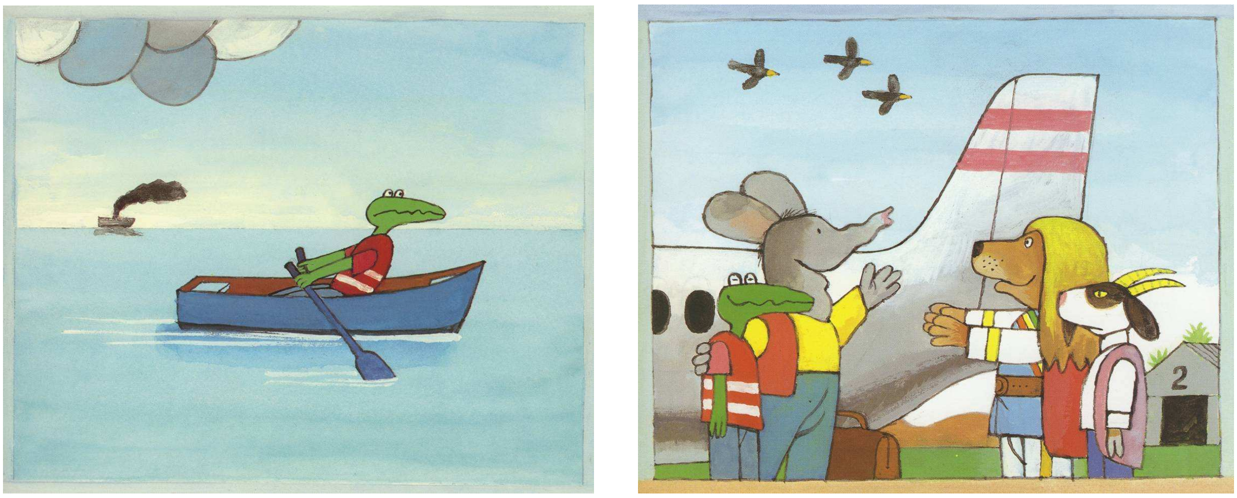
Test Materials

Target word selection. A total of 64 target words and accompanying illustrations were selected from 8 Dutch picture storybooks (see Appendix B). The target word selection was based on a previous pilot study in adults (N = 19, 32% males, age range 17-47, M = 24.05, SD = 6.21). In the pilot study the participants had to fill in the missing word in a sentence from a picture storybook. The missing words were possible target words for the present study selected by the researchers and the sentences were presented first without and then with the accompanying illustration from the picture storybook. Two criteria were taken into account by the final selection of the target words:

1. The word had to be too difficult to guess without illustration.
2. The illustration had to increase the chance of correctly guessing the target word.

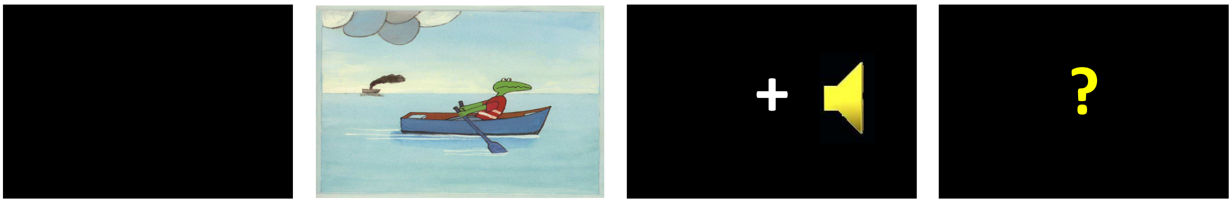
Based on the pilot study a selection was made of 64 words from eight different picture storybooks (8 per book).

Target word presentation. Each target word was presented in the pre- and posttest EEG recording sessions in a congruent and an incongruent condition. In the congruent condition, target words were combined with the picture that is simultaneously presented with the target word in the storybook. In the incongruent condition the target word is combined with a picture that does not match the target word. The target words were combined with a picture from the same book in both conditions. An example of a congruent and an incongruent picture-word combination is provided in Figure 1. Furthermore, Figure 2 provides an overview of a single trail in the EEG recording sessions.



‘roeien’ [rowing], congruent ‘roeien’ [rowing], incongruent

Figure 1. Example of a congruent and an incongruent picture-word combination.



1250 + 10**X* ms 3000 ms 850 ms + 2400 ms Till response

Figure 2. Example of a trial for the EEG experiment. The black screen intervened between trials and *X* indicates a random number between 0 and 50. The fixation cross appeared after the picture and the target word followed 850 ms after the appearance of the fixation cross, here indicated by the speaker symbol. The next trial started after the participant responded. *Note*: the speaker symbol is not visible in the actual experiment.

Procedure

Storybook reading. During the storybook reading session the books were presented in a randomized order for each participant and they watched the books the second time in the same order as the first time. Watching each picture story book took about five minutes, and the total duration of the intervention was approximately 50 minutes. Participants were instructed to carefully watch the dynamic pictures while the story was read. During the intervention participants were either alone in the room or with the experimenter. Participants participated in the intervention and in the posttest within seven days after the pretest.

Pre- and posttest. All participants took part in two EEG sessions, a pretest prior to the intervention and a posttest following the intervention. Both EEG recording sessions consisted of two blocks and picture-word combinations were randomized such that each block contained 32 congruent and 32 incongruent picture-word combinations. Furthermore, the block order was counterbalanced in the pre- and posttest. Finally, the picture-word combinations within one block were randomly presented.

The picture-word combinations were presented using E-prime version 2.0.8.22 (Psychology Software Tools). Participants were tested in a sound attenuated booth in the EEG-lab in a comfortable chair in a dimly lit room. The distance between the participant and the monitor was approximately 75 cm and the target words were presented at 65 dB. Participants responded to 128 trials with short breaks after 16 trials and a longer break between the first and second block of the experiment (after 64 trials). During the breaks the experimenter repeated instructions to the participant if necessary (e.g. ‘try not to move too much’). Responses for congruent and incongruent matches were provided with the index and the middle finger of the preferred hand. Each EEG recording session lasted approximately 22 minutes.

EEG Recording and Preprocessing

EEG recordings were conducted with the 129-channel hydrocel geodesic sensor net version 1.0 and amplified with a NetAmps300 amplifier (Electrical Geodesics, Inc.). Sampling rate was 250 Hz and Cz (VREF in EGI) was used as the reference electrode. Impedances for all electrodes were kept below 100 kΩ.

Offline data processing has been conducted using Brain Vision Analyzer version 2.0 (Brain Products). A passband range of 0.5-30 Hz was used to filter the data and segments from 200 ms prior to the presentation of the target word till 800 ms after the target word onset were selected. These segments were rereferenced to the average activity in all channels and corrected for ocular artifacts based on an independent component analysis (ICA, vertical EOG electrodes 127-25, horizontal EOG electrodes 128-125). Entire segments were removed when the difference between the minimum and the maximum activity exceeded 70 µV in the vertical EOG channel (electrodes 126-8) and 50 µV in the horizontal EOG channel (electrodes 1-32). Furthermore, individual electrode signals were removed if the difference between the minimum and maximum activity exceeded 200 µV. Finally, all remaining segments were baseline corrected to a 200 ms baseline immediately preceding the target word onset.

All experimental sessions in which less than 60% of the trials remained were excluded from further analyses. This exclusion criterion resulted in the total exclusion of one participant and the additional loss of two pretest sessions and one posttest session for other participants. In other words, the data from 9 out of 14 sessions was of sufficient quality to include in further analyses. For these sessions, on average 96.00 ($SD = 6.27$) segments in which the participant correctly decided if the picture-word combination was congruent or incongruent remained in the pretest and 96.40 ($SD = 15.36$) in the posttest. All incorrect segments were excluded in order to ensure that only trials are included in which the participant noticed the semantic congruency or incongruency of the picture-word combination.

In order to provide a reliable sample of a region's activity, electrodes were clustered in 11 clusters corresponding to the international 10/20 system (see Table 1 and for a comparable approach Yang, Perfetti & Schmalhofer, 2007). Clustering has the advantage that spurious interactions within a single electrode location are reduced (Dien & Santuzzi, 2005). The location of the clusters and the accompanying EGI electrode numbers are provided in Figure 3. Electrode 62 was originally part of the Pz cluster, however, visual inspection of the electrodes revealed high frequency noise on this channel for multiple participants. Therefore, the electrode was excluded from further analyses.

Table 1.

EGI electrode clusters according to the international 10/20 system.

EGI electrode number	24	11	124	36	VREF	104	60	72	85	45	108
10/20 system location	F3	Fz	F4	C3	Cz	C4	P3	Pz	P4	T3	T4

Data Analysis

Two complementary analyses were applied to examine the ERP results (e.g. Yang et al., 2007). Topographic maps were created to explore the spatiotemporal dynamics of the ERP data in the congruent (412 segments) and in the incongruent (454 segments) conditions. Furthermore, for the time window in which the N400 effect occurred (350-550 ms after target word onset), the mean amplitude voltage of the difference wave was exported. A difference wave was chosen due to the relative rather than the absolute negativity of the N400 (Huffmeijer, Tops, Alink, Bakermans-Kranenburg & Van IJzendoorn, 2011). The difference wave was constructed by subtracting the congruent voltage from the incongruent voltage. This operationalization implicated that negative values indicated a stronger negativity in the incongruent than in the congruent condition. The difference wave extracted from the N400 timeframe was compared to the difference wave in the time frame from 50-250 ms after target word onset, in which no difference between the two conditions was expected.

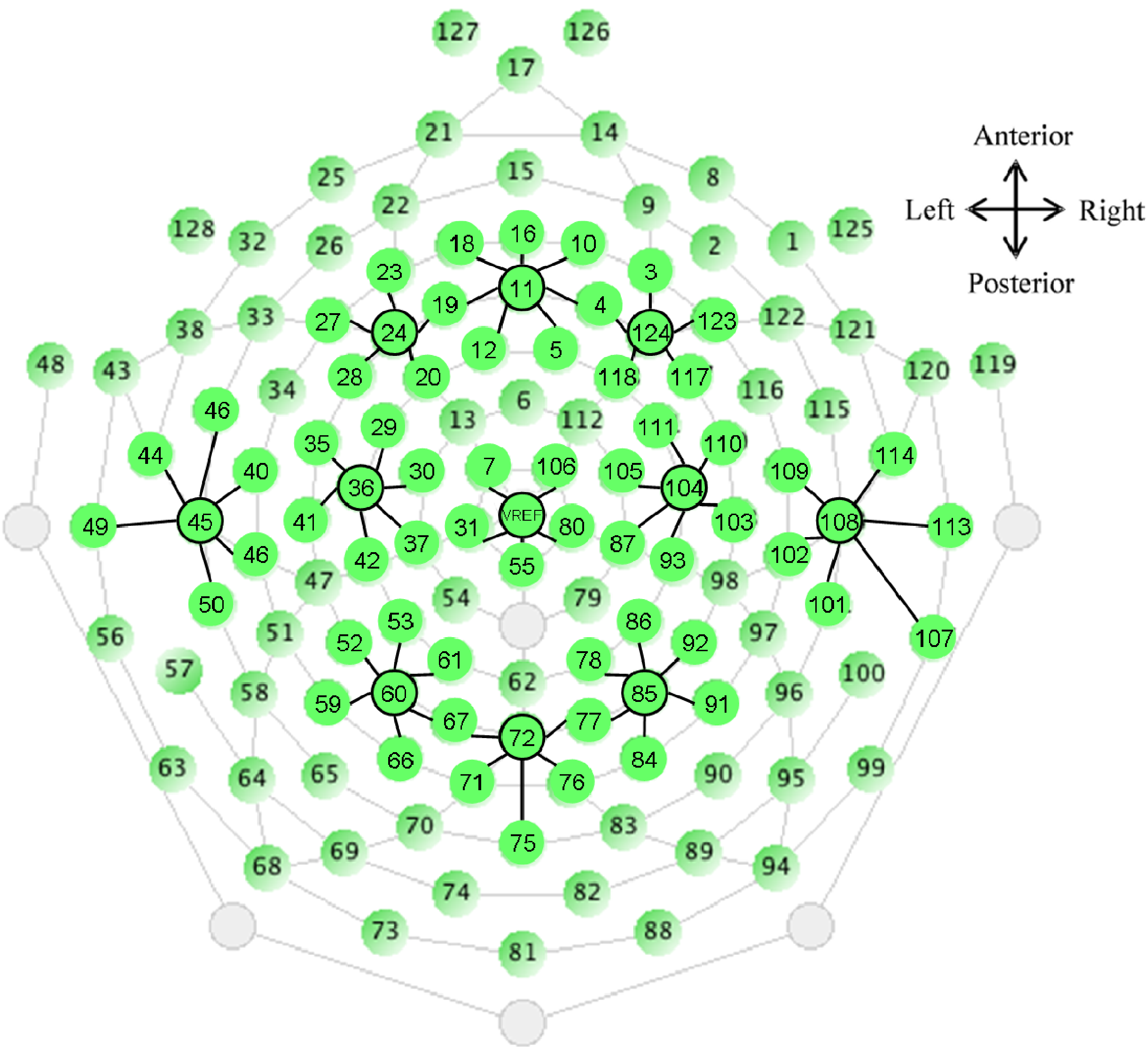


Figure 3. The 11 electrode clusters with the accompanying EGI electrode numbers.

Mean amplitude voltage for the electrodes in the difference wave was exported and a single microvolt value for each cluster per time frame was calculated by computing an average in SPSS. Finally, two repeated measures ANOVAs were performed for the mean amplitude voltages per cluster: one for the medial electrode sites with the factors Condition (50-250 ms/350-550 ms) and Electrode location (Fz/Cz/Pz) and one for the lateral electrode sites with the factors Condition, Electrode location (F3-F4, C3-C4, P3-P4, and T3-T4) and Hemisphere (left/right).

Results

Behavioral Results

The number of correct semantic decisions in the pre- (M = 52.17, SD = 4.71) and posttest (M = 54.00, SD = 2.68) did not differ for the books in the control condition ($t(5) = -1.23$, $p = .274$). However, for the books in the intervention condition, a higher number of correct semantic decisions was made in the posttest (M = 57.67, SD = 5.79) compared to the pretest (M = 50.33, SD = 2.42) ($t(5) = -3.223$, $p = .023$). This indicates that participants were better able to decide if a picture-word combination was congruent or incongruent after reading the storybooks twice.

EEG Data Analysis

As a consequence of the small sample size, it was not possible to examine the intervention effect using the EEG-recordings. The number of correctly answered trials in the intervention and control condition separately was too small to reliably average out noise and maintain the ERP waveform (Luck, 2005). In order to increase the power of the analyses and extract the ERP of interest, the data obtained from all correctly answered trials (pre- & posttest; control & intervention condition) are combined.

Data Inspection

Prior to the ERP data analysis, a data inspection procedure was used to examine cases outside the 95% confidence interval around the mean. A total number of 7 outliers was detected and winsorized (Naus, 1975); three in the congruent and four in the incongruent condition. The repeated measures ANOVA assumptions of normality of the distributions were not entirely met, a number of minor skewness and kurtosis violations were observed. However, the F distribution is fairly robust for deviations of normality (e.g. Maxwell & Delaney, 2004; Lix, Keselman & Keselman, 1996). Furthermore, the assumption of sphericity was met for both repeated measures ANOVAs as indicated by Mauchly's test of sphericity ($p > .01$ for all contrasts). It was therefore concluded that the F-test was valid to analyze the present data.

Topographic Maps

Topographic maps for the congruent and incongruent condition in the early and late time windows are displayed in Figure 4. A negative peak in the central and parietal regions was visible in the late time window (350-550 ms) in the incongruent condition. This peak was not observed in the earlier time window (50-250 ms), or in the congruent condition. Furthermore, the topographic plot for the incongruent condition in the late time window shows a voltage distribution in which the negative peak is larger in the right hemisphere.

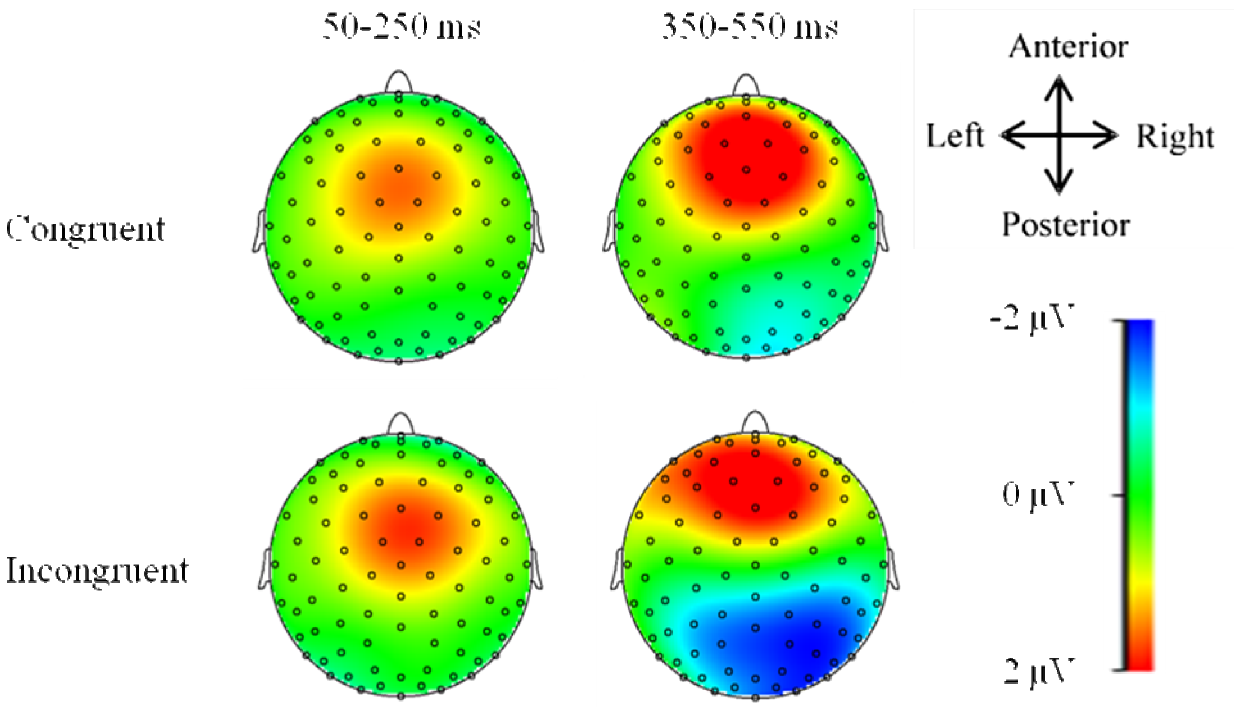


Figure 4. Topographic map for the congruent and incongruent condition in the two time windows.

Mean Voltage Amplitude Analyses

Statistics for the two repeated measures ANOVAs are displayed in Table 2. For the medial ANOVA the effect of location (Fz/Cz/Pz) on the mean difference wave amplitude in the two conditions (50-250 ms/350-550 ms) was examined. Furthermore, the lateral ANOVA addressed the effect of location (Fz/Cz/Pz/Tz) and hemisphere (left/right) on the mean difference wave amplitude in the two conditions (50-250 ms /350-550 ms). No significant main effects were found for the repeated measures ANOVAs meaning that there were no direct effects for condition, location or hemisphere. However, for both the medial and the lateral ANOVA an interaction effect of condition and location was found (see Figure 5).

Table 2.

Statistics for the medial and lateral repeated measures ANOVAs on the mean difference wave amplitude for the early (50-250 ms) and N400 (350-550 ms) time windows.

Medial ANOVA	*df*	*F*	*p*	η^2
Condition	1,5	2.72	.160	.06
Location	2,10	.40	.682	.03
Condition*Location	4,20	4.80	.035	.18
Lateral ANOVA				
Condition	1,5	1.09	.344	<.01
Location	3,15	2.64	.087	.15
Hemisphere	1,5	.28	.617	<.01
Condition*Location	5,25	9.04	.001	.16
Condition*Hemisphere	3,15	.69	.444	<.01
Location*Hemisphere	5,25	2.37	.111	.04
Condition*Location*Hemisphere	7,35	.78	.521	<.01

Note. The medial ANOVA examines the effect of electrode site (Fz/Cz/Pz) on the difference wave amplitude in the two conditions (50-250&350-550 ms). The lateral ANOVA examines the effect of electrode location (F3-F4/C3-C4/P3-P4/T3-T4) and hemisphere (left: F3/C3/P3/T3 & right: F4/C4/P4/T4) on the difference wave amplitude in the two conditions (50-250 & 350-550 ms).

In order to further specify the significant interaction effect in both the medial and the lateral ANOVA, a series of paired-samples t-tests was conducted to further specify the nature of the interaction effects. For the medial ANOVA t-tests were conducted to compare the means of the two time windows for each separate location. A difference in the mean voltage amplitude was found for the frontal medial (Fz) electrode cluster only ($t(5) = -3.94$, $p = .011$), a stronger negativity was observed in the early time window (50-250 ms, $M = -.43$, $SD = .15$) than in the late time window (350-550 ms, $M = .00$, $SD = .38$).

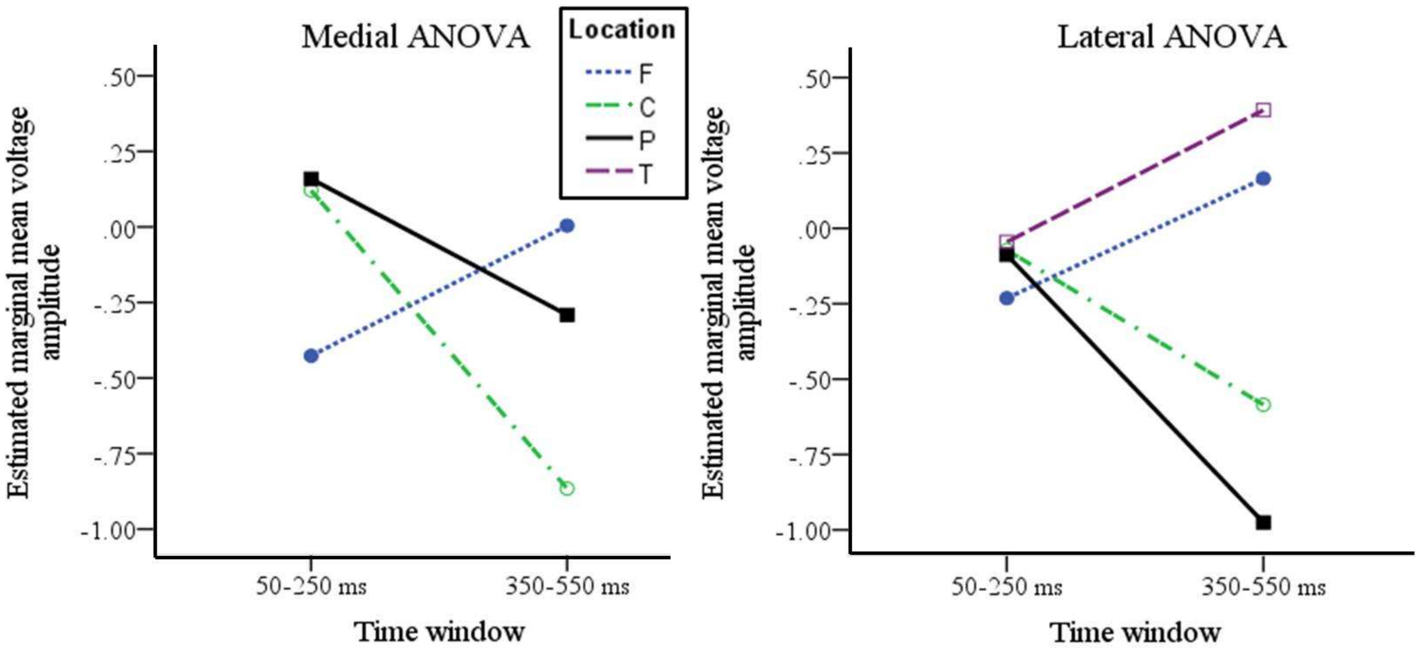


Figure 5. Interaction effect for the two time windows and the locations for both the medial and the lateral repeated measures ANOVAs. The estimated marginal mean voltage amplitude is displayed on the y-axis and the early (50-250 ms) and N400 (350-550 ms) time windows are displayed on the x-axis. *Note:* F = Frontal, C = Central, P = Parietal and T = Temporal.

Whereas no significant effect for hemisphere was found, the regions in the left and right hemisphere (e.g. F3 and F4) can be combined to further examine the interaction effect of condition and location in the lateral ANOVA. Therefore, the mean voltage amplitude for each spatial region (frontal, central, parietal and temporal), independent of hemisphere, was used in further analyses. In line with the follow-up tests for the medial ANOVA the means for each spatial region were compared for the early time window and the N400 time window. For both the central ($t(5) = 3.03$, $p = .029$) and the parietal ($t(5) = 2.60$, $p = .048$) regions a difference in the mean voltage amplitude was found. In the central regions a larger negativity was found in the N400 time window (*M* = -.59, *SD* = .17) than in the early time window (*M* = -.07, *SD* = .39). A comparable effect was found for the difference between the early (*M* = -.09, *SD* = .26) and late (*M* = -.98, *SD* = 1.03) time window in the parietal regions.

Discussion

The adult participants in the present study were able to make a correct semantic decision for the majority of the picture-word combinations. On average, more correct decisions were made in the posttest as compared to the pretest, but only for the books that the participants read, which indicates a positive effect of the book reading session. Unfortunately, the sample was too small to examine these findings at the neuropsychological level.

The results of the analyses of the EEG-data in the present pilot study support the hypothesis of a time window specific difference between the ERP signal in the congruent and in the incongruent condition. The N400 effect as found in a large number of studies using a broad range of stimuli (e.g. Liao, Su, Wu & Qiu, 2011; Kutas & Hillyard, 1980; Caldera, Jermann, Lopez Aranga & Van der Linden, 2004), is also apparent in the present study's design with highly complex visual stimuli from picture storybooks matched with target words from the same books. A significant negativity in central and parietal electrode clusters was found in the incongruent condition. As expected, this effect was only present in the late time window and for the lateral electrode regions. Furthermore, the spatiotemporal distribution as shown in Figure 4 resembled all characteristics of the N400: the peak occurs at approximately 400 ms (Kutas & Hillyard, 1980), is apparent in the central-parietal regions (Khateb, Pegna, Landis, Mouthon & Annoni, 2010) and the negativity seems to be slightly stronger in the right hemisphere (Kutas & Federmeijer, 2011).

The stronger positive signal which was found in the frontal electrode cluster for the late time window as compared to the early time window was an unexpected finding. This frontal positivity is generally not included in descriptions of the N400 effect (e.g. Kutas & Hillyard, 1980; Kutas & Federmeijer, 2011). Furthermore, the lack of a main effect for the condition factor was not expected. However, these two findings might be a consequence of the bipolar nature of EEG recordings. Indeed, Johnson and Hamm (2000) report the occurrence of a positive deflection during an N400 at approximately 400 ms for the more extreme located electrode sites in especially the temporal lobes (e.g. electrode 49 and 114). It is therefore possible that the electrode clustering not only resulted in less spurious interactions (Dien & Santuzzi, 2005), but in a positivity after approximately 400 ms for the temporal and, in this study, frontal electrode clusters as well. This hypothesis is supported by the results of the lateral and medial ANOVAs as displayed in Figure 5 and the topographic maps as displayed in Figure 4. An increase in voltage is observed for the frontal and temporal electrode clusters in the lateral ANOVA (although it is not significant) and for the frontal electrode cluster in the medial ANOVA (significant). The main effect for condition as expected based on the increased negativity in the central and parietal regions might be weakened (and therefore not reach significance) by the increased positivity in the temporal and frontal regions.

Integration of Verbal and Nonverbal Information

The occurrence of the N400 when a word is presented that does not match the context of the picture storybook illustration, indicates that there is a relationship on the neurological level between verbal and nonverbal information processing. Whereas previous studies illustrated the importance of illustrations in story comprehension and bridging inferences (Brookshire, Scharff & Moses, 2002; Pike et al., 2010) and the advantages of a rich nonverbal environment within digital picture storybooks (Verhallen et al., 2006; Verhallen & Bus, 2010), the neurological processes underlying the powerful combination of pictures and text remained largely unknown. The results of the present pilot study provide a first insight in the integrated processing of verbal and complex nonverbal information at the neuropsychological level.

The typical N400 response observed as a reaction to incongruent picture-word combinations support the mechanisms described in Paivio's dual-coding theory (2007) and the concept of multisensory integration as proposed by Stein and colleagues (2010). Apparently the picture and the word are simultaneously processed at the neurological level and referential connections are formed to construct meaning (Paivio, 2007; Sipe, 1998).

Combining Verbal and Nonverbal Information Supports Comprehension

Based on the outcome of this and previous studies (e.g. Mayer & Sims, 1994; Schlag & Ploetzner, 2011), it can be concluded that nonverbal information can support the comprehension of verbal information. However, it is important to emphasize the following point when nonverbal information is used to facilitate the comprehension of verbal information.

The nonverbal information should match the verbal information. Whereas the participants in the present study were able to notice inconsistencies between the two types of information, the target words were not especially difficult for the adults in this study. This is reflected in the high scores on the behavioral measures in the pretest as well. Prior to the intervention, adults were able to make a correct semantic decision in the majority of the presented picture-word combinations (81% correct). It is unlikely that a high accuracy will also be the case if target words are used that are unknown to the participants. Since the support of nonverbal information is especially important when a text is difficult for a reader (Pike et al., 2010) and part of the verbal information will then be unknown, it is critical that the nonverbal information is congruent with the verbal information. Presenting congruent verbal and nonverbal information is likely to result in referential connections between the two types of information (Paivio, 2007; Stein et al., 2010) and thereby in enhanced comprehension. From this point of view a combination of non-matching verbal and nonverbal information might lead to incorrect referential connections and hamper vocabulary learning. However, more research regarding the inferences made from nonverbal information to comprehend unknown verbal information is required.

Limitations and Future Directions

The main aim of the present study was to gain more insight in the processes underlying information processing when a combination of verbal and nonverbal information is processed. Complex visual stimuli as used in the present study result in a large number of eye movement artifacts, however, this kind of stimuli are necessary for the ecological validity. The disadvantage is that it is not possible to obtain high-quality EEG-recordings when a person is actually processing complex nonverbal information simultaneously with verbal information. Therefore, the design requires a balance between the complex visual picture storybook stimuli and the recording of high-quality, relatively artifact free, EEG-data. In the present experiment this resulted in the sequential presentation of nonverbal and verbal information, a solution used in previous research as well (e.g. Ganis et al., 1996). Future studies should examine the possibilities to simultaneously present nonverbal and verbal information.

Another limitation of the present study is the sample size. Whereas robust evidence for the occurrence of the N400 was found, it was not possible to examine the intervention effectiveness with the small sample in the present study. It can be hypothesized that the intervention will result in a larger perceived discrepancy between congruent and incongruent picture-word combinations, which is in turn reflected by a larger N400 (Kutas & Hillyard, 1980). However, the current sample was too small to examine the subtle differences between the intervention and the control condition. Nevertheless, despite the small sample size, a N400 effect was found which shows the robustness of the occurrence of this effect. Future studies including larger samples and more trials per participant will provide more insight in the subtle differences that are expected to occur.

Conclusion

The present pilot study is the first study to provide insight in the neurological processes underlying information processing from picture storybooks. In line with Paivio’s dual-coding theory (2007) and the concept of multisensory integration from Stein and colleagues (2010) the present study suggests that verbal and nonverbal information processing is related and that relevant nonverbal information supports the comprehension of verbal information and vice versa. A well-established ERP peak (the N400) was found if an incongruent picture-word combination was presented, but not if a congruent picture-word combination was presented. Future studies might provide more insight into the expected subtle differences (e.g. between an intervention and a control condition). Furthermore, an experimental design might be used that is more ecologically valid while at the same time fulfilling the requirements to obtain high-quality EEG-data.

In sum the present pilot study shows that the use of EEG-recordings can provide important insights in the neurological processes underlying the combined processing of verbal and nonverbal information. Based on the results of this study it can be concluded that there is neuropsychological evidence for the forming of referential connections in the brain. This finding indicates that a combination of verbal and nonverbal information (as in picture storybooks), is not only a matter of beautification, but is important for the child’s cognitive development and specifically vocabulary learning as well.

References

Barret, S. E., & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition*, *14*, 201-212. doi: 10.1016/0278-2626(90)90029-N

BrainVision Analyzer (Version 2.0) [Computer software]. Gilching, Germany: Brain Products.

Brookshire, J., Scharff, L. F. V., & Moses, L. E. (2002). The influence of illustrations on children’s book preferences and comprehension. *Reading Psychology*, *23*, 323-339. doi: 10.1080/02702710290061391

Bus, A. G., Van IJzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, *65*, 1-21. doi: 10.2307/1170476

Caldera, R., Jermann, F., Lopez Aranga, G., & Van der Linden, M. (2004). Is the N400 category specific? A face and language processing study. *NeuroReport, 15*, 2589-2593. doi: 10.1097/00001756-200412030-00006

Dien, J., & Santuzzi, A. M. (2005). Application of repeated measures ANOVA to high-density ERP datasets: A review and tutorial. In T.C. Handy (Ed.), *Event-related potentionals: A methods handbook* (pp. 57-82). Cambridge, MA: The MIT Press.

E-prime (Version 2.0.8.22) [Computer software]. Pittsburgh, PA: Psychology Software Tools.

Fukuda, M., Niwa, S., Hiramatsu, K., Hata, A., Saitoh, O., Hayashida, S., ... Itoh, K. (1997). Behavioral and P3 amplitude enhancement in schizophrenia following feedback training. *Schizophrenia Research, 25,* 231-242. doi: 10.1016/S0920-9964(97)00028-5

Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for “common sense”: An electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience, 8*, 89-106. doi: 10.1162/jocn.1996.8.2.89

Hemphill, L., & Tivnan, T. (2008). The importance of early vocabulary for literacy achievement in high-poverty schools. *Journal of Education for Students Placed at Risk, 13*, 426-451. doi: 10.1080/10824660802427710

Huffmeijer, R., Tops, M., Alink, L. R. A., Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H. (2011). Love withdrawal is related to heigthened processing of faces with emotional expressions and incongruent emotional feedback: Evidence from ERPs. *Biological Psychology, 86*, 307-313. doi: 10.1016/j.biopsycho.2011.01.003

Huotilainen, M., Lovio, R., Kujala, T., Tommiska, V., Karma, K., & Fellman, V. (2011). Could audiovisual training be used to improve cognition in extremely low birth weight children? *Acta Paediatrica, 100*, 1489-1494. doi: 10.1111/j.1651-2227.2011.02345.x

Itagaki, S., Yabe, H., Mori, Y., Ishikawa, H., Takanashi, Y., & Nawi, S. (2011). Event-related potentials in patients with adult attention-deficit/hyperactivity disorder versus schizophrenia. *Psychiatry Research*, *189*, 288-291. doi: 10.1016/j.psychres.2011.03.005

Johnson, B. W., & Hamm, J. P. (2000). High-density mapping in an N400 paradigm: Evidence for bilateral temporal lobe generators. *Clinical Neurophysiology, 111*, 532-545. doi: 10.1016/S1388-2457(99)00270-9

Khateb, A., Pegna, A. J., Landis, T., Mouthon, M. S., & Annoni, J. M. (2010). On the origin of the N400 effects: An ERP waveform and source localization analysis in three matching tasks. *Brain Topography*, *23*, 311-320. doi: 10.1007/s10548-010-0149-7

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203-205. doi: 10.1126/science.7350657

Kutas, M., & Federmeier, K. D. (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. doi: 10.1146/annurev.psych.093008.131123

Liao, S., Su, Y., Wu, X., & Qiu, J. (2011). The Poggendorff illusion effect influenced by top-down control: Evidence from an event-related brain potential study. *Neuroreport*, *22*, 739-743. doi: 10.1097/WNR.0b013e32834ab40b

Lix, L. M., Keselman, J. C., & Keselman, H. J. (1996). Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance *F* test. *Review of Educational Research, 66*, 579-619. doi: 10.3102/00346543066004579

Luck, S. J. (2005). *An introduction to the event-related potential technique.* Cambridge, MA: The MIT Press.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Experimental Psychology, 86,* 389-401. doi**:** 10.1037/0022-0663.86.3.389

Mol, S. E., & Bus, A. G., & De Jong, T. M. (2009). Interactive book reading in early education: A tool to stimulate print knowledge as well as oral language. *Review of Educational Research, 79*, 979-1007. doi: 10.3102/0034654309332561

Naus, J. I. (1975). *Data quality control and editing.* New York: Marcel Dekker, Inc.

Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Pike, M. M., Barnes, M. A., & Barron, R. W. (2010). The role of illustrations in children's inferential comprehension. *Journal of Experimental Child Psychology, 105,* 243-255. doi: 10.1016/j.jecp.2009.10.006

Schlag, S., & Ploetzner, R. (2011). Supporting learning from illustrated texts: Conceptualizing and evaluating a learning strategy. *Instructional Science, 39*¸ 921-937. doi: 10.1007/s11251-010-9160-3

Silva-Pereyra, J., Conboy, B. T., Klarman, L., & Kuhl, P. K. (2007). Grammatical processing without semantics? An event-related brain potential study of preschoolers using jabberwocky sentences. *Journal of Cognitive Neuroscience, 19*, 1050-1065. doi: 10.1162/jocn.2007.19.6.1050

Singer, M., Andrusiak, P., Reisdorf, P., & Black, N. L. (1992). Individual differences in bridging inference processes. *Memory & Cognition, 20*, 539-548. doi: 10.3758/BF03199586

Sipe, L. R. (1998). How picture books work: A semiotically framed theory of text-picture relationships. *Children's Literature in Education, 29*, 97-108. doi: 10.1023/A:1022459009182

Stein, B. E., Burr, D., Constantinidis, C., Laurienti, P. J., Meredith, M. A., Perrault, T., … Lewkowicz, D. J. (2010). Semantic confusion regarding the development of multisensory integration: A practical solution. *European Journal of Neuroscience, 31*, 1713-1720. doi: 10.1111/j.1460-9568.2010.07206.x

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643-662. doi: 10.1037/h0054651

Van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11,* 657-671. doi: 10.1162/089892999563724

Verhallen, M. J. A. J., & Bus, A. G. (2010). Low-income immigrant pupils learning vocabulary through digital picture storybooks. *Journal of Educational Psychology, 102,* 54-61. doi: 10.1037/a0017133

Verhallen, M. J. A. J., Bus, A. G., & De Jong, M. T. (2006). The promise of multimedia stories for kindergarten children at risk. *Journal of Educational Psychology, 98,* 410-419. doi: 10.1037/0022-0663.98.2.410

Woltering, S., Granic, I., Lamm, C., & Lewis, M. D. (2011). Neural changes associated with treatment outcome in children with externalizing problems. *Biological Psychiatry, 70*, 873-879. doi: 10.1016/j.biopsych.2011.05.029

Yang, C. L., Perfetti, C. A., & Schmalhofer, F. (2007). Event-related potential indicators of text integration across sentence boundaries. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 55-89. doi: 10.1037/0278-7393.33.1.55

Appendix A: Picture storybooks

Intervention condition

Boonen, S., & Meijer, M. (2004). *Met opa op de fiets [With grandpa on the bike]*. Hasselt, The Netherlands: Clavis.

De Wijs, I., & Van den Hurk, N. (2001). *Rokko Krokodil [Rocco Crocodile]*. Arnhem, The Netherlands: Ziederis.

Van Genechten, G. (2009). *Kleine kangoeroe [Little Kangaroo]*. Hasselt, The Netherlands: Clavis.

Velthuijs, M. (2008). *De beer en het varkentje [The bear and the pig]*. Amsterdam: Leopold.

Control condition

Jutte, J. (2001). *Ruimtereis [Space journey]*. Amsterdam: Leopold.

Van Haeringen, A. (2004). *Beer is op vlinder [Bear has a crush on butterfly]*. Amsterdam: Leopold.

Veldkamp, T., & De Boer, K. (2008). *Na-apers [Apers]*. Arnhem, The Netherlands: Lannoo.

Velthuijs, M. (1998). *De trouwe dienaar [The loyal servant]*. Amsterdam: Leopold.

Appendix B: Target words selected from the 8 picture storybooks

'Beer is op vlinder' [Bear has a crush on butterfly]	
'Durft'	[Dares]
'Gebroken'	[Devastated]
'Geplukt'	[Picked]
'Hoogtevrees'	[Fear of heights]
'Kapot'	[Broken]
'Lied'	[Song]
'Verlegen'	[Shy]
'Wakkert'	[Stokes up]

'De beer en het varkentje' [The bear and the pig]	
'Afscheid'	[Goodbye]
'Dichtgevroren'	[Frozen]
'Graven'	[Dig]
'Luie'	[Lazy]
'Rivier'	[River]
'Sjouwde'	[Toted]
'Verdrink'	[Drown]
'Voorraad'	[Stock]

'Met opa op de fiets' [With grandpa on the bike]	
'IJsje'	[Ice cream]
'Kasteel'	[Castle]
'Koeien'	[Cows]
'Lek'	[Leak]
'Nek'	[Neck]
'Vuil'	[Dirt]
'Waslijn'	[Clothesline]
'Wei'	[Meadow]

'De trouwe dienaar' [The loyal servant]	
'Bericht'	[Message]
'Brievenbus'	[Mailbox]
'Haaien'	[Sharks]
'Oerwoud'	[Jungle]
'Roeien'	[Rowing]
'Schreef'	[Wrote]
'Taxi'	[Taxi]
'Vliegtuig'	[Airplane]

‘Ruimtereis’ [Space journey]	
‘Achterop’	[On the back]
‘Alleen’	[Alone]
‘Gedroomd’	[Dreamed]
‘Gegrom’	[Growling]
‘Golven’	[Waves]
‘Laat’	[Late]
‘Planeten’	[Planets]
‘Ruimte’	[Space]

‘Rokko krokodil’ [Rokko crocodile]	
‘Cadeautjes’	[Presents]
‘Klim’	[Climb]
‘Krokodillentranen-stoet’	[Crocodile tears procession]
‘Mist’	[Misses]
‘Ruim’	[Clean]
‘Steiger’	[Pier]
‘Trieste’	[Sad]
‘Verschijnen’	[Appear]

‘Na-apers’ [Apers]	
‘Buiten’	[Outside]
‘Buren’	[Neighbors]
‘Directeur’	[Manager]
‘Handen’	[Hands]
‘Na-aapten’	[Aped]
‘Nette’	[Neat]
‘Ouders’	[Parents]
‘Slingerden’	[Swinged]

‘Kleine kangoeroe’ [Little Kangaroo]	
‘Duwen’	[Push]
‘Grappig’	[Funny]
‘Leek’	[Seemed]
‘Ogen’	[Eyes]
‘Trots’	[Proud]
‘Uitgelaten’	[Exuberant]
‘Uitgeput’	[Exhausted]
‘Wasbeurt’	[Washed]