Published in final edited form as:

Dev Sci. 2015 March ; 18(2): 254–269. doi:10.1111/desc.12212.

The N400 and the Fourth Grade Shift

Donna Coch

Dartmouth College

Abstract

While behavioral and educational data characterize a fourth grade shift in reading development, neuroscience evidence is relatively lacking. We used the N400 component of the event-related potential waveform to investigate the development of single word processing across the upper elementary years, in comparison to adult readers. We presented third graders, fourth graders, fifth graders, and college students with a well-controlled list of real words, pseudowords, letter strings, false font strings, and animal name targets. Words and pseudowords elicited similar N400s across groups. False font strings elicited N400s similar to words and letter strings in the three groups of children, but not in college students. The pattern of findings suggests relatively adult-like semantic and phonological processing by third grade, but a long developmental time course, beyond fifth grade, for orthographic processing in this context. Thus, the amplitude of the N400 elicited by various word-like stimuli does not reflect some sort of shift or discontinuity in word processing around the fourth grade. However, the results do suggest different developmental time courses for the processes that contribute to automatic single word reading and the integrative N400.

According to Chall's (1983) developmental theory of reading, fourth grade marks a shift from "learning to read" to "reading to learn"; that is, a shift from focusing on lower-level reading skills such as grapheme-to-phoneme correspondences to focusing on higher-level reading skills such as comprehension. Chall and Jacobs (2003) suggested that automaticity and fluency, which they defined as quick and accurate word recognition, were key to the fourth grade shift. The other side of this shift, the fourth grade slump, has been discussed in the reading education literature for decades (e.g., Chall, Jacobs, & Baldwin, 1990; Stanovich, 1986). It characterizes students who have appeared to master basic phonological and orthographic decoding skills in the primary grades, but struggle with the shift that typifies the requirements of a fourth grade curriculum, particularly in terms of new word meanings (e.g., Chall & Jacobs, 2003). While behavioral and educational data are indicative of a fourth grade shift in reading, neuroscience evidence is relatively lacking. Here, we used the N400 component of the event-related potential (ERP) waveform as a proxy to investigate the development of automatic written word processing across the fourth grade.

Corresponding author: Donna Coch, Dartmouth College, Department of Education, Reading Brains Lab, 3 Maynard Street, Raven House, HB 6103, Hanover, NH, 03755. Telephone: 603.646.3282. FAX: 603.646.3968. donna.coch@dartmouth.edu..

Supplementary Materials

Grand average ERP waveforms elicited by words (black), pseudowords (red), letter strings (dark blue), and false fonts (light blue) for the four groups (college students, fifth graders, fourth graders, and third graders) across lateral and medial recording sites. Within the panel for each group, more anterior sites are toward the top, while more posterior sites are toward the bottom; left hemisphere sites are on the left and right hemisphere sites are on the right; lateral sites are toward the outer edges and medial sites are toward the middle of the panel; each vertical tick marks 100 ms; and negative is plotted up. The calibration bar marks 4.0 μ V.

Automaticity in Word Reading: Behavioral Evidence

Particularly in classrooms using a phonics approach, early stages of learning to read involve a focus on lower-level processes such as letter name knowledge and development of orthographic skills, phonemic awareness and development of phonological skills, and an evolving understanding of the relations between orthography and phonology (e.g., Adams, 1990; Barron, 1986; Stanovich, 1980; Treiman, 2000; Wagner & Torgesen, 1987). As decoding skills develop and integration across the orthographic and phonological domains becomes less effortful, word recognition becomes increasingly fast, obligatory, and autonomous, requiring only limited use of cognitive resources (e.g., Adams, 1990; Bakker, 1981; LaBerge & Samuels, 1974; Wolf, 1991). Theoretically, fluent readers have automatized not only a number of subskills contributing to word recognition, but also their integration (LaBerge & Samuels, 1974). Thus, across developmental time and reading experience, resources are freed to allow for meaningful analysis of words and integration of words and their contexts, in turn affording better comprehension (e.g., Adams, 1990; Stanovich, 1980). Behavioral findings indicate that this automaticity develops gradually throughout the early school years and reaches adult levels, at least for simple words, by the fourth grade (e.g., see Adams, 1990; Barron, 1981; Stanovich, 1980).

For example, there is evidence of a shift in, or increasingly flexible use of, the unit of processing around the fourth grade. While beginning readers tend to be letter-by-letter readers, fourth graders have begun to respond more quickly and accurately to syllables than to single letters (e.g., Friedrich, Schadler, & Juola, 1979). This is accompanied by a marked decrease in the word length effect in the fourth grade as compared to the primary grades, particularly with word center fixation (e.g., Aghababian & Nazir, 2000; Samuels, LaBerge, & Bremer, 1978). Others have also shown that effects of word length in letters, phonemes, and syllables decrease with increasing age from third grade to fifth grade to college; in a naming task, the greatest performance gain is between third and fifth grade (Bijeljac-Babic, Millogo, Farioli, & Grainger, 2004). This overall pattern of findings is consistent with a “developmental transition” from more serial (perhaps more effortful) to more parallel (perhaps more automatic) processing of printed words (Bijeljac-Babic et al., 2004, p. 425) around the fourth grade.

The concept of automaticity is closely related to the concept of word-likeness, as a key component of automaticity is specialized processing for lexical items. Developmentally, already during first grade, real words are responded to more quickly and accurately than pronounceable pseudowords when the stimuli are simple consonant-vowel-consonant strings (e.g., Doehring, 1976; Gibson, Osser, & Pick, 1963; Henderson & Chard, 1980; McCaughey, Juola, Schadler, & Ward, 1980). But it is not until the third or fourth grade that children are able to use phonology and spelling patterns to distinguish between pseudowords and nonsense letter strings in an adult-like way (e.g., Henderson & Chard, 1980; Lefton & Spragins, 1974; Rosinski & Wheeler, 1972). Thus, the concept of word-likeness develops across the primary grades, along with automaticity, as phonological and orthographic skills continue to grow, with some evidence of a shift around the fourth grade (Henderson & Chard, 1980, p. 101).

While these various behavioral findings are consistent with the notion of a shift in reading around the fourth grade, behavioral methods cannot provide information about on-line, real time processing during reading. Given that automaticity in lexical processing is theoretically a key feature of the fourth grade shift (e.g., Chall & Jacobs, 2003), it seems imperative to use methods, such as the recording of ERPs, which can directly index on-line processing in the study of the putative shift.

Automaticity in Word Reading: The N400 Component

Consistent with the behavioral evidence, from a neurocognitive perspective, reading words involves a sequence of overlapping and interactive processes; some of these processes can be indexed by ERPs (e.g., Grainger & Holcomb, 2009; Holcomb & Grainger, 2006). Generally, the N400 component of the ERP waveform (a negative-going deflection peaking at about 400 ms after stimulus onset) seems to serve as an index of word processing in terms of lexical access and/or post-access contextual integration processes, which are both in part dependent on lower-level processing (e.g., Khateb, Pegna, Landis, Mouthon, & Annono, 2010; Kutas & Federmeier, 2011; Kutas & Federmeier, 2000; Lau, Phillips, & Poeppel, 2008).

One possibility is that the N400 serves as an index of a higher-level integration process (e.g., Brown & Hagoort, 1993; Doyle, Rugg, & Wells, 1996; Holcomb, 1993) that builds, from the products of automatic lower-order processes, representations that provide the basis for comprehension (Holcomb, 1988; 1993, p. 60). These would seem to be the very processes implicated in the fourth grade shift. In this view, the N400 reflects the ease of linking across levels of representation, or “processing at a form-meaning interface that is sensitive to the compatibility of co-activated form (orthographic and phonological) and meaning representations” (e.g., Grainger & Holcomb, 2009; Holcomb, Reder, Misra, & Grainger, 2005, p. 170). Accordingly, it seems that the N400 could serve as an index of the integrative automaticity for word processing that is thought to characterize the fourth grade shift.

In addition, ERP studies with adults have shown that the N400 component is sensitive to lexicality or word-likeness: The N400 to pseudowords is similar to or larger than the N400 to real words, while nonpronounceable letter strings elicit markedly reduced or no N400 activity (e.g., Bentin, 1987; Bentin, McCarthy, & Wood, 1985; Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999; Deacon, Dynowska, Ritter, & Grose-Fifer, 2004; Holcomb, Grainger, & O'Rourke, 2002; Nobre & McCarthy, 1994; Rugg & Nagy, 1987). This pattern of findings – that only orthographically and phonologically legal letter strings elicit substantial N400 activity in adults – has led to the conclusion that N400 effects are “largely confined to items that are capable of giving rise to significant activation within the lexical system” (Rugg & Nagy, 1987, p. 479). This pattern suggests that it may be possible to use the N400 as a metric of what sorts of stimuli are processed as lexical items, providing an on-line measure of the automatic lexical processing that typifies fluent reading.

There is also some evidence from studies with adults that the N400 is sensitive to both phonology and orthography. For example, Newman and Connolly (2004) reported that pseudohomophones of expected words (e.g., *The ship disappeared in the thick phog [fog]*)

elicited a significantly reduced N400 in comparison to phonologically and semantically anomalous words. The authors interpreted this to indicate that “integrating word meaning with sentential context is influenced by the phonological representation of the presented letter string... the phonological representation of the nonword [pseudohomophone] is formed prior to and contributes to semantic integration” (Newman & Connolly, 2004, pp. 94, 97). The N400 has also been shown to be sensitive to lexical neighborhood size, an orthographic measure, for both words and pseudowords: Stimuli with many lexical neighbors elicit larger N400s than stimuli with smaller lexical neighborhoods in adults (Holcomb et al., 2002; Laszlo & Federmeier, 2011; Vergara-Martínez & Swaab, 2012). This is consistent with the hypothesis that the N400 in part could be an indirect index of the automatic lower-level orthographic and phonological form processing (e.g., Coch & Holcomb, 2003) theoretically critical to the fourth grade shift.

Developmentally, there is some evidence that the N400 could serve as such an index, although no extant study has directly addressed this issue. For example, in a study of visual sentence processing with 7- to 26-year-olds, Holcomb and colleagues (1992) reported an N400 to semantic anomalies (i.e., sentence-terminal words that did not make sense in context) in all age groups, which declined in amplitude with increasing age. Interestingly, both semantic anomalies and canonical terminal words elicited marked N400s (but the N400 to anomalies was larger) only in the youngest age groups (7-8 and 9-10). This pattern suggests that the N400 may be sensitive to the development of reading skills, with a shift around the fourth grade. Benau and colleagues (2011) have also reported that, in 10-year-old children, but not adults, moderately incongruent and strongly incongruent sentence terminal words elicit similar N400s.

Some findings from single word reading studies with children have also suggested developmental differences in the N400. For example, Dykman, Ackerman, Loizou, and Casey (2000) reported that 9- to 15-year-olds showed larger N400s to random letter strings than to words, inconsistent with findings from research with adults. Similarly, in a study with 10- and 11-year-olds comparing picture and word processing and using real words, pseudowords, nonpronounceable letter strings, and strings of letter-like symbols (false fonts), Coch, Maron, Wolf, and Holcomb (2002) found that all types of stimuli elicited an N400, although words and pseudowords elicited larger N400s than letter strings and false font strings. And in a word list study with first-grade girls, Coch and Holcomb (2003) reported that better readers who scored above the fourth grade level showed a substantial N400 to all stimulus types (known words, unknown words, difficult words, and letter strings), while poorer readers showed relatively little N400 activity for any type of stimulus. Overall, the pattern across these studies shows development in the sensitivity of the N400 from primary to upper elementary grades.

Taken together, these ERP findings suggest that children, as compared to adults, may be less efficient at distinguishing legal from illegal strings of letters, consistent with the behavioral evidence reviewed above. This could reflect a developing word processing system more open to linguistic possibility, such that all sorts of strings of letter-like symbols undergo some processing as potential lexical items. Strings may be processed as if they were words that children simply had not encountered before, reflecting a lack of automaticity in

processing for word-likeness (e.g., Smith & Dixon, 1971). As noted, this interpretation is consistent with the behavioral findings indicating that, with increasing age and reading experience, the “concept of word-likeness becomes more refined” (Henderson & Chard, 1980, p. 101). Theoretically, that refinement is tuned around the fourth grade (e.g., Chall, 1983). Here, we explored the possibility of the N400 as an index of that fourth grade shift.

The Present Study

Groups of undergraduate students, fifth graders, fourth graders, and third graders viewed real words, pseudowords, nonpronounceable letter strings, false font strings, and animal names in a word list paradigm as ERPs were recorded. A semantic categorization task required participants to respond to animal names. Similar sets of stimuli have been used in other neuroimaging studies (e.g., Petersen, Fox, Snyder, & Raichle, 1990) to index orthographic (letter strings compared to false fonts), phonological (pronounceable pseudowords compared to nonpronounceable letter strings), and semantic (words compared to pseudowords) processing.

Given the research reviewed above, we predicted that, in college students, real words and pseudowords would elicit marked N400s, letter strings would elicit a diminished N400, and false fonts would not elicit an N400. Whether there is a shift in processing more word-like stimuli (words and pseudowords) differently from less word-like stimuli (letter strings and false fonts) in terms of N400 amplitude across the third, fourth, and fifth grades was the primary research question. Given the scant developmental ERP data and the educational and behavioral evidence for a fourth grade shift reviewed above, we predicted that third graders would process all stimulus types more similarly; fifth graders would show a more differentiated, adult-like pattern of processing; and fourth graders would fall somewhere in between.

Methods

Participants

Participants included 24 undergraduate students (12 female, average age 20;3 years, *SD* 21 months), 24 fifth grade students (12 female, average 11;1, *SD* 4 months), 24 fourth grade students (12 female, average 10;0, *SD* 6 months), and 24 third grade students (13 female, average 8;9 years, *SD* 5 months). For children who participated in the summer months, group was determined by the grade completed the previous spring. All participants were right-handed (Oldfield, 1971), monolingual English speakers (who learned English as a first language and were not fluent in any other language) with no history of neurological dysfunction or language or reading disorders by self or parent report.¹ Also by self or parent report, none were currently taking any medications that would affect brain function. All had normal or corrected-to-normal binocular vision (20/40 or better) as tested with a standard Snellen chart. All participants were volunteers paid \$20 for their time; children also received

¹Parents of 4 third graders reported a delay in speech or reading, early motor issues, Title I services in first grade, and brief stutter around age 3; a parent of a fourth grader reported an early pronunciation issue and slow reading development; and parents of 4 fifth graders reported Title I services in first grade (1) and late talking (3). None of these children received a diagnosis or services under an IFSP or IEP, and none had any current language or reading issues; therefore, all were included in analyses.

small prizes. The study could be completed in one or two sessions (one-hour standardized testing battery, not reported on here due to space limitations, and ERP recording portions together or separately).

ERP Stimuli

The stimulus list consisted of 60 instances of each of five types of stimuli: real words, pseudowords, nonpronounceable letter strings, false font strings, and animal names (see Table 1). Real word stimuli (e.g., *bed*, *bring*) were the 20 most frequent 3-, 4-, and 5-letter words in grades 1 through 5 (Zeno, Ivens, Millard, & Duvvuri, 1995) that were regular and decodable, single syllable, open-class, and two to five phonemes. Pseudowords were matched for control to the real words (on number of letters and phonemes, orthographic neighborhood size, and constrained and unconstrained bigram and trigram counts, Medler & Binder, 2005, all p 's $> .41$ except constrained trigram count, $p = .093$), and were formed by changing one letter of each real word to make a decodable string with no homophones and a single preferred pronunciation (e.g., *bem*, *fring*). Letter strings (e.g., *mbe*, *nrfgi*) were matched for control to the pseudowords and were formed by randomly rearranging letters from the pseudowords into orthographically and phonologically illegal strings that were not common acronyms or typically pronounceable ($p < .001$ on all orthographic neighborhood, bigram, and trigram measures comparing pseudowords and letter strings; no significant difference in number of phonemes). False font stimuli were matched to the pseudowords as well and were composed of characters based on each letter in a pseudoword, systematically altered to retain the features of each letter, rearranged (e.g., BEM,,FRING); thus, physical variables such as spatial frequency and luminance remained constant across real letters and letter-like false fonts (see Grossi & Coch, 2005, for details). Animal names (e.g., *cat*, *duck*) served as targets that required a simple button-press response. Stimuli were presented singly, in list form, in pseudorandom order such that related stimuli were separated by at least four stimuli (e.g., *bed*, *bem*, *mbe*, and BEM were each separated by at least four intervening stimuli).

Procedure

Participants were given a brief overview of the study and tour of the lab and any questions were addressed before undergraduates signed a consent form or parents signed a permission form and children signed an assent form (which was read aloud to them). Participants were fitted with an electrode cap (Electro-Cap International, Eaton, Ohio) for electroencephalogram (EEG) recording. Active electrodes included Fz, Cz, Pz, FP1/2, F7/8, FT7/8, F3/4, FC5/6, C3/4, C5/6, T3/4, CT5/6, P3/4, T5/6, TO1/2, and O1/2; recordings from Fz, Pz, and Cz are not analyzed separately here, but data from these sites are included in the voltage maps. Mastoid electrodes served as reference; on-line recordings were referenced to the right mastoid and recordings were re-referenced to averaged mastoids in the final data averaging. Electrodes located below the right eye and at the outer canthi of the left and right eyes were used to identify blinks (in conjunction with recordings from FP1/2) and horizontal eye movements, respectively. Mastoid and scalp electrode impedances were maintained below 5 K Ω , and eye electrode impedances below 10 K Ω . Once electrode preparation was complete, participants were seated in a comfortable chair in a sound-attenuating and

electrically shielded booth for the ERP semantic categorization task. An experimenter was seated in the booth on a stool next to each child participant throughout ERP recording.

The stimuli were presented using Presentation software (Neurobehavioral Systems) at the center of a 19-inch LCD monitor approximately 66 inches in front of each participant, in white Times New Roman font on a black background. The stimuli subtended about 0.64° of vertical visual angle and 1.5° of horizontal visual angle, minimizing the need for scanning eye movements. The sequence of events began with the presentation of an outline of a white rectangle, within which a stimulus appeared 500 ms later; duration of the stimulus was 350 ms for college students and 500 ms for children. The white rectangle outline remained on the screen for 700 ms after the stimulus disappeared, followed by an asterisk for 2000 ms, followed by a blank screen for 500 ms, and the beginning of the next trial. Participants were instructed that they would see different kinds of words appear on the screen and to press a button on a game controller whenever they saw an animal name; response hand was counterbalanced across participants. They were further instructed to sit as still and relaxed as possible, to keep their eyes at the center of the screen, to try not to blink or move while the white rectangle was on the screen, and to save their blinks and wiggles for when the star (asterisk) was on the screen. A total of 300 stimuli were presented, with recording time of approximately 20 minutes. Breaks were given at each quarter, and more often as needed. A practice list with 10 items, not including any stimulus used in the actual experiment, was run prior to presentation of the experimental list.

A brief post-test followed the ERP recording session. Participants were given a sheet of paper with half the stimuli from the ERP task printed in five columns, and asked to circle the real words on the sheet as quickly and accurately as possible. Two post-test forms, with alternate halves of the stimulus list, were counterbalanced within and across groups.

Data Analysis

EEG was amplified with SA Instrumentation bioamplifiers (bandpass 0.01 to 100 Hz) and digitized on-line (sampling rate 4 ms). ERPs were time-locked to the onset of each stimulus. Off-line, separate ERPs to words, pseudowords, letter strings, false fonts, and animal names were averaged for each subject at each electrode site over a 1000 ms epoch, using a 200 ms pre-stimulus-onset baseline. Trials contaminated by eye movements, muscular activity, or electrical noise were not included in analyses. Standard artifact rejection parameters were initially employed, and data were subsequently analyzed on an individual basis for artifact rejection. The average number of trials included in individual ERP averages for each condition by group is summarized in Table 2. Pair-wise comparisons (Bonferroni-corrected $p = .0016$) indicated that the fourth grade group had fewer word trials than the college student group on average ($p = .001$), and that both the third ($p = .0001$) and fourth ($p = .0001$) grade groups had fewer animal name trials than the college student group on average, in part due to the sometimes vigorous button pressing of some of the younger children. ERP responses to the animal name targets were not included in further analyses, as they were characterized across groups by a target P300 (e.g., Polich, 2007) that affected N400 amplitude.

Mean amplitude of the N400 component was measured within the 300-550 ms time window; this epoch, consistent with previous studies, was determined by visual inspection of both individual and grand average waveforms. Analyses were first conducted within groups, followed by comparisons between groups using normalized and difference wave data. The within-groups ANOVAs focused only on the main effect of condition (real words, pseudowords, letter strings, false fonts). Planned follow-ups were conducted with Bonferroni adjustment for multiple comparisons.

For analyses between groups, to control for the typical overall larger waves in children, amplitude data were normalized based on the formula (score-mean/stdev), where score was an ERP average amplitude value (one for each condition and scalp site for each subject), mean was the mean amplitude across all subjects in each age group, and stdev was the standard deviation of the mean amplitude (see Coch, Grossi, Skendzel, & Neville, 2005; Holcomb et al., 1992). ANOVAs conducted with normalized data included the factors group (college students, fifth graders, fourth grades, third graders), condition, anterior/posterior [6 levels: frontal (F7/8, F3/4), fronto-temporal (FT7/8, FC5/6), temporal (T3/4, C5/6), central (CT5/6, C3/4), temporoparietal (T5/6, P3/4), and occipital (TO1/2, O1/2)], lateral/medial, and hemisphere (left, right). Significant findings involving both condition and group are reported below.

In addition, in order to compare the size of the differences in N400 amplitude between conditions between groups, difference waves were created from the non-normalized data; the rationale for these subtractions mirrored previous work (e.g., Petersen et al., 1990). Because words and pseudowords were matched on a number of orthographic and phonological variables here (number of letters, number of phonemes, orthographic neighborhood size, and constrained and unconstrained bigram and trigram counts), the main difference between these conditions was semantics (i.e., words had meaning while pseudowords did not); thus, subtracting the ERPs to pseudowords from the ERPs to words was considered to reflect primarily the effect of semantics. Because letter strings were matched for control with pseudowords, including containing the same letters, rearranged, the main difference between these conditions was pronounceability (i.e., pseudowords were pronounceable while letter strings were not); thus, subtracting the ERPs to letter strings from the ERPs to pseudowords was considered to reflect primarily the effect of phonology. Because false font strings, like letter strings, were matched for control with pseudowords, and were composed of characters based on each letter in the pseudowords (the same letters as in the letter strings), the main difference between letter strings and false font strings was orthography (i.e., letter strings were composed of real letters while false font strings were not); thus, subtracting the ERPs to false fonts from the ERPs to letter strings was considered to reflect primarily the effect of orthography. ANOVAs conducted with the difference wave data included the factors group, effect (semantics, phonology, orthography), anterior/posterior, lateral/medial, and hemisphere. Significant findings involving both effect and group are reported below. In order to visualize the distribution of these effects by group, topographical voltage maps based on the mean voltages measured in the difference waves at each electrode location within the 300-550 ms time window were created using a spherical spline interpolation (Perrin, Pernier, Bertrand, & Echallier, 1989).

The Greenhouse-Geisser correction was applied to all within-subjects measures with more than one degree of freedom. Results are significant at the .05 level unless otherwise noted.

Results

Post-Test Lexical Decision Task

Although accuracy on the post-test was high overall, the hit rate varied between groups, $F(3, 96) = 4.52, p < .01$: Average accuracy for third graders was 96.0% (SD 4.0%); for fourth graders, 97.6% (SD 3.4%); for fifth graders, 98.4% (SD 1.8%); and for college students, 98.8% (SD 1.3%). In follow-up paired comparisons (Bonferroni-corrected $p = .008$), only third graders were less accurate than college students, $t(46) = -3.24, p = .002$. The groups did not differ significantly in terms of false alarms (i.e., circling an item that was not a real word), $p = .25$.

Accuracy and Response Time on the ERP Semantic Categorization Task

Accuracy in pressing the button to animal names during the ERP semantic categorization task also varied by group, $F(3, 96) = 8.91, p < .001$, with average accuracy for third graders 89.9% (SD 7.7%); for fourth graders, 93.0% (SD 6.2%); for fifth graders, 94.1% (SD 4.4%); and for college students, 98.1% (SD 2.4%). In follow-up paired comparisons (Bonferroni-corrected $p = .008$), the third, $t(46) = -4.97, p = .001$, fourth, $t(46) = -3.77, p = .001$, and fifth, $t(46) = -3.89, p = .001$, grade groups were all less accurate than the college student group, but accuracy did not differ significantly amongst the child groups.

Response time also differed by group, $F(3, 96) = 20.20, p < .001$, with average time of 904.8 ms (SD 92.2) for third graders, 807.6 ms (SD 139.2) for fourth graders, 777.2 ms (SD 79.5) for fifth graders, and 670.5 ms (SD 100.4) for college students. In follow-up paired comparisons (Bonferroni-corrected $p = .008$), each of the child groups was significantly slower in responding than college students, third/college, $t(46) = 8.43, p = .001$, fourth/college, $t(46) = 3.92, p = .001$, fifth/college, $t(46) = 4.09, p = .001$. Third graders were also slower than both fourth, $t(46) = 2.85, p = .006$, and fifth, $t(46) = 5.14, p = .001$, graders.

ERPs within Groups: The Effect of Condition

The findings from these analyses are illustrated in Figures 1 and 2 and summarized in Table 3 (figures including all electrode sites measured are available in Supplementary Materials).

College Students—In college students, pseudowords elicited the most negative N400 (mean $-1.15 \mu V, SE .59$), followed by words (mean $-.50 \mu V, SE .48$), letter strings (mean $1.09 \mu V, SE .42$), and false fonts (mean $2.30 \mu V, SE .47$), $F(3, 69) = 51.25, p < .001$. As expected, planned comparisons indicated that words and pseudowords elicited similar-amplitude N400s in undergraduates ($p = .34$), while both words and pseudowords elicited more negative N400s than letter strings ($p = .001$ for both) and false fonts ($p = .001$ for both), and letter strings elicited a more negative N400 than false fonts ($p = .001$). Thus, the predicted pattern of more word-like stimuli eliciting significantly larger (more negative) N400s than less word-like stimuli was observed in college students.

Fifth Graders—In fifth graders, pseudowords elicited the most negative N400 (mean $-5.41 \mu\text{V}$, $SE .69$), followed by words (mean $-4.77 \mu\text{V}$, $SE .62$), false fonts (mean $-3.07 \mu\text{V}$, $SE .63$), and letter strings (mean $-2.54 \mu\text{V}$, $SE .63$), $F(3, 69) = 12.10$, $p < .001$. Planned comparisons indicated that words and pseudowords ($p = .56$) and words and false fonts ($p = .11$) elicited similar-amplitude N400s, while words elicited a more negative N400 than letter strings ($p = .004$). Pseudowords also elicited a more negative N400 than both letter strings ($p = .001$) and false fonts ($p = .004$). Letter strings and false fonts elicited similar-amplitude N400s ($p = 1$). Thus, the pattern of more word-like stimuli eliciting larger (more negative) N400s than less word-like stimuli was not observed in fifth graders, as false fonts and words elicited similar-amplitude N400s; this was not the pattern observed in college students here.

Fourth Graders—In fourth graders, pseudowords elicited the most negative N400 (mean $-5.89 \mu\text{V}$, $SE .75$), followed by words (mean $-5.02 \mu\text{V}$, $SE .73$), false fonts (mean $-3.85 \mu\text{V}$, $SE .84$), and letter strings (mean $-3.75 \mu\text{V}$, $SE .82$), $F(3, 69) = 8.03$, $p < .01$. Planned comparisons indicated that the amplitude of the N400 to words was not significantly different than the amplitude of the N400 to pseudowords ($p = .15$), letter strings ($p = .12$), or false fonts ($p = .21$). The N400 to pseudowords was more negative than the N400 to letter strings ($p = .01$) and false fonts ($p = .02$). Letter strings and false fonts elicited similar-amplitude N400s ($p = 1$). Thus, the pattern of more word-like stimuli eliciting larger (more negative) N400s than less word-like stimuli was not observed in fourth graders, as the overall amplitude of the N400 to words was not significantly different from the overall amplitude of the N400 to letter strings or false fonts. This was not the pattern observed in either college students or fifth graders here.

Third Graders—In third graders, pseudowords elicited the most negative N400 (mean $-5.77 \mu\text{V}$, $SE .90$), followed by words (mean $-4.78 \mu\text{V}$, $SE 1.02$), false fonts (mean $-3.86 \mu\text{V}$, $SE .70$), and letter strings (mean $-3.09 \mu\text{V}$, $SE .79$), $F(3, 69) = 8.78$, $p < .001$. Planned comparisons indicated that words elicited an N400 of similar amplitude to pseudowords ($p = .58$) and false fonts ($p = 1$), but a more negative N400 than letter strings ($p = .05$). Pseudowords elicited a more negative N400 than both letter strings ($p = .001$) and false fonts ($p = .01$). Letter strings and false fonts elicited similar-amplitude N400s ($p = .92$). Thus, the pattern of more word-like stimuli eliciting larger (more negative) N400s than less word-like stimuli was not observed in third graders, as false fonts and words elicited similar-amplitude N400s. This was not the pattern observed in college students or fourth graders here, but was similar to the pattern observed in fifth graders.

ERPs between Groups: Normalized Data

Between-groups analyses allowed for a direct comparison of the condition effects found within each group. A between-groups analysis with the raw ERP data yielded a main effect of group, $F(1, 92) = 13.85$, $p < .001$, likely in part reflecting the overall larger waveforms in the child groups. Therefore, a between-groups analysis with normalized ERP data was conducted to control for the typical larger waves in children. The analysis of the normalized data confirmed the impression from the within-group results that the various stimulus types were processed differently by the different groups, condition $\times$ group, $F(9, 276) = 5.26$, $p < .$

001, condition $\times$ anterior/posterior $\times$ group, $F(45, 1380) = 1.75, p < .05$, condition $\times$ anterior/posterior $\times$ lateral/medial $\times$ group, $F(45, 1380) = 2.39, p < .001$. Follow-up analyses (Bonferroni-corrected $p = .0125$) were conducted in order to investigate group differences in N400 by condition.

Real words elicited similar N400s across the four groups; for the word condition, there were no significant effects involving group.

For pseudowords, the distribution of the N400 differed by group, anterior/posterior $\times$ lateral/medial $\times$ group, $F(15, 460) = 2.40, p = .008$. In the three groups of children, pseudowords elicited an N400 that was widespread across the scalp but most negative at central, lateral sites. In contrast, the N400 to pseudowords in college students was widespread across the scalp anterior to the occipital sites, and maximal at frontal, medial sites.

The distribution of the N400 to letter strings also differed by group, anterior/posterior $\times$ lateral/medial $\times$ group, $F(15, 460) = 2.57, p = .004$. The N400 elicited by letter strings was marked and widespread across the scalp in all three groups of children; however, it was most negative at central, lateral sites in third graders and central, medial sites in fourth and fifth graders. In contrast, in college students, an N400 to letter strings was observed at all sites anterior to the occipital sites, with increasing amplitude moving anteriorly and the most negative N400s at frontal, medial sites.

The distribution of the N400 to false fonts differed by group, as well, anterior/posterior $\times$ lateral/medial $\times$ group, $F(15, 460) = 4.33, p = .001$. In each of the child groups, a marked N400 to false fonts was widespread anterior to the occipital sites, especially at medial sites. In contrast, an N400 to false fonts was absent at posterior sites, but evident at more anterior sites, in college students.

ERPs between Groups: Difference Waves

Difference waves were created to compare the effects of semantics (words - pseudowords), phonology (pseudowords - letter strings), and orthography (letter strings - false fonts) across groups (see Figure 3). Analyses of the difference waves indicated that the distribution of these effects differed by group, effect $\times$ group $\times$ hemisphere, $F(6, 184) = 2.68, p < .05$, effect $\times$ group $\times$ anterior/posterior, $F(30, 920) = 2.71, p < .01$. Follow-up analyses by effect (Bonferroni-corrected $p = .0167$) were conducted in order to compare each effect across the groups.

The difference between the N400 elicited by words and the N400 elicited by pseudowords was similar across groups; there were no significant interactions involving group for the semantic effect.

The difference between the N400 to pseudowords and the N400 to letter strings was differently distributed across groups, group $\times$ hemisphere, $F(3, 92) = 4.01, p = .010$. The phonological effect was more widespread over the left hemisphere in third graders and over the right hemisphere in fourth graders, but had a more bilateral distribution in fifth graders and college students.

Overall, the difference in N400 amplitude between letter strings and false fonts (the orthographic effect) was greatest in college students (mean $-1.22 \mu\text{V}$, $SE .26$), followed by third graders (mean $.76 \mu\text{V}$, $SE .52$), fifth graders (mean $.52 \mu\text{V}$, $SE .53$), and fourth graders (mean $.10 \mu\text{V}$, $SE .38$), $F(3, 92) = 4.10$, $p = .009$. Follow-up comparisons (Bonferroni-corrected $p = .008$) indicated that the orthographic effect had a similar average size among the child groups, while the third ($p = .001$), fourth ($p = .006$), and fifth ($p = .005$) grade groups all showed a smaller difference between N400 amplitude to letter strings and false fonts than the college student group. The orthographic effect was localized to occipital and lateral temporoparietal sites in the three child groups, and distributed less posteriorly and more laterally in the college students, anterior/posterior $\times$ group, $F(15, 460) = 6.40$, $p = .001$, lateral/medial $\times$ group, $F(3, 92) = 6.21$, $p = .001$, anterior/posterior $\times$ lateral/medial $\times$ group, $F(15, 460) = 3.72$, $p = .001$.

Discussion

We presented third graders, fourth graders, fifth graders, and college students with a well-controlled list of real words, pseudowords, letter strings, false font strings, and animal name targets in order to investigate the putative fourth grade shift in terms of the development of the N400 across the upper elementary years, in comparison to adult readers. The ERP data provided little support for the notion of a fourth grade shift. However, the findings did suggest different developmental time courses for the orthographic, phonological, and semantic processes central to reading and indexed by the N400.

As expected, both words and pseudowords elicited marked N400s in each of the groups. The higher-level integration processes theoretically indexed by the N400 are applied with pseudowords because early orthographic and phonological processing can be completed, and semantic processes are likely partially activated (e.g., through partial activation of similar real word representations, Holcomb et al., 2002). Thus, because both words and pseudowords are orthographically and phonologically legal and are likely to activate, at least partially, lexical representations, both were expected to elicit N400s in college students (e.g., Rugg & Nagy, 1987). Analyses of normalized data indicated no differences among groups in the amplitude of the N400 elicited by words, suggesting similar higher-level integration processing for simple, single words across groups; this is consistent with behavioral data indicating specialized word processing (for very simple words) already during first grade (e.g., Doehring, 1976; Gibson et al., 1963; Henderson & Chard, 1980; McCaughey et al., 1980). In contrast, the N400 elicited by pseudowords had a more central, lateral distribution in the child groups and a more frontal, medial distribution in college students. Speculatively, this implies a difference between children and college students in the networks involved in partial lexical activation; this may be a reflection of comparatively sparse lexical networks in the children. However, despite this distributional difference, the effect of semantics on the N400 (words - pseudowords) was not significantly different across groups.

Given the similar N400s to words and pseudowords across groups, there was little ERP evidence of a fourth grade shift related to word meaning (cf. Chall & Jacobs, 2003). However, the real words used here were simple, high frequency, single syllable words (not the more difficult words that characterize a fourth grade curriculum), and the pseudowords

were matched with the real words. The results suggest that the integrative processing indexed by the N400 for such words and pseudowords is relatively adult-like by the third grade (e.g., Eddy, Grainger, Holcomb, Mitra, & Gabrieli, 2014), even earlier than behavioral data suggesting automatic processing for simple, high frequency single words by the fourth grade (e.g., Adams, 1990; Barron, 1981).

Turning to the letter string condition, as expected, letter strings overall elicited a significantly smaller N400 than words or pseudowords in college students in within-group analyses (e.g., Bentin et al., 1985; Nobre & McCarthy, 1994, 1995). This pattern confirms the sensitivity of N400 amplitude to more and less word-like stimuli. In the case of letter strings, automatic lower-level phonological and orthographic pattern processing might be attempted, but cannot be completed (the strings by definition are unpronounceable and contain unfamiliar sequences of letters), and little semantic processing is activated, precluding the integration of the products of these processes and reducing the amplitude of the N400.

In within-group analyses of the condition effect in children, letter strings overall also elicited smaller N400s than words in third and fifth graders, while N400 amplitude did not distinguish between words and letter strings overall in fourth graders; given the similarity between the third and fifth grade groups, this anomaly does not seem to be an indication of some sort of shift in the fourth grade. The findings from third and fifth graders are consistent with a previous finding of marked N400s, but smaller than the N400s elicited by words, elicited by letter strings in 10- and 11-year-olds (Coch et al., 2002). A similar effect for letter strings as compared to pseudowords was also reported (Coch et al., 2002); this was also observed here in the within-group analyses: Pseudowords elicited a more negative N400 than letter strings overall, across groups. This is consistent with behavioral findings indicating differentiation of pseudowords and letter strings by the third or fourth grade (e.g., Lefton & Spragins, 1974; Rosinski & Wheeler, 1972), but the ERP evidence of marked N400s suggests that letter strings still afford some processing within the lexical system (e.g., Laszlo & Federmeier, 2011).

Nonetheless, the findings regarding letter string processing are relatively consistent in indicating that these less word-like stimuli are differentiated by N400 amplitude from more word-like stimuli in each of the groups. However, there was evidence of developmental differences: Analyses of normalized data showed a more central distribution for the N400 elicited by letter strings in the child groups, and a more frontal, medial distribution in college students. In combination with the similar distributional findings from the pseudoword analyses, this pattern suggests that the generators of the N400 (e.g., see Lau et al., 2008; Nobre, Allison, & McCarthy, 1994) may be somewhat differently oriented in children than in adults. The difference wave analysis of the phonological effect within the N400 time window (pseudowords - letter strings) provided confirmatory evidence for developmental distributional differences: The effect was more widespread over the left hemisphere in third graders, over the right hemisphere in fourth graders, and was more bilaterally distributed in fifth graders and college students.

Despite these apparent differences, the overall pattern of findings indicates that N400 amplitude is sensitive to the differences between letter strings and pseudowords (i.e., primarily phonological information) across age groups. For each group, the lack of phonological information in letter strings as compared to pseudowords reduced N400 amplitude. This is consistent with previous reports of sensitivity of the N400 to phonological information from print in adults (e.g., Newman & Connolly, 2004), and in children as young as age 6 (e.g., Coch, Mitra, George, & Berger, 2011; Grossi, Coch, Coffey-Corina, Holcomb, & Neville, 2001). It is also consistent with the view that the N400 reflects higher-level integration across various levels of representations (e.g., Grainger & Holcomb, 2009; Holcomb et al., 2005), and thus can provide an indirect index of the automaticity of the contributing lower-level processes (e.g., Coch & Holcomb, 2003). Our data suggest that automatic phonological processing (or the reduction in N400 amplitude due to the lack thereof in processing letter strings) is relatively adult-like by the third grade, providing little evidence of some sort of specific shift in processing letter strings across the fourth grade.

Taken together, the ERP findings regarding N400 processing of words, pseudowords, and letter strings do not support the notion of a marked fourth grade shift in word processing, as each of the child groups showed N400 patterns similar to the college student group. The data from the false font condition also do not support the notion of a marked fourth grade shift, but for a different reason: All of the child groups in part showed a different pattern than the college student group. In within-group analyses of condition, false fonts overall elicited smaller N400s than pseudowords in each of the groups, consistent with both previous findings in 10- and 11-year-olds with a different set of false font stimuli (Coch et al., 2002) and the hypothesis that N400 amplitude distinguishes more from less word-like stimuli. However, false fonts elicited smaller N400s than words and letter strings only in college students; in third, fourth, and fifth graders, overall N400 amplitude did not distinguish false fonts from words or letter strings.

These were unexpected findings suggesting that children across groups processed false fonts, strings of letter-like characters, both as word-like (in comparison to words) and as not-word-like (in comparison to letter strings). Does this mean that the children thought that the false font strings were words? Both the evident un-word-likeness of the false font stimuli and anecdotal comments from the children (e.g., “What were those weird ones?”) belie this, as do the post-test data indicating no differences in false alarms across groups (i.e., children did not circle false font items as real words on the post-test). Another possibility is that, given greater sensitivity to context in children as compared to adults (e.g., Stanovich, 1980), the local context of the list including real words, animal names, pseudowords, and letter strings influenced children to process false font strings as possible strings of letters, perhaps in an unfamiliar font (e.g., see Coch et al., 2002). Still another, potentially related, possibility is that children processed false font strings as some sort of meaningful objects, while adults, with years more experience reading and fine-tuning word processing, did not attempt to find meaning in false font strings (e.g., see neuronal recycling, Dehaene & Cohen, 2007).

Direct comparisons between groups confirmed that children processed false font strings differently than college students in terms of the N400. In analyses with normalized data,

false fonts elicited a marked, widespread N400 at sites anterior to the occipital sites in the child groups, while college students showed an N400 to false fonts only at more anterior sites. That there was an N400 to false font strings at all in college students may be a reflection of the nature of these specific false font stimuli, with re-arranged pieces of letters allowing for some basic visual familiarity. The difference wave analyses reflected both the amplitude and distributional differences between groups. The orthographic effect (i.e., the difference in N400 amplitude to letter strings and false font strings) was significantly larger in college students than in children. Further, the orthographic effect was localized primarily to the most posterior, occipital sites in the child groups, but was more centrally distributed (reflecting the relative absence of an N400 to false fonts and presence of an N400 to letter strings) in college students. Thus, of the ERP measures analyzed here, it was the orthographic control false fonts that distinguished each of the child groups from the college student group - but not from each other. Again, the developmental differences did not provide evidence for some sort of shift around the fourth grade; indeed, the evidence suggests a shift in false font processing to become more adult-like sometime after the fifth grade.

Based on a view of the N400 as an index of interactive word processing at multiple, cascaded levels of representation (e.g., Coch & Holcomb, 2003; Kutas & Federmeier, 2011; Laszlo & Federmeier, 2011, p. 185), the false font findings suggest a lack of automaticity in contributing lower-level orthographic processing in the child groups. Semantic analysis may be attempted for all sorts of orthographic inputs (Laszlo & Federmeier, 2011, p. 177), but the automaticity of this process fails with strings of letter-like characters because the contributing lower-level orthographic processing cannot be automatic for these unfamiliar characters. In fluently reading college students, this is reflected in the minimal N400 to false fonts – smaller overall than the N400 to letter strings, consistent with the suggestion of failure of not only contributing automatic lower-level phonological processing, but also orthographic processing. Taken together, these findings are compatible with the speculation that the marked N400 to false fonts in the child groups could be an indication of non-automatic, effortful orthographic processing. This would, in turn, be consistent with other ERP findings showing a relatively lengthy time course for the development of automaticity in orthographic processing, extending beyond age 11 into adolescence (e.g., Brem et al., 2006; Coch, Mitra, & George, 2012; Eddy et al., 2014).

In contrast to the ERP findings, the behavioral response time and post-test accuracy findings might be interpreted as consistent with a shift at the fourth grade, as previous behavioral data have suggested (e.g., Adams, 1990; Barron, 1981; Stanovich, 1980). In terms of response times on the semantic categorization ERP task, all children were slower than college students, but, within the child groups, only the third graders were slower than the fourth and fifth graders. This suggests some sort of shift at the fourth grade. However, given that this measure likely involves factors related to basic motor development, it is difficult to conclude that this shift is specific to word processing; this highlights the importance of using both behavioral and neural measures. On the post-test, only the third graders were less accurate than the college students, indicating a shift towards adult-like response accuracy at the fourth grade. In contrast, all of the child groups were similarly accurate on the semantic categorization task, but less accurate than college students, inconsistent with a fourth grade

shift. Thus, while some of the behavioral measures might suggest a fourth grade shift, these data overall are equivocal.

In conclusion, previous behavioral and educational data have suggested a fourth grade shift related to increasing automaticity and fluency in word recognition (e.g., Adams, 1990; Chall & Jacobs, 2003). Higher-level, integrative word processing is in part dependent on the automaticity of contributing lower-level processes, such as orthographic and phonological processes (e.g., LaBerge & Samuels, 1974). The N400 may index such processing at the interface between lower-level and higher-level processes (e.g., Grainger & Holcomb, 2009; Holcomb et al., 2005). Here, for the processing of simple, single words, pseudowords, letter strings, and false font strings, we found little electrophysiological evidence to support the notion of a fourth grade shift. Words and pseudowords elicited similar-amplitude N400s within each group, suggesting relatively adult-like semantic processing in this context by the third grade. Pseudowords elicited larger N400s than letter strings within each age group, suggesting relatively adult-like phonological processing by the third grade. And while letter strings elicited larger N400s than false font strings in college students, letter strings and false fonts elicited similar-amplitude N400s overall in each of the child groups, suggesting development in automatic orthographic processing beyond the fifth grade.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Grateful thanks are due to the children, parents, and college students who elected to participate in this study, and to the local schools, libraries, businesses, and programs that allowed us to share information about the study with families and children. The project could not have been completed without help from undergraduate research assistants Jennifer Bares, Clarisse Benoit, Natalie Berger, Sarah Brim, Ayesha Dholakia, Elyse George, Emily Jasinski, Gabriela Meade, Priya Mitra, Erin Rokey, and Anna Roth on various aspects of paradigm preparation, scheduling, data collection, and initial data analysis, and help with programming from Ray Vukcevic and Mark Dow. Grant R03HD058613 from the National Institutes of Health, Eunice Kennedy Shriver National Institute of Child Health and Human Development, supported this research.

References

- Adams, MJ. Beginning to read: thinking and learning about print. MIT Press; Cambridge: 1990.
- Aghababian V, Nazir TA. Developing normal reading skills: aspects of the visual processes underlying word recognition. *Journal of Experimental Child Psychology*. 2000; 76:123–150. doi: 10.1006/jecp.1999.2540. [PubMed: 10788306]
- Bakker, DJ. A set of brains for learning to read.. In: Diller, KA., editor. Individual differences and universals in language learning aptitude. Newbury House; Rowley: 1981. p. 65-71.
- Barron, RW. Development of visual word recognition: A review.. In: MacKinnon, GE.; Waller, TG., editors. Reading research: Advances in theory and practice. Vol. 3. Academic Press; New York: 1981. p. 119-158.
- Barron RW. Word recognition in early reading: A review of the direct and indirect hypotheses. *Cognition*. 1986; 24:93–119. doi: 10.1016/0010-0277(86)90006-5. [PubMed: 3539489]
- Benau EM, Morris J, Couperus JW. Semantic processing in children and adults: incongruity and the N400. *Journal of Psycholinguistic Research*. 2011; 40:225–239. doi: 10.1007/s10936-011-9167-1. [PubMed: 21479571]

- Bentin S. Event-related potentials, semantic processes, and expectancy factors in word recognition. *Brain and Language*. 1987; 31:308–327. doi: 10.1016/0093-934X(87)90077-0. [PubMed: 3620905]
- Bentin S, McCarthy G, Wood CC. Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology*. 1985; 60:343–355. doi: 10.1016/0013-4694(85)90008-2. [PubMed: 2579801]
- Bentin S, Mouchetant-Rostaing Y, Giard MH, Echallier JF, Pernier J. ERP manifestations of processing printed words at different psycholinguistic levels: time course and scalp distribution. *Journal of Cognitive Neuroscience*. 1999; 11(3):235–260. doi: 10.1162/089892999563373. [PubMed: 10402254]
- Bijeljac-Babic R, Millogo V, Farioli F, Grainger J. A developmental investigation of word length effects in reading using a new on-line word identification paradigm. *Reading and Writing*. 2004; 17:411–431. doi: 10.1023/B:READ.0000032664.20755.af.
- Brem S, Bucher K, Halder P, Summers P, Dietrich T, Martin E, Brandeis D. Evidence for developmental changes in the visual word processing network beyond adolescence. *Neuroimage*. 2006; 29:822–837. doi: 10.1016/j.neuroimage.2005.09.023. [PubMed: 16257546]
- Brown E, Hagoort P. The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*. 1993; 5:34–44. doi: 10.1162/jocn.1993.5.1.34. [PubMed: 23972118]
- Chall, JS. Stages of reading development. McGraw-Hill; New York: 1983.
- Chall JS, Jacobs VA. Poor children's fourth-grade slump. *American Educator*. 2003; 27(1):14–15. 44.
- Chall, JS.; Jacobs, VA.; Baldwin, LE. The reading crisis: Why poor children fall behind. Harvard University Press; Cambridge, MA: 1990.
- Coch D, Grossi G, Skendzel W, Neville H. ERP nonword rhyming effects in children and adults. *Journal of Cognitive Neuroscience*. 2005; 17(1):168–182. doi: 10.1162/0898929052880020. [PubMed: 15701247]
- Coch D, Holcomb PJ. The N400 in beginning readers. *Developmental Psychobiology*. 2003; 43(2): 146–166. doi: 10.1002/dev.10129. [PubMed: 12918093]
- Coch D, Maron L, Wolf M, Holcomb PJ. Word- and picture-processing in children: An event-related potential study. *Developmental Neuropsychology*. 2002; 22(1):373–406. doi: 10.1207/S15326942dn2201_3. [PubMed: 12405510]
- Coch D, Mitra P, George E. Behavioral and ERP evidence of word and pseudoword superiority effects in 7- and 11-year-olds. *Brain Research*. 2012; 1486:68–81. doi: 10.1016/j.brainres.2012.09.041. [PubMed: 23036274]
- Coch D, Mitra P, George E, Berger N. Letters rhyme: electrophysiological evidence from children and adults. *Developmental Neuropsychology*. 2011; 36(3):302–318. doi: 10.1080/87565641.2010.549985. [PubMed: 21462009]
- Deacon D, Dynowska A, Ritter W, Grose-Fifer J. Repetition and semantic priming of nonwords: implications for theories of N400 and word recognition. *Psychophysiology*. 2004; 41:60–74. doi: 10.1111/1469-8986.00120. [PubMed: 14693001]
- Dehaene S, Cohen L. Cultural recycling of cortical maps. *Neuron*. 2007; 56:384–398. doi: 10.1016/j.neuron.2007.10.004. [PubMed: 17964253]
- Doehring DG. Acquisition of rapid reading responses. *Monographs of the Society for Research in Child Development*. 1976; 41(2) Serial No. 165 doi: 10.2307/1166001.
- Doyle MC, Rugg MD, Wells T. A comparison of the electrophysiological effects of formal and repetition priming. *Psychophysiology*. 1996; 33:132–147. doi: 10.1111/j.1469-8986.1996.tb02117.x. [PubMed: 8851241]
- Dykman RA, Ackerman P, Loizou PC, Casey PH. An event-related potential study of older children with an early history of failure to thrive. *Developmental Neuropsychology*. 2000; 18(2):187–212. doi: 10.1207/S15326942DN1802_3. [PubMed: 11280964]
- Eddy MD, Grainger J, Holcomb PJ, Mitra P, Gabrieli JDE. Masked priming and ERPs dissociate maturation of orthographic and semantic components of visual word recognition in children. *Psychophysiology*. 2014; 51(2):136–141. doi: 10.1111/psyp.12164. [PubMed: 24313638]
- Friedrich FJ, Schadler M, Juola JF. Developmental changes in units of processing in reading. *Journal of Experimental Child Psychology*. 1979; 28:344–358. doi: 10.1016/0022-0965(79)90094-8.

- Gibson EJ, Osser H, Pick AD. A study of the development of grapheme-phoneme correspondences. *Journal of Verbal Learning and Verbal Behavior*. 1963; 2:142–146. doi: 10.1016/S0022-5371(63)80078-X.
- Grainger J, Holcomb PJ. Watching the word go by: on the time-course of component processes in visual word recognition. *Language and Linguistics Compass*. 2009; 3(1):128–156. doi: 10.1111/j.1749-818x.2008.00121.x. [PubMed: 19750025]
- Grossi G, Coch D. Automatic word form processing in masked priming: an ERP study. *Psychophysiology*. 2005; 42:343–355. doi: 10.1111/j.1469-8986.2005.00286.x. [PubMed: 15943688]
- Grossi G, Coch D, Coffey-Corina S, Holcomb PJ, Neville HJ. Phonological processing in visual rhyming: A developmental ERP study. *Journal of Cognitive Neuroscience*. 2001; 13(5):610–625. doi: 10.1162/089892901750363190. [PubMed: 11506660]
- Henderson, L.; Chard, J. The reader's implicit knowledge of orthographic structure.. In: Frith, U., editor. *Cognitive processes in spelling*. Academic Press; NY: 1980. p. 85-116.
- Holcomb PJ. Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language*. 1988; 35(1):66–85. doi: 10.1016/0093-934X(88)90101-0. [PubMed: 3179703]
- Holcomb PJ. Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology*. 1993; 30:47–61. doi: 10.1111/j.1469-8986.1993.tb03204.x. [PubMed: 8416062]
- Holcomb PJ, Coffey SA, Neville HJ. Visual and auditory sentence processing: A developmental analysis using event-related brain potentials. *Developmental Neuropsychology*. 1992; 8:203–241. doi: 10.1080/87565649209540525.
- Holcomb PJ, Grainger J. On the time course of visual word recognition: an event-related potential investigation using masked repetition priming. *Journal of Cognitive Neuroscience*. 2006; 18(10):1631–1643. doi: 10.1162/jocn.2006.18.10.1631. [PubMed: 17014368]
- Holcomb PJ, Grainger J, O'Rourke T. An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *Journal of Cognitive Neuroscience*. 2002; 14(6):938–950. doi: 10.1162/089892902760191153. [PubMed: 12191460]
- Holcomb PJ, Reder L, Misra M, Grainger J. The effects of prime visibility on ERP measures of masked priming. *Cognitive Brain Research*. 2005; 24:155–172. doi: 10.1016/j.cogbrainres.2005.01.003. [PubMed: 15922167]
- Khateb A, Pegna AJ, Landis T, Mouthon MS, Annono J-M. On the origin on the N400 effects: an ERP waveform and source localization analysis in three matching tasks. *Brain Topography*. 2010; 23:311–320. doi: 10.1007/s10548-010-0149-7. [PubMed: 20549553]
- Kucera, H.; Francis, WN. *Computational analysis of present-day American English*. Brown University Press; Providence, RI: 1967.
- Kutas M, Federmeier K. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*. 2011; 62:621–647. doi: 10.1146/annurev.psych.093008.131123.
- Kutas M, Federmeier KD. Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*. 2000; 4(12):463–470. doi: 10.1016/S1364-6613(00)01560-6. [PubMed: 11115760]
- LaBerge D, Samuels SJ. Toward a theory of automatic information processing in reading. *Cognitive Psychology*. 1974; 6:293–323. doi: 10.1016/0010-0285(74)90015-2.
- Laszlo S, Federmeier K. The N400 as a snapshot of interactive processing: evidence from regression analyses of orthographic neighbor and lexical associate effects. *Psychophysiology*. 2011; 48:176–186. doi: 10.1111/j.1469-8986.2010.01058.x.
- Lau EF, Phillips C, Poeppel D. A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*. 2008; 9:920–933. doi: 10.1038/nrn2532.
- Lefton LA, Spragins AB. Orthographic structure and reading experience affect the transfer from iconic to short-term memory. *Journal of Experimental Psychology*. 1974; 103:775–781. doi: 10.1037/h0037200.

- McCaughey MW, Juola JF, Schadler M, Ward NJ. Whole-word units are used before orthographic knowledge in perceptual development. *Journal of Experimental Child Psychology*. 1980; 30:411–421. doi: 10.1016/0022-0965(80)90047-8. [PubMed: 7205138]
- Medler, DA.; Binder, JR. MCWord: an on-line orthographic database of the English language. 2005. <http://www.neuro.mcw.edu/mcword/>
- Newman RL, Connolly JF. Determining the role of phonology in silent reading using event-related brain potentials. *Cognitive Brain Research*. 2004; 21:94–105. doi: 10.1016/j.cogbrainres.2004.05.006. [PubMed: 15325417]
- Nobre AC, Allison T, McCarthy G. Word recognition in the human inferior temporal lobe. *Nature*. 1994; 372:260–263. doi: 10.1038/372260a0. [PubMed: 7969469]
- Nobre AC, McCarthy G. Language-related ERPs: Scalp distributions and modulation by word-type and semantic priming. *Journal of Cognitive Neuroscience*. 1994; 6:233–255. doi: 10.1162/jocn.1994.6.3.233. [PubMed: 23964974]
- Nobre AC, McCarthy G. Language-related field potentials in the anterior-medial temporal lobe: II. Effects of word type and semantic priming. *Journal of Neuroscience*. 1995; 15:1090–1098. [PubMed: 7869085]
- Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*. 1971; 9:97–113. doi: 10.1016/0028-3932(71)90067-4. [PubMed: 5146491]
- Perrin F, Pernier J, Bertrand O, Echallier JF. Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*. 1989; 72:184–187. doi: 10.1016/0013-4694(89)90180-6. [PubMed: 2464490]
- Petersen SE, Fox PT, Snyder A, Raichle ME. Activation of prestriate and frontal cortical activity by words and word-like stimuli. *Science*. 1990; 249:1041–1044. doi: 10.1126/science.2396097. [PubMed: 2396097]
- Polich J. Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*. 2007; 118:2128–2148. doi: 10.1016/j.clinph.2007.04.019. [PubMed: 17573239]
- Rosinski RR, Wheeler KE. Children's use of orthographic structure in word discrimination. *Psychonomic Science*. 1972; 26:97–98. doi: 10.3758/BF03335445.
- Rugg MD, Nagy ME. Lexical contribution to nonword-repetition effects: Evidence from event-related potentials. *Memory and Cognition*. 1987; 15:473–481. doi: 10.3758/BF03198381. [PubMed: 3695941]
- Samuels SJ, LaBerge D, Bremer CD. Units of word recognition: evidence for developmental changes. *Journal of Verbal Learning and Verbal Behavior*. 1978; 17:715–720. doi: 10.1016/S0022-5371(78)90433-4.
- Smith RC, Dixon TR. Frequency and the judged familiarity of meaningful words. *Journal of Experimental Psychology*. 1971; 88:279–281. doi: 10.1037/h0031002.
- Stanovich KE. Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*. 1980; 16(1):32–71. doi: 10.2307/747348.
- Stanovich KE. Matthew Effects in reading: some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*. 1986; 21(4):360–407. doi: 10.1598/RRQ.21.4.1.
- Treiman R. The foundations of literacy. *Current Directions in Psychological Science*. 2000; 9(3):89–92. doi: 10.1111/1467-8721.00067.
- Vergara-Martínez M, Swaab TY. Orthographic neighborhood effects as a function of word frequency: an event-related potential study. *Psychophysiology*. 2012; 49:1277–1289. doi: 10.1111/j.1469-8986.2012.01410.x.
- Wagner RK, Torgesen JK. The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*. 1987; 101(2):192–212. doi: 10.1037//0033-2909.101.2.192.
- Wolf M. Naming speed and reading: The contribution of the cognitive neurosciences. *Reading Research Quarterly*. 1991; 26(2):123–141. doi: 10.2307/747978.
- Zeno, SM.; Ivens, SH.; Millard, RT.; Duvvuri, R. The educator's word frequency guide. Touchstone Applied Science Associates, Inc.; Brewster, NY: 1995.

Research Highlights

- College students and third, fourth, and fifth graders viewed real words, pseudowords, letter strings, false font strings, and animal name targets in an event-related potential (ERP) semantic categorization task
- Words and pseudowords elicited similar N400s across groups, but false font strings elicited N400s similar to words and letter strings only in the three groups of children
- The ERP findings suggested relatively adult-like semantic and phonological processing by third grade, but a lengthy developmental time course, beyond fifth grade, for orthographic processing
- The N400 findings were not consistent with a shift or discontinuity in word processing around the fourth grade, the putative fourth grade shift characterized in behavioral and educational literatures

Figure 1.

A bar graph illustrating overall N400 amplitude to words (black), pseudowords (red), letter strings (dark blue), and false fonts (light blue) across the four groups (college students, fifth graders, fourth graders, and third graders). Negative is plotted up and standard error is indicated.

Figure 2.

Grand average ERP waveforms elicited by words (black), pseudowords (red), letter strings (dark blue), and false fonts (light blue) for the four groups (college students, fifth graders, fourth graders, and third graders) at a single electrode site (central, medial C3). For each plot, each vertical tick along the x-axis marks 100 ms, negative is plotted up, the calibration bar (y-axis) marks 4.0 μV , and the N400 is identified.

Figure 3.

Topographical voltage maps illustrating the semantic (words - pseudowords), phonological (pseudowords - letter strings), and orthographic (letter strings - false fonts) effects for the four groups. A spherical spline interpolation (Perrin et al., 1989) was used to interpolate the potential on the surface of an idealized, spherical head based on the voltages measured from the difference waves at each electrode site within the 300-550 ms time window.

Table 1

Stimulus characteristics [mean, (SD)]

Examples	# Letters	# Phonemes	Frequency (Zeno) ^a	Frequency (KF) ^b	Orthographic Neighborhood (N) ^c	Constrained Bigram Count ^c	Constrained Trigram Count ^c	Unconstrained Bigram Count ^c	Unconstrained Trigram Count ^c
real words	4.0 (.8)	3.3 (.7)	811.1 (609.4)	548.6 (453.0)	11.0 (6.8)	27.1 (17.1)	4.8 (3.8)	2987.0 (1509.0)	350.7 (480.0)
pseudowords	4.0 (.8)	3.3 (.7)	n/a	n/a	10.2 (6.1)	25.6 (17.5)	3.6 (4.0)	2748.2 (1676.8)	306.8 (545.6)
letter strings	4.0 (.8)	3.6 (.8)	n/a	n/a	0.7 (1.1)	2.4 (3.2)	0.1 (0.2)	1403.5 (1073.3)	49.7 (130.4)
false fonts	BEM,,FRING	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
animal names	cat, duck	3.2 (.6)	68.1 (106.1)	12.8 (19.9)	10.1 (6.3)	24.6 (16.9)	4.2 (3.2)	2359.0 (1305.7)	221.4 (220.8)

^aBased on grades 1 through 5 in Zeno et al. (1995).

^bBased on the Kucera and Francis (1967) corpus.

^cFrom MCWord (Medler & Binder, 2005).

Table 2

Average number of trials (SD) included in individual ERP averages for each stimulus category (condition) by group (maximum 60)

	Real Words	Pseudo-words	Letter Strings	False Fonts	Animal Names
College Students	51.0 (6.4)	52.3 (7.0)	51.5 (7.5)	52.8 (6.8)	53.3 (5.7)
Fifth Graders	49.1 (6.3)	50.0 (6.3)	49.5 (7.8)	50.1 (6.9)	48.3 (6.6)
Fourth Graders	43.0 (9.6)	45.1 (8.9)	45.9 (9.4)	44.7 (9.5)	41.3 (8.0)
Third Graders	46.6 (7.0)	45.9 (8.3)	45.8 (7.4)	47.6 (6.5)	44.0 (7.0)

Table 3

Summary of pair-wise comparisons between conditions in within-group analyses of the main effect of condition

	Words / Pseudo-words	Words / Letter Strings	Words / False Fonts	Pseudo-words / Letter Strings	Pseudo-words / False Fonts	Letter Strings / False Fonts
College Students	=	>	>	>	>	>
Fifth Graders	=	>	=	>	>	=
Fourth Graders	=	=	=	>	>	=
Third Graders	=	>	=	>	>	=

Note. > indicates that the first member of the pair elicited a more negative (larger) N400 overall than the second; = indicates that the amplitudes of the N400s elicited by each member of the pair were not significantly different overall