HHS Public Access
Author manuscript
Lang Cogn Neurosci. Author manuscript; available in PMC 2021 January 01.

Published in final edited form as:
Lang Cogn Neurosci. 2020 ; 35(5): 641–657. doi:10.1080/23273798.2018.1503309.

Harry Potter and the Chamber of *What*?: The impact of what individuals know on word processing during reading

Melissa Troyer, **Marta Kutas**
Department of Cognitive Science, University of California, San Diego, United States of America

Abstract

During reading, effects of contextual support indexed by N400—a brain potential sensitive to semantic activation/retrieval—amplitude are presumably mediated by comprehenders' world knowledge. Moreover, variability in knowledge may influence the contents, timing, and mechanisms of what is brought to mind during real-time sentence processing. Since it is infeasible to assess the entirety of each individual's knowledge, we investigated a limited domain—the narrative world of Harry Potter (HP). We recorded event-related brain potentials while participants read sentences ending in words more/less contextually supported. For sentences about HP, but not about general topics, contextual N400 effects were graded according to individual participants' HP knowledge. Our results not only confirm that context affects semantic processing by ~250 ms or earlier, on average, but empirically demonstrate what has until now been assumed—that N400 context effects are a function of each individual's knowledge, which here is highly correlated with their reading experience.

Keywords

N400; individual differences; knowledge; sentence processing

Language comprehenders rapidly use many sources of contextual information to make sense of written and spoken words in sentences. For example, a word's processing is facilitated when it is preceded by a supportive (vs. unsupportive) sentence or discourse context, as indicated by electrophysiological and behavioral measures (e.g., Kutas & Hillyard, 1980; Altmann & Kamide, 1999). Word processing is also facilitated by non-linguistic context, such as a co-present visual scene (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Altmann & Kamide, 1999) or speaker identity (Van Berkum, van den Brink, Tesink, Kos, & Hagoort, 2008; Borovsky & Creel, 2014; see Brown-Schmidt, Yoon, & Ryskin, 2015, for a review).

A common method for estimating the strength of contextual support for a specific word is offline cloze probability norming, wherein a group of participants provide continuations for sentence fragments. The cloze probability of a word is defined as the proportion of

Corresponding author: Melissa Troyer, 9500 Gilman Drive, La Jolla, CA 92039, mtroyer@ucsd.edu.
Conflicts of interest
The authors declare no conflicts of interest.

participants who provide that word for a given context. A word's cloze probability is inversely correlated with the amplitude of a centro-parietal, negative-going event-related brain potential (ERP) between ~300-500 ms referred to as the N400 (Kutas & Hillyard, 1980, 1984; DeLong, Urbach, & Kutas, 2005; Wlotko & Federmeier, 2012). Empirically, large N400 amplitude appears to be part of the "default" response to words, with N400 amplitude reductions occurring for words in later, vs. earlier, sentence positions, taken to reflect greater ease of access with contextual accrual (Van Petten & Kutas, 1990), and for words semantically related to a sentence context and/or a predictable (though never encountered) sentence continuation (Kutas & Hillyard, 1984; Federmeier & Kutas, 1999; Metusalem, Kutas, Urbach, Hare, McRae, & Elman, 2012; Amsel, DeLong, & Kutas, 2015). Although accounts of the functional significance of N400 effects differ (see Kutas & Federmeier, 2011 and Kuperberg, 2016 for recent reviews), most seem to agree that variation in N400 amplitude reflects aspects of semantic processing.

As individuals process a word in context, contextual information presumably activates world knowledge, which is rapidly combined to form higher-level representations, allowing individuals to incrementally update their interpretation of the sentence and to activate, or pre-activate, a current or upcoming word's meaning (e.g., DeLong et al., 2005; Boudewyn, Long, & Swaab, 2015; see Kuperberg, 2016, for a recent discussion). On such an assumption, the relationship between a word's contextual support and N400 amplitude must be mediated by an individual's world knowledge.

Consistent with this assumption, Hagoort and colleagues (Hagoort, Hald, Bastiaansen, & Petersson, 2004) found that violations of culturally-specific knowledge elicited N400 potentials of similar scalp topography, timing, and magnitude to those of more general semantic violations. When Dutch participants read sentences such as '*The Dutch trains are {yellow / white / sour}*,' words like '*yellow*' (a supported continuation) elicited reduced N400 activity (from 250-550 ms at centro-parietal electrodes) compared to '*sour*' (unsupported), replicating many studies. However, the N400 response to words like '*white*' —a word inconsistent with culturally-specific knowledge that Dutch trains are yellow— elicited an N400 indistinguishable from that of '*sour*.' These findings, among others (e.g., Hald, Steenbeek-Planting, & Hagoort, 2007; Filik & Leuthold, 2013) underscore that world knowledge, gleaned from actual experience in the world, determines aspects of processing reflected in N400 amplitudes.

Even within the same cultural context, however, different individuals know different things and to varying extents. Consequently, measures of contextual support like cloze probability computed over groups of participants do not necessarily correspond to the contents of any given individual's knowledge or the likelihood that they entertained a given word or concept during moment-by-moment processing (see Verhagen, Mos, Backus, & Schilperoord, 2018, for discussion). Moreover, ERPs themselves are typically averaged over a group of participants and are not generally linked to individual participants' knowledge (though see Coronel & Federmeier, 2016, who established a relationship between knowledge of personal preferences and N400 effects at the individual level).

Across many domains, the content of an individual's knowledge can influence performance on memory tasks (e.g., de Groot, 1965; Simon & Chase, 1973), text comprehension (e.g., Chiesi, Spilich, & Voss, 1979; Spilich, Vesonder, Chiesi, & Voss, 1979), and other cognitive tasks (see Chi, 2006, for a review). Yet, although variability in individuals' experiences and, therefore, knowledge seems likely to influence real-time processing (as argued by Kutas, 2006), few studies have examined the link between level of content knowledge and real-time sentence processing. More typically, the experimental focus has been on how differences in more generalised cognitive processing abilities, such as verbal working memory (WM) and cognitive control (Münte, Schiltz, & Kutas, 1998; Nakano, Saron, & Swaab, 2010; Boudewyn, Long, & Swaab, 2012), or language-specific abilities, such as language proficiency (McLaughlin, Osterhout, & Kim, 2004; McLaughlin, Tanner, Pitkänen, Frenck-Mestre, Inoue, Valentine, & Osterhout, 2010; Pakulak & Neville, 2010; Tanner, Inoue, & Osterhout, 2014), relate to individual differences in aspects of sentence processing (see Boudewyn, 2015, for a review).

Clearly a major hurdle for investigating individual differences in knowledge is infeasibility of capturing all of an individual's knowledge using standard laboratory procedures. As an approximation, measures of print exposure and general knowledge (Stanovich & West, 1989; Stanovich & Cunningham, 1993) have been tentatively linked to real-time brain potential measures of sentence processing (e.g., Metusalem et al., 2012); however, these provide only coarse approximations of general knowledge and not precise topics or domains where individuals may vary in how much they know.

A potential solution may be to restrict knowledge to a single domain. In the vast literature on expertise and expert behaviour, researchers interested in how level of knowledge impacts perception and cognition have focused on specific domains of knowledge—e.g., chess (de Groot, 1965; Simon & Chase, 1973), physics (e.g., Chi, Feltovich, & Glaser, 1981), children's knowledge of dinosaurs or spiders (e.g., Pearson, Hansen, & Gordon, 1979; Gobbo & Chi, 1986). Few, however, lend themselves easily to the study of real-time language processing, which requires a domain with a rich set of verbal descriptions. Moreover, based on a relatively low signal-to-noise ratio, ERP studies typically require a number of linguistic stimuli on the order of a hundred or more per study.

To mitigate these challenges, we zeroed in on the narrative world of Harry Potter, based on the book series by J.K. Rowling. This domain is linguistically rich, making it an excellent choice for studying the interface of knowledge and language. In addition, it includes many novel organizational structures—including new categories, like magical creatures and spells, as well as multi-faceted events. At the same time, it is a constrained domain, allowing us to estimate each individual's level of HP knowledge. Finally, as present-day college students grew up with this series, they constitute a pool of participants who are likely to naturally vary in their HP knowledge.

Using this tractable domain of knowledge, we asked whether or not, and, if so, to what extent and when individual differences in knowledge about HP would specifically influence processing of sentences about the domain of HP (and not general topics). We therefore recorded ERPs while participants read sentences about general topics (control sentences) and Harry Potter (HP sentences) that ended in contextually supported or unsupported words (Table 1). These participants also completed offline questionnaires that assessed their knowledge about Harry Potter.

Many studies seem to assume that real-world knowledge at the individual participant level modulates the effect of contextual support on the earliest stages of semantic processing, as captured during the N400 time period. However, no study has directly tested the assumption that specific knowledge, at the individual-subject level, modulates the effect of contextual support. Here, we explicitly tested this assumption by using a single domain of knowledge to carefully manipulate the availability of knowledge on a subject-by-subject basis. We predicted that HP knowledge would specifically influence the size of N400 context effects in HP, but not control, sentences. Such a finding would provide direct evidence that an individual's level of knowledge of a domain modulates the influence of context during the earliest stages of semantic processing.

As manipulations that affect N400 amplitude often influence the size of so-called "post-N400" positivities (PNPs) at times, we also examined a later time period. Depending on the nature of the linguistic manipulation, words which are unsupported in context may elicit parietal and/or frontal positivities (for reviews, see Van Petten & Luka, 2012; Brouwer, Fitz, & Hoeks, 2012; DeLong, Troyer & Kutas, 2014). We therefore had reason to suspect that there might be late positivities of larger amplitude for unsupported (compared to supported) critical words in the control sentences, and, for individuals knowledgeable about HP (and therefore sensitive to the contextual support manipulation), in the HP sentences.

Methods

Participants

41 undergraduate students / members of the UCSD community (mean age = 20, range = 18-24; 26 women, 15 men) took part in the study for partial course credit or payment of $9 / hour. Of these, one participant was excluded from data analysis due to excessive eye movements. All participants provided informed consent reviewed by the Institutional Review Board at the University of California, San Diego. To ensure that some participants would have high knowledge of the Harry Potter domain, a subset (n = 12) was recruited via an announcement that specifically required having read all seven Harry Potter books and/or seen all eight Harry Potter films.

Materials

Sentence materials—During the EEG portion of the experiment, participants read three blocks of control sentence pairs followed by three blocks of Harry Potter (HP) sentence pairs. The final word of the second sentence was always the critical word, which was either supported or unsupported by the context. Two lists were created; each participant only read one version of each sentence pair.

Control sentences. 108 control sentence pairs (first sentence ranging 3-18 words, mean = 7; second ranging 4-10 words, mean = 7) described everyday topics and events. All sentence pairs were highly constraining (mean cloze of best completion = 94%; range = 87-100%).

For control sentences, supported words were defined as the best completion. To create unsupported words, plausible continuations were selected that were semantically related to the best completion but were never produced during cloze norming.

Harry Potter sentences.—108 Harry Potter (HP) sentence pairs were constructed as follows. Using freely available materials (including Wikipedia and Harry Potter fan sites) along with the text of the Harry Potter books, the first authaor created a set of single sentences that accurately described events and entities from the series. The final, "supported" word of these sentences was designed to be 100% predictable given perfect knowledge of the book series. To verify that this was the case, a norming study was conducted on a separate group of participants. This group included some participants who were highly knowledgeable about the world of Harry Potter (determined by a trivia quiz; see "Harry Potter Quiz" section below). 32-34 participants provided a final word for each sentence. To be included in the study, a sentence needed to be completed with the supported word by a minimum of 65% of the most knowledgeable respondents (that is, those who scored in the top quartile on the HP knowledge quiz). Across all norming participants, mean cloze for supported words was 51% (range = 26-84%).

These single sentences were then broken into two sentences so that the first could be presented all at once. On average, the first sentence was 9 words long (range = 4 to 18 words); the second was 7 words long (range = 3 to 12 words). The second sentence ended in a critical word that either was supported by the context (see above) or was unsupported. To create unsupported endings, supported words were replaced by words that defied "ground truth" from the HP stories but seemed similarly plausible (for those with little to no HP knowledge). To achieve this, we used words that were from a the same/a similar category. For example, a standard English word describing an animal, like "rat," was replaced with another animal, "dog"; a magical object specific to HP like "Bludgers" was replaced with another magical object specific to HP, "Spellotape"; and an HP-specific proper name like "Kreacher" was replaced with another HP-specific proper name like "Hermes" (see these and other examples in Table 1). After data collection, we discovered that two HP sentence pairs contained factual errors in both the supported/unsupported versions. These materials were dropped from subsequent analyses for a total of 106 HP items and 108 Control items.

Memory tests—Immediately after the EEG portion of the experiment, participants completed two memory quizzes, one for control sentences and one for HP sentences. The primary purpose of the memory tests was to establish that participants had paid attention when reading during the EEG experiment. Participants were asked to circle the words they remembered seeing as a final word of the second sentences from the experiment—first for control sentences and then for HP sentences. Each quiz contained a total of 90 words—30 new, 30 critical words from supported contexts, and 30 critical words from unsupported contexts.

Additional tasks and measures

Overview of tasks.: We collected several other measures of individual differences besides HP knowledge to better understand any group differences between individuals with high vs.

low HP knowledge. We developed a measure of Harry Potter experience (self-report questionnaire) and collected other measures (see Appendix, in the supplementary materials) including general print/reading experience (media and reading habits questionnaire (MRH) and author and magazine recognition tests (ART/MRT), Stanovich & West, 1989); measures of general knowledge (a general knowledge trivia quiz (GKQ) that we developed from freely available materials and cultural knowledge checklists (CLC/MCLC, Stanovich & Cunningham, 1993)); vocabulary (PPVT, Dunn & Dunn, 2007), and verbal WM (sentence span, Daneman & Carpenter, 1980). Finally, we administered a debriefing questionnaire.

HP self-report questionnaire.: We asked participants questions about their experience with the Harry Potter book series and related materials, e.g., how many times they had read each of the HP books, seen each of the HP movies, and additional ways in which they might have engaged with Harry Potter. As an estimate of overall experience with Harry Potter, a numeric score was determined by summing the total number of times an individual had read each book, seen each movie, and so on. In addition, we report statistics on the raw number of times participants read each book, on average (e.g., if an individual read the first book 3 times and each other book just once, their score would be 1.286).

Harry Potter quiz.: We estimated participants' knowledge of Harry Potter from their score on a trivia-style quiz containing ten multiple choice questions; for example, *To gain access to the kitchens, one must tickle the following fruit: (a) Pear, (b) Orange, (c) Grape, (d) Banana*. HP quiz score (henceforth referred to as "HP knowledge") was the number of correct answers out of ten. For regression analyses, we z-transformed these scores.

Aggregate measures.: An aggregate measure of reading experience was based on an average of z-transformed ART scores, MRT scores, total MRH score, and number of favorite authors listed on the MRH. An aggregate measure of general knowledge of common topics was based on an average of z-transformed CLC scores, MCLC scores, and general knowledge test scores.

Debriefing questionnaire.: On a debriefing questionnaire, many participants indicated they noticed that sentences about Harry Potter were sometimes inaccurate. This was to be expected, as half of the sentences were designed to be inaccurate portrayals of "ground truth" based on the HP book series, and all participants had at least minimal knowledge of the HP series. After observing this trend, we asked all but the first two participants to complete additional debriefing questions, estimating approximately how many sentences they thought were true/accurate, and, of these, how many they thought they had known ahead of time. Participants reported that 60% of the Harry Potter sentences had been true (range = 30-100%). Of these, participants reported that they had known an average of 64% (range = 0-100%).

Procedures

Ordering of tasks.—During set-up for the EEG experiment, participants completed the ART and MRT. After EEG recording, participants completed the memory tests followed by other questionnaires, with the order corresponding to that of their description in the preceding section.

EEG experiment.—Before the study, participants were asked to remain relaxed and still to minimize muscle artifact. They were told they would be reading short, two-sentence stories (first three blocks about general topics, then three blocks about the world of Harry Potter) for meaning and that they would be asked questions about what they read at the end of the EEG recording session. Participants then read four practice items.

During the EEG experiment, participants sat approximately 100 cm in front of a cathode-ray tube monitor. The background of the screen was black and words were presented in white type. Each trial began with a blank screen for two seconds. Then, the first sentence of each pair appeared on the screen until the participant pressed a button to advance to the next sentence. After their button press, a crosshair appeared in the center of the screen for a duration which varied randomly between 1050 and 1450 ms. Participants were instructed to focus on the crosshair and not to move their eyes or blink while it was on the screen. The second sentence was then presented one word at a time right above the crosshair. Each word was presented for 200 ms with an interstimulus interval of 300 ms. After the sentence-final word disappeared, the crosshair stayed on the screen for a duration that randomly varied between 750 and 1150 ms. Control sentences were presented across three blocks, with short breaks in between, followed by three blocks of HP sentences. Within each block of the study, sentences with supported and unsupported endings were randomly interspersed.

EEG recording

The electroencephalogram (EEG) was recorded from 26 electrode sites arranged geodesically in an Electro-cap (as described in Ganis, Kutas, & Sereno, 1996; see Fig. 9). For all cap electrodes, online recording was referenced to the left mastoid; these electrodes were re-referenced offline to an average of the left and right mastoid. Electrodes were placed lateral to the outer canthus of each eye to create a bipolar recording used to monitor eye movements. Electrodes placed under each eye were referenced to the left mastoid and were used to monitor blinks. Throughout the experiment, all electrode impedances were maintained under 5 kΩ. The signal was amplified with Grass amplifiers which were set at a bandpass of .01 to 100 Hz; the sampling rate was 250 Hz.

EEG data analysis

Trials contaminated by eye movements, blinks, muscle activity, blocking, or other artifact were removed from subsequent analysis. This resulted in an exclusion of 17% of trials: HP-Supported: 17%; HP-Unsupported: 17%; Control-Supported: 18%; Control-Unsupported: 17%. ERPs were created by averaging from 200 ms before the onset of a critical word until 900 ms post-critical word. Then, for each electrode, a baseline was computed by averaging potentials from 200 ms before the word to the start of the word; this baseline was subtracted from the waveform.

Because our study is the first to directly compare ERPs to words in sentences about general topics with those in sentences about a fictional, narrative world, it was important to characterize overall influences of contextual support and sentence type across all participants. We therefore conducted a traditional, whole-head analysis prior to examining individual differences based on HP knowledge and other covariates. Of primary interest was a time period surrounding the typical peak of the N400 brain potential (~375 ms; e.g., Federmeier & Kutas, 1999) from 250 ms to 500 ms post-stimulus; we also examined a late positivity time period from 500-750 ms post-stimulus. We subjected mean amplitudes of the ERP waveforms in these time periods to a whole-head ANOVA to determine effects of sentence and/or ending type across all participants, including repeated measures of electrode (26 levels), ending type (2 levels: supported/unsupported), and sentence type (2 levels: HP/control) as well as a between-subjects factor of list (2 levels). For all ANOVAs, we applied the Greenhouse-Geisser epsilon correction for F-tests with more than one degree of freedom in the numerator and report the corrected p-value, unadjusted degrees of freedom, and value of the Greenhouse-Geisser epsilon.

Our primary research questions hinged on whether and when HP knowledge interacted with contextual support. We therefore examined the relationship between HP knowledge, sentence type, and ending type in an ROI where N400 effects are typically largest, averaging mean amplitude between 250 and 500 ms across eight centro-parietal electrodes (MiCe, LMCe, RMCe, MiPa, LDPa, RDPa, LMOc, and RMOc) for each sentence and ending type. HP knowledge was defined as z-transformed performance on the HP knowledge quiz. For these analyses, we used hierarchical mixed-effects linear regression models (Baayen, Davidson, & Bates, 2008). All categorical fixed effects (sentence type, ending type) were sum coded ([−1, 1]). All models were fit to subject-averaged ERPs, with random intercepts for subject. Mixed effects models were implemented using the lme4 (version 1.1-12; Bates, Maechler, Bolker, & Walker, 2015) and lmerTest (version 2.0-30; Kuznetsova, Brockhoff, & Christensen, 2017) packages in R. P-values were calculated using lmerTest, with the Satterthwaite option for denominator degrees of freedom for F statistics.

Results

Behavioral data

Memory task—recognition accuracy—On the control recognition test, participants correctly recognised an average of 15.26 out of 60 words (~25%) and false alarmed to an average of 2.08 out of 30 words (~7%). On the HP recognition test, participants correctly recognised an average of 29.69 out of 60 words (~49%) and false alarmed to an average of 2.28 out of 30 words (~8%). Participants were therefore able to discriminate between words they had and had not seen for both the Control and HP recognition tests.

For statistical analyses, we computed a d-prime sensitivity index for each participant and condition based on the false alarms for each recognition test (control, HP) and the number of items correctly recognised from each condition. We used a mixed-effects model to predict d-prime based on sentence type (control, HP), ending type (supported, unsupported), and HP knowledge (Supplemental table 1). This model revealed a significant effect of sentence type, indicating higher accuracy for HP compared to control sentences. In addition, the interaction between HP knowledge and sentence type was significant. HP knowledge was correlated with d-prime for HP words ($r^2 = .136$, $p < .05$), but not for control words ($r^2 = .001$, n.s.).

Additional tasks—Table 2 reports descriptive statistics for scores on the HP knowledge quiz and other individual differences measures completed by participants. Intercorrelations among these measures are provided in Table 3.

ERP data

Figure 1 shows the grand average ERPs for all participants across 26 scalp electrodes from 200 ms before the onset of the critical word to 900 ms post-critical word. Across most electrodes, ERPs to critical words in both control and HP sentences are characterised by two early sensory components, a negative-going peak around 100 ms (N1) and a positive-going peak around 200 ms (P2). Across all participants, for supported words, the P2 is followed by a positivity in the N400 time period (250-500 ms). For unsupported endings, the P2 is followed by a relative negativity in this same window.

Since we were specifically interested in the unique effects of HP knowledge on processing HP, compared to control, sentences, whole-head plots for the high-knowledge group ($n = 15$) and low-knowledge group ($n = 19$) for each sentence type are provided in Figure 2.

N400: 250-500 ms post-stimulus

Whole-head analyses.—Results from the whole-head ANOVA for the N400 time period are provided in Table 4. As expected, there was a main effect of ending type, with supported endings leading to more positive-going waves (i.e., reduced negativities) than unsupported endings. Visual inspection of topographic scalp maps (Figure 1) revealed that an interaction between electrode and ending type was driven by a broadly distributed, centro-parietal N400 context effect. The distribution was roughly similar for both sentence types, though the context effect seemed somewhat more broadly distributed for HP, compared to control, sentences.[1]

Three- and four-way interactions with list suggest there may have been subtle differences in N400 amplitude based on which list participants saw. However, as the relevant values were in the same direction for each of the relevant comparisons (that is, unsupported values were more negative than supported values for each sentence type and for each of the two lists), we do not further pursue this point.

ROI analyses.—Figure 3 shows ERPs from the centro-parietal ROI used in regression analyses. Results from a linear mixed-effects regression analysis are provided in Table 5. This regression confirmed the effect of ending type (i.e., contextual support) observed in the whole-head analysis. In addition, there was a significant interaction between sentence type and HP knowledge, with more knowledgeable individuals tending to have overall more positive-going N400s compared to less knowledgeable individuals. Critically, the three-way interaction between HP knowledge, sentence type, and ending type was significant. To follow up, we conducted planned analyses separately for each sentence type (Table 6).

[1]For N400 context effects, we followed up on interactions with electrode in a distribution analysis containing a subset of 16 electrodes (following the procedure in Kutas & Federmeier, 1999) which supported this interpretation. For both sentences types, N400 context effects were centro-posterior; for control, but not HP, sentences, the N400 context effect was somewhat right lateralized.

For control sentences, ending type was a significant predictor, but, critically, there were no effects of HP knowledge. For HP sentences, however, ending type, HP knowledge, and their interaction were all significant predictors. To further explore the relationship between N400 HP context effects and HP knowledge score, we conducted planned simple regressions predicting mean amplitude for supported, as well as unsupported, endings in this time period based on HP knowledge score, finding that this relationship seemed to be driven by variance in the supported ($r^2 = .167$, $p < .01$) but not unsupported ($r^2 = .021$, n.s.) endings.

Therefore, across the centro-parietal ROI, effects of contextual support for sentences about Harry Potter were sensitive to HP knowledge, but, critically, effects of contextual support for sentences about general topics were not. Moreover, this relationship reflected a difference in the brain's response to *supported*, not unsupported, endings.

Next, we asked whether other differences between participants might modulate the influence of contextual support. We used mixed-effects linear regression to predict each mean amplitude based on four variables: HP knowledge, verbal WM, reading experience, and general knowledge, as well as their interaction with ending type (see Supplementary Table 2).

Confirming previous results, ending type was a significant predictor for both sentence types. For control sentences, none of the individual-differences measures nor their interaction with ending type were significant predictors. For HP sentences, however, the interaction between reading experience and ending type was a significant predictor. However, the aggregate measure of reading experience was strongly correlated with HP knowledge scores (at $r = .54$; see Table 3 and Supplementary Figure 1); this multicollinearity makes it impossible to fully dissociate effects of each predictor variable on the size of the effect of contextual support (Kutner, Nachtsheim, Neter, & Li, 2005). We return to this point in the discussion.

Late positivity: 500-750 ms post-stimulus

Whole-head analyses.—Results from the whole-head ANOVA for the late positivity time period are provided in Table 4. A main effect of sentence indicated more positive-going waves for HP sentences compared to control sentence endings. An interaction of ending type × sentence type indicated a crossover effect such that the direction of the contextual support effect differed between the two sentence: whereas for control sentences, unsupported endings led to more positive-going waves compared to supported endings, the numerically reverse pattern obtained for HP sentences, with unsupported endings leading to less positive-going waves compared to supported endings. Based on visual inspection, the higher-order interactions with electrode seemed to reflect a frontal, left positivity greater for unsupported compared to supported endings for the control, but not HP, sentences.[2]

ROI analyses.—A mixed-effects linear regression analysis predicting potentials from the centro-parietal ROI based on HP knowledge, sentence type, and ending type (Table 5)

[2]For post-N400 context effects, we followed up on interactions with electrode in a distribution analyses containing a subset of 16 electrodes (as for our N400 context effects). These analyses confirmed that for control sentences, N400 context effects were left/frontal, while no context effects were present on the late positivity for HP sentences.

confirmed the effect of sentence type observed in the whole-head analysis, with more positive-going potentials for HP compared to control sentences. No other predictors were significant.

Discussion

Summary of findings

We used the narrative world of Harry Potter to ask whether level of domain knowledge could predict sentential context effects for individuals. We observed N400 context effects for sentences about Harry Potter, as well as for control sentences about general topics, but the size of the context effects—only for the HP sentences—was significantly predicted by an individual's level of HP knowledge. These effects were observed during the earliest stages of semantic processing (between ~250-500 ms post-stimulus), and the relationship between context effects and knowledge seemed to be driven primarily by supported, not unsupported, words, suggesting greater ease of access to, or retrieval of, pertinent information for individuals with high knowledge scores. Indeed, differential experience in a domain can lead to differences not only in which "chunks"—i.e., "coherent patterns" (Chi, 2006, p. 181; cf. Miller, 1956; Simon, 1974)—are available to individuals but also in the depth of information that comes to mind when a cue (or chunk), like a picture or word, is processed (Chi, 2006). We speculate that our HP knowledge measure may act as a proxy for some combination of chunk size, depth of information access, or other differential organization of information.

N400 context effects

As expected based on the extensive N400 literature (Kutas & Federmeier, 2011), we observed N400 context effects for sentences ending in supported vs. unsupported words, whether about general topics or about Harry Potter, a well-known fictional domain with which many young adults are familiar. The control and HP N400 context effects both had similar morphology, onset (~200 ms), and timing (peaking around ~375 ms). These similarities are consistent with previous reports that contextual information combines with knowledge to almost immediately influence semantic processing (e.g., Hald et al., 2007), even for knowledge of fictional characters from popular culture (Filik & Leuthold, 2013) or for fictional descriptions that override veridical real-world knowledge (Nieuwland & Van Berkum, 2006). Our findings of N400 context effects in sentences about the people, places, objects, and ideas from the fictional world of Harry Potter affirm that experiences gleaned from many sources—including the printed word or images on the movie screen, as well as the real world—combine to form the knowledge that influences the earliest stages of semantic processing during word by word reading.

Since the original reports of N400 amplitude modulation based on contextual support (Kutas & Hillyard, 1980; Kutas & Hillyard, 1984; Kutas, Lindamood, & Hillyard, 1984), many have demonstrated that a word's N400 response is highly sensitive to factors related to semantic retrieval, including whether the word is a content word vs. function word (Kutas, Van Petten, & Besson, 1988); the word's lexical frequency (Van Petten & Kutas, 1990; Kutas & Federmeier, 2000); how much context has accrued over the course of a sentence (Van Petten & Kutas, 1990); and the word's relationship to a preceding context, whether the context be a single word (Heinze, Muente, & Kutas, 1998; Federmeier, Kutas, & Schul, 2010), a picture (Friedrich & Friederici, 2010; Knoeferle, Urbach, & Kutas, 2011), or a whole sentence (e.g., Kutas & Hillyard, 1984; DeLong et al., 2005; Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007). Presumably, the relationship between a word and context is mediated by knowledge in the head of the comprehender, which has heretofore been approximated using cloze probability for sentence contexts (e.g., DeLong et al., 2005). However, offline cloze probabilities do not provide precise estimates of any one individual's knowledge nor of what an individual brings to mind in the moment.

Here, by testing participants who varied in their knowledge of Harry Potter, we were able to directly and systematically investigate the availability of relevant knowledge during real-time sentence processing. We found that N400 context effects for HP, but not for control, sentences were graded with respect to the HP knowledge measure, with minimal to no effects for individuals with the lowest scores and the largest effects for those with the highest scores. By using a constrained domain in which we could estimate individual differences in knowledge, we explicitly demonstrated what has been to date implicitly assumed—namely, that specific knowledge relevant to interpreting written sentences is brought to bear—i.e., activated, or perhaps even pre-activated—relatively early (by ~200 ms or so) as words are encountered during real-time sentence processing.

Our interpretation is based on our assumption that HP knowledge level (determined by score on a trivia quiz) was directly proportional to the likelihood that an individual knew the information described in any given HP sentence pair in real time, during the EEG study. Beyond this relationship, if an HP expert and an HP novice both know a given fact (e.g., that Harry's scar has the form of a lightning bolt), they might nonetheless process the information differently, bringing more/less information to mind or doing so with a different time course. That is, differences in the functional organization of knowledge (e.g., Federmeier & Kutas, 1999) may impact processing above and beyond *whether* an individual knows a specific fact. To investigate these possibilities, we would want to compare individuals of differing knowledge levels on trials we know that they know vs. do not know (see Brothers, Swaab, & Traxler, 2015, 2017, who utilize trial-by-trial approaches to investigate effects of contextual support and judgments of lexical prediction).

We also asked whether the effect of contextual support observed in the HP sentences was influenced by other individual differences, including measures of reading experience, general knowledge, and verbal working memory. In the presence of these additional variables, we found that reading experience modulated the effect of contextual supported observed for HP sentences; HP knowledge had an overall effect, but did not interact with ending type.

It is important to note that HP knowledge and reading experience are correlated at $r = .54$ in our sample, thereby limiting our ability to determine precisely which drives the individual variation in N400 amplitude. This relatively high correlation is not surprising, as HP knowledge comes in large part from reading. For the moment, that HP knowledge (considered separately from reading experience) is predictive of contextual support effects for HP sentences but *not* for control sentences about general topics (see Table 5) leads us to speculate that HP knowledge mediates the observed relationship between reading experience and HP effects of contextual support on N400 amplitude. In future studies, we aim to better dissociate HP knowledge and reading experience, either by testing a sample of participants in whom the two measures are less correlated, or, better yet, by simultaneously investigating multiple domains of knowledge, allowing individuals who are experienced in one domain but inexperienced in the other to serve as their own controls.

Late positivity context effects

Although late positivity effects were not a focus of our study, we did observe systematic post-N400 positivity context effects for control sentences in the whole-head analysis, with unsupported words eliciting larger left anterior positivities compared to supported words. In our study, supported words were the best completion, and unsupported words were low-cloze yet plausible endings related in meaning to the best completion. These sentences were similar in nature to a subset of the sentences in Thornhill and Van Petten (2012), who likewise observed a frontal positivity for low-cloze, yet plausible words (both related and unrelated to the best continuation), compared to the best completion. Thornhill and Van Petten suggested anterior PNPs reflected the processing of lexically unexpected words; elsewhere, anterior PNPs have been linked to the processing of low-cloze congruent/ plausible words, with posterior PNPs linked to the processing of low-cloze incongruent/ implausible words (Van Petten & Luka, 2012; DeLong, Quante, & Kutas, 2014). Our findings from control sentences are consistent with both sets of hypotheses.

Perhaps surprisingly, we did not observe any effects of contextual support, nor any interaction between HP knowledge and contextual support (in the ROI analysis), for the HP sentences during the late time period. Because we designed our HP sentences such that, given little to no knowledge of HP, unsupported endings would seem similarly plausible to supported endings, we did not expect effects of the contextual support manipulation for low-knowledge individuals. It is, however, unclear why we did not observe effects of contextual support on late positivities for high-knowledge participants. Future work investigating trial-by-trial variation in sentence-specific knowledge along with individual differences in domain knowledge may shed light on this.

In both the whole-head and ROI analyses, post-N400 positivities were overall larger for HP than control sentence endings. HP sentences, by contrast to control sentences which described generalized situations, often described episodes/events in HP that had occurred. Individuals thus may have entertained specific, episodic information for some of the HP sentences. In the memory literature, late parietal positivities have been associated with episodic retrieval (cf. “old/new” effects, reviewed in Rugg & Curran, 2007). In future studies, asking participants to provide more information on a trial-by-trial level (e.g., whether they had known the information ahead of time, or whether they had brought episodic information to mind during reading) might shed light on knowledge-based individual differences of this nature.

Limitations and future directions

A limitation of our correlational approach is that participants were not randomly assigned levels of HP exposure, and there are other individual differences besides HP knowledge. Moreover, in a model including several measures of individual differences, it was reading experience, and not HP knowledge, that interacted with ending type. As reading experience and HP knowledge were positively correlated, we were unable to tease apart pure effects of HP knowledge from overall differences in reading experience in this study, though we leaned to the former. Future studies including multiple domains of knowledge simultaneously (e.g., separate book series with which individuals have differing amounts of exposure/knowledge) may be able to better identify individual differences due to specific domain knowledge vs. those differences due to differences in general levels of reading experience.

That individuals with differing levels of domain knowledge represent and process information within that domain differentially has been substantiated across many areas of expertise (see reviews in Ericsson, Charness, Hoffman, & Feltovich, 2006). However, the precise nature of differences in the representation of knowledge as a function of expertise remains an open question. How experts chunk, store, process, retrieve, and/or otherwise use information seem likely to differ depending on the nature of the information (motoric, perceptual, procedural, declarative, etc.) and the goals of the task at hand (e.g., winning a chess game, completing a physics problem, or enjoying a narrative) (see Chi, 2006, for discussion). In future work, we aim to better understand how differences in specific domain knowledge alter the nature of information that is readily cued from linguistic input and brought to mind during real-time sentence comprehension—including differences in the presence/level of detail brought to mind when reading about events (e.g., Metusalem et al., 2012; Amsel et al., 2015) or differences in how the features/categories related to a word are conceived of and accessed during real-time processing (e.g., Federmeier & Kutas, 1999; Federmeier, McLennan, De Ochoa, & Kutas, 2002).

Conclusions

In this study, we went beyond group-level measures of contextual support (i.e., cloze probability) by estimating individuals' knowledge in a specific domain, the narrative world of Harry Potter. We find that such knowledge estimates seem to predict patterns of neural activity, at the earliest neural stages of semantic processing, as reflected in N400 amplitude modulations. Future work can combine measures of individual-level knowledge in a domain with trial-by-trial estimates. Such an approach could prove powerful for investigating the relative contributions of (a) the functional organization of knowledge, which is likely to substantively differ between domain experts and novices, and (b) the likelihood that an individual knows any given item, which is probabilistically related to, but not necessarily determined by, an individual's overall level of knowledge. In sum, our results lay the groundwork for investigating how inter-individual differences in organization of knowledge (amount, content, or other aspects of internal organization/connectivity) influence aspects of knowledge use (including timing and depth/level of semantic processing) in real time sentence processing.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This research was supported by NICHD grant RO1HD022614 to MK and Frontiers of Innovation in Science Program grant to MT. We thank Katherine DeLong, Tom Urbach, Jeffrey Elman, Seana Coulson, Victor Ferreira, and Zhuowen Tu for feedback on this work.

References

Altmann GTM & Kamide Y (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. Cognition, 73, 79–87. doi: 10.1016/S0010-0277(99)00059-1

Amsel BD, DeLong KD, & Kutas M (2015). Close, but no garlic: Perceptuomotor and event knowledge activation during language comprehension. Journal of Memory and Language, 82, 118–132. doi: 10.1016/j.jml.2015.03.009 [PubMed: 25897182]

Baayen RH, Davidson DJ, & Bates DM (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language, 59, 390–412. doi; 10.1016/j.jml.2007.12.005

Bates D, Maechler M, Bolker B, & Walker S (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1), 1–48. doi: 10.18637/jss.v067.i01

Borovsky A & Creel S (2014). Children and adults integrate talker and verb information in online processing. Developmental Psychology, 50(5), 1600–1613. doi: 10.1037/a0035591 [PubMed: 24611671]

Boudewyn MA (2015). Individual differences in language processing: electrophysiological approaches. Language and Linguistics Compass, 9/10, 406–419. doi: 10.1111/lnc3.12167

Boudewyn MA, Long DL, & Swaab TY (2012). Cognitive control influences the use of meaning relations during spoken sentence comprehension. Neuropsychologia, 50, 2659–2668. doi: 10.1016/j.neuropsychologia.2012.07.019 [PubMed: 22842106]

Boudewyn MA, Long DL, & Swaab TY (2015). Graded expectations: Predictive processing and the adjustment of expectations during spoken language comprehension. Cognitive and Affective Behavioral Neuroscience, 15, 607–624. doi: 10.3758/s13415-015-0340-0

Brouwer H, Fitz H, & Hoeks J (2012). Getting real about semantic illusions: rethinking the functional role of the P600 in language comprehension. Brain Research, 1446, 127–143. doi: 10.1016/j.brainres.2012.01.055 [PubMed: 22361114]

Brothers T, Swaab TY, & Traxler MJ (2015). Effects of prediction and contextual support on lexical processing: Prediction takes precedence. Cognition, 136, 135–149. doi: j.cognition.2014.10.017 [PubMed: 25497522]

Brothers T, Swaab TY, & Traxler MJ (2017). Goals and strategies influence lexical prediction during sentence comprehension. Journal of Memory and Language, 93, 203–216. doi: 10.1016/j.jml.2016.10.002

Brown-Schmidt S, Yoon SO, & Ryskin RA (2015). People as contexts in conversation. Psychology of Learning and Motivation, 62, 59–99. doi: 10.1016/bs.plm.2014.09.003

Chi MTH (2006) Laboratory methods for assessing experts' and novices' knowledge. In Ericsson et al. (Eds.), The Cambridge Handbook of Expertise and Expert Performance. Cambridge University Press. doi: 10.1017/CBO9780511816796.010

Chi MTH, Feltovich PJ, & Glaser R (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5, 121–152. doi: 10.1207/s15516709cog0502_2

Chiesi HL, Spilich GJ, & Voss JF (1979). Acquisition of domain-related information in relation to high and low domain knowledge. Journal of Verbal Learning and Verbal Behavior, 18, 257–273. doi: 10.1016/S0022-5371(79)90146-4

Coronel JS & Federmeier KD (2016). The N400 reveals how personal semantics is processed: Insights into the nature and organization of self-knowledge. Neuropsychologia, 84, 36–43. doi: 10.1016/j.neuropsychologia.2016.01.029 [PubMed: 26825011]

Daneman M & Carpenter P (1980). Individual differences in working memory and reading. Journal of Verbal Memory and Verbal Behavior, 19(4), 450–466. doi: 10.1016/S0022-5371(80)90312-6

de Groot AD (1965). Thought and choice in chess Mouton, The Hague. doi: 10.5117/9789053569986

DeLong KA, Quante L, & Kutas M (2014). Predictability, plausibility, and two late ERP positivities during written sentence comprehension. Neuropsychologia, 61, 150–162. doi: 10.1016/j.neuropsychologia.2014.06.016 [PubMed: 24953958]

DeLong KA, Troyer M, & Kutas M (2014). Pre-processing in sentence comprehension: Sensitivity to likely upcoming meaning and structure. Language and Linguistics Compass, 8(12), 631–645. doi: 10.1111/lnc3.12093 [PubMed: 27525035]

DeLong KA, Urbach TP, & Kutas M (2005). Probabilistic word pre-activation during language comprehension inferred from brain activity. Nature Neuroscience, 8(8), 1117–1121. doi: 10.1038/nn1504 [PubMed: 16007080]

Dunn DM & Dunn LM (2007). Peabody picture vocabulary test. Minneapolis, MN: NCS Pearson.

Ericsson KA, Charness N, Feltovich PJ, Hoffman RR (Eds). (2006). The Cambridge Handbook of Expertise and Expert Performance. Cambridge, UK: Cambridge University Press.

Federmeier KD & Kutas M (1999). A rose by any other name: Long-term memory structure and sentence processing. Journal of Memory and Language, 41, 469–495. doi: 10.1006/jmla.1999.2660

Federmeier KD, Kutas M, & Schul R (2010). Age-related and individual differences in the use of prediction during language comprehension. Brain and Language, 115(3), 149–161. doi: 10.1016/j.bandl.2010.07.006 [PubMed: 20728207]

Federmeier KD, McLennan DB, De Ochoa E, & Kutas M (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. Psychophysiology, 39, 133–146. doi: 10.1111/1469-8986.3920133 [PubMed: 12212662]

Federmeier KD, Wlotko EW, De Ochoa-Dewald E, & Kutas M (2007). Multiple effects of sentential constraint on word processing. Brain Research, 1146, 75–84. doi: 10.1016/j.brainres.2006.06.101 [PubMed: 16901469]

Filik R & Leuthold H (2013). The role of character-based knowledge in online narrative comprehension: Evidence from eye movements and ERPs. Brain Research, 1506, 94–104. doi: 10.1016/j.brainres.2013.02.017 [PubMed: 23419895]

Friedrich M & Friederici AD (2010). N400-like semantic incongruity effect in 19-month-olds: Processing known words in picture contexts. Journal of Cognitive Neuroscience, 16(8), 1465–1477. doi: 10.1162/0898929042304705

Ganis G, Kutas M, & Sereno MI (1996). The search for “common sense”: An electrophysiological study of the comprehension of words and pictures in reading. Journal of Cognitive Neuroscience, 8(2), 89–106. doi: 10.1162/jocn.1996.8.2.89 [PubMed: 23971417]

Gobbo C & Chi M (1986). How knowledge is structured and used by expert and novice children. Cognitive Development, 1, 221–237. doi: 10.1016/S0885-2014(86)80002-8

Hagoort P, Hald L, Bastiaansen M, Petersson KM (2004). Integration of word meaning and world knowledge in language comprehension. Science, 304, 438–441. doi: 10.1126/science.1095455 [PubMed: 15031438]

Hald LA, Steenback-Planting EG, Hagoort P (2007). The interaction of discourse context and world knowledge in online sentence comprehension. Evidence from the N400. Brain Research, 1146, 210–218. doi: 10.1016/j.brainres.2007.02.054 [PubMed: 17433893]

Heinze H-J, Muente T-F, & Kutas M (1998). Context effects in a category verification task as assessed by event-related brain potential (ERP) measures. Biological Psychology, 47, 121–135. doi: 10.1016/S0301-0511(97)00024-0 [PubMed: 9554184]

Knoeferle P, Urbach TP, & Kutas M (2011). Comprehending how visual context influences incremental sentence processing: Insights from ERPs and picture-sentence verification. Psychophysiology, 48, 495–506. doi: 10.1111/j.1469-8986.2010.01080.x [PubMed: 20701712]

Kuperberg GR (2016). Separate streams or probabilistic inference? What the N400 can tell us about the comprehension of events. Language, Cognition, and Neuroscience, 31(5), 602–616. doi: 10.1080/23273798.2015.1130233

Kutas M (2006). One lesson learned: frame language processing—literal and figurative—as a human brain function. Metaphor and Symbol, 21(4), 285–325. Doi: 10.1207/s15327868ms2104_5

Kutas M & Federmeier KD (2000). Electrophysiology reveals semantic memory use in language comprehension. Trends in Cognitive Sciences, 4(12), 463–470. doi: 10.1016/S1364-6613(00)01560-6 [PubMed: 11115760]

Kutas M & Federmeier KD (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). Annual Review of Psychology, 62(14), 1–27. doi: 10.1146/annurev.psych.093008.131123

Kutas M & Hillyard SA (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. Science, 207, 203–205. doi: 10.1126/science.7350657 [PubMed: 7350657]

Kutas M & Hillyard SA (1984). Brain potentials during reading reflect word expectancy and semantic association. Nature, 307(12), 161–163. doi: 10.1038/307161a0 [PubMed: 6690995]

Kutas M, Lindamood TE, & Hillyard SA (1984). Word expectancy and event-related brain potentials during sentence processing In Kornblum S and Requin J (Eds.), Preparatory States and Processes (pp. 217–237). Hillsdale, New Jersey: Lawrence Erlbaum.

Kutas M, Van Petten C, & Besson M (1988). Event-related potential asymmetries during the reading of sentences. Electroencephalography and Clinical Neurophysiology, 69, 218–233. doi: 10.1016/0013-4694(88)90131-9 [PubMed: 2450003]

Kutner MH, Nachtsheim CJ, Neter J, & Li W (2005). Applied Linear Statistical Models, Fifth *Edition* McGraw-Hill Irwin.

Kuznetsova A, Brockhoff PB, & Christensen RHB (2017). lmerTest Package: Tests in linear mixed effects models. Journal of Statistical Software, 82(13), 1–26. doi: 10.18637/jss.v082.i13

McLaughlin J, Osterhout L, & Kim A (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. Nature Neuroscience, 7(7), 703–704. doi: 10.1038/nn1264 [PubMed: 15195094]

McLaughlin J, Tanner D, Pitkänen I, Frenck-Mestre C, Inoue K, Valentine G, & Osterhout L (2010). Brain potentials reveal discrete stages of L2 grammatical learning. Language Learning, 60, 123–150. doi: 10.1111/j.1467-9922.2010.00604.x

Metusalem R, Kutas M, Urbach TP, Hare M, McRae K, & Elman JL (2012). Generalized event knowledge activation during online sentence comprehension. Journal of Memory and Language, 66(4), 545–567. doi: 10.1016/j.jml.2012.01.001 [PubMed: 22711976]

Miller GA (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63(2), 81–97. doi: 10.1037/h0043158 [PubMed: 13310704]

Münte TF, Schiltz K, & Kutas M (1998). When temporal terms belie conceptual order. Letters to Nature, 395, 71–73. doi: 10.1038/25731

Nakano H, Saron C, & Swaab T (2010). Speech and span: Working memory capacity impacts the use of animacy but not of world knowledge during spoken sentence comprehension. Journal of Cognitive Neuroscience, 22(12), 2886–2898. doi: 10.1162/jocn.2009.21400 [PubMed: 19929760]

Nieuwland MS & Van Berkum JJA (2006). When peanuts fall in love: N400 evidence for the power of discourse. Journal of Cognitive Neuroscience, 18(7), 1098–1111. doi: 10.1162/jocn.2006.18.7.1098 [PubMed: 16839284]

Pakulak E & Neville H (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. Journal of Cognitive Neuroscience, 22(12), 2728–2744. doi: 10.1162/jocn.2009.21393 [PubMed: 19925188]

Pearson PD, Hansen J, & Gordon C (1979). The effect of background knowledge on young children's comprehension of explicit and implicit information. Journal of Reading Behavior, 11(3), 201–209.

Rugg MD & Curran T (2007). Event-related potentials and recognition memory. TRENDS in Cognitive Science, 11(6), 251–257. doi: 10.1016/j.tics.2007.04.004

Simon HA (1974). How big is a chunk? Science, 183(4124), 428–488. [PubMed: 4358075]

Simon HA & Chase WG (1973). Skill in chess. American Scientist, 61(4), 394–403. doi: 10.1007/978-1-4757-1968-0_18

Simonson R & Walker S (1988). Multi-cultural literacy. Gray Wolf Press: Saint Paul.

Spilich GJ, Vesonder GT, Chiesi HL, & Voss JF (1979). Text processing of domain-related information for individuals with high and low domain knowledge. Journal of Verbal Learning and Verbal Behavior, 18, 275–290. doi: 10.1016/S0022-5371(79)90155-5

Stanovich KE & Cunningham AE (1993). Where does knowledge come from? Specific associations between print exposure and information acquisition. Journal of Educational Psychology, 85(2), 211–229. doi: 10.1037/0022-0663.85.2.211

Stanovich KE & West RF (1989). Exposure to print and orthographic processing. Reading Research Quarterly, 24(4), 402–433. doi: 10.2307/747605

Tanenhaus MK, Spivey-Knowlton MJ, Eberhard KM, & Sedivy JC (1995). Integration of visual and linguistic information in spoken language comprehension. Science, 268, 1632–1634. doi: 10.1126/science.7777863 [PubMed: 7777863]

Tanner D, Inoue K, & Osterhout L (2014). Brain-based individual differences in online L2 grammatical comprehension. Bilingualism: Language and Cognition, 17(2), 277–293. doi: 10.1017/S1366728913000370

Thornhill DE & Van Petten C (2012). Lexical versus conceptual anticipation during sentence processing: Frontal positivity and N400 ERP components. International Journal of Psychophysiology, 83, 382–392. doi: 10.1016/j.ijpsycho.2011.12.007 [PubMed: 22226800]

Van Berkum JJA, van den Brink D, Tesink CMJY, Kos M, & Hagoort P (2008). The neural integration of speaker and message. Journal of Cognitive Neuroscience, 20(4), 580–591. doi: 10.1162/jocn.2008.20054 [PubMed: 18052777]

Van Petten C & Kutas M (1990). Interactions between sentence context and word frequency in event-related brain potentials. Memory and Cognition, 18(4), 380–393. doi: 10.3758/BF03197127 [PubMed: 2381317]

Van Petten C & Luka BJ (2012). Prediction during language comprehension: Benefits, costs, and ERP components. International Journal of Psychophysiology, 83,176–190. doi: 10.1016/j.ijpsycho.2011.09.015 [PubMed: 22019481]

Verhagen V, Mos M, Backus A, & Schilperoord J (2018). Predictive language processing revealing usage-based variation. Language and Cognition, 10, 329–373. doi: 10.1017/langcog.2018.4

Wlotko EW & Federmeier KD (2012). So that's what you meant! Event-related potentials reveal multiple aspects of context use during construction of message-level meaning. NeuroImage, 62, 356–366. doi: 10.1016/j.neuroimage.2012.04.054 [PubMed: 22565202]

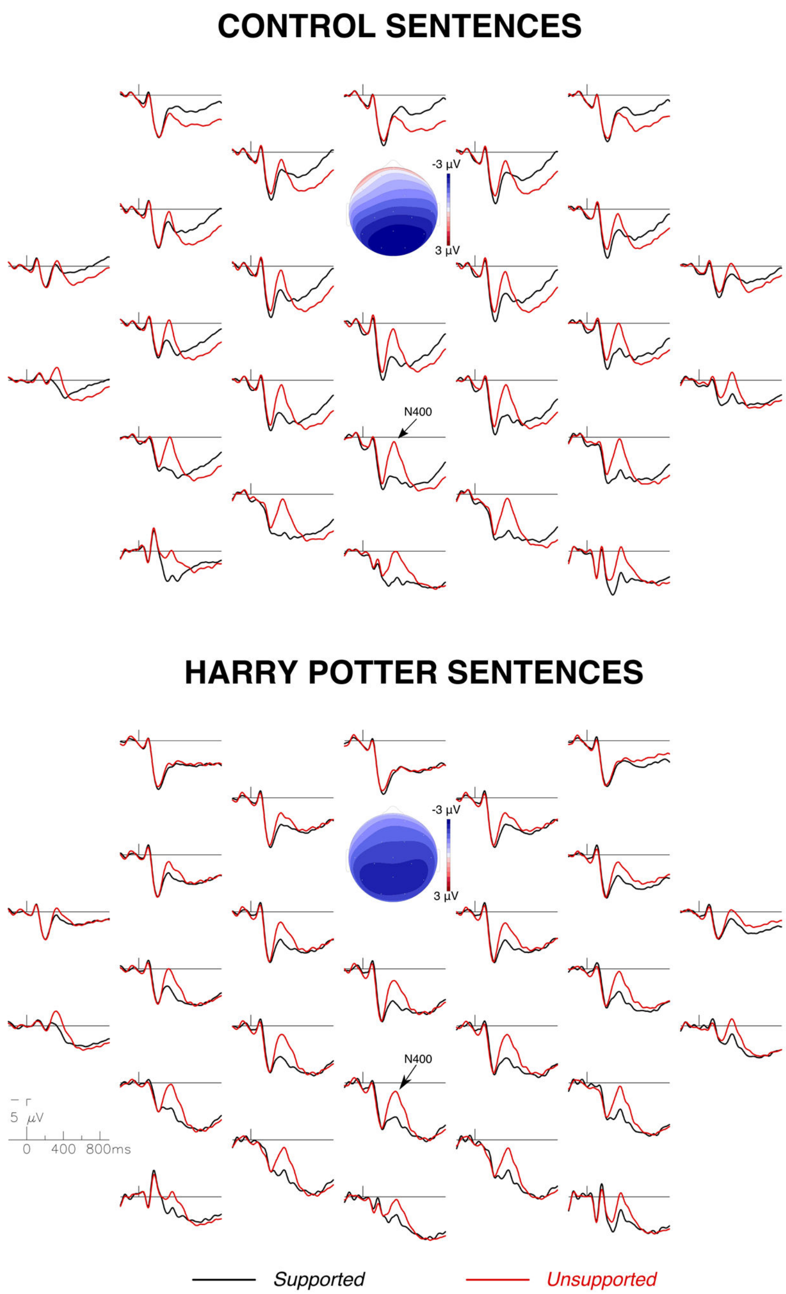
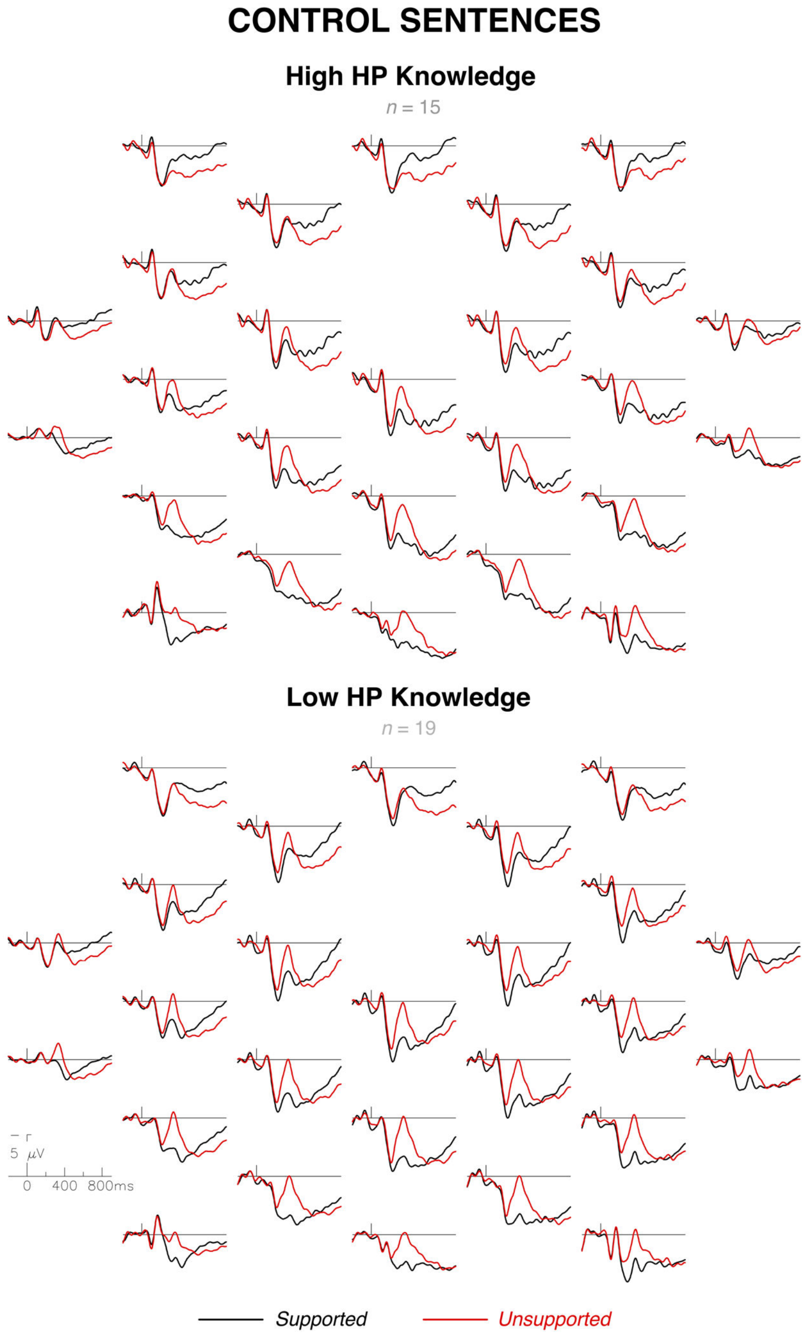


Figure 1.
Grand average ERPs across all participants for critical words of each type (supported, unsupported) for control and HP sentences. Topographical scalp plots show the N400 effect of contextual support (unsupported minus supported) from 250 to 500 ms.


CONTROL SENTENCES
High HP Knowledge
n = 15
Low HP Knowledge
n = 19
5 µV
0 400 800ms
Supported
Unsupported

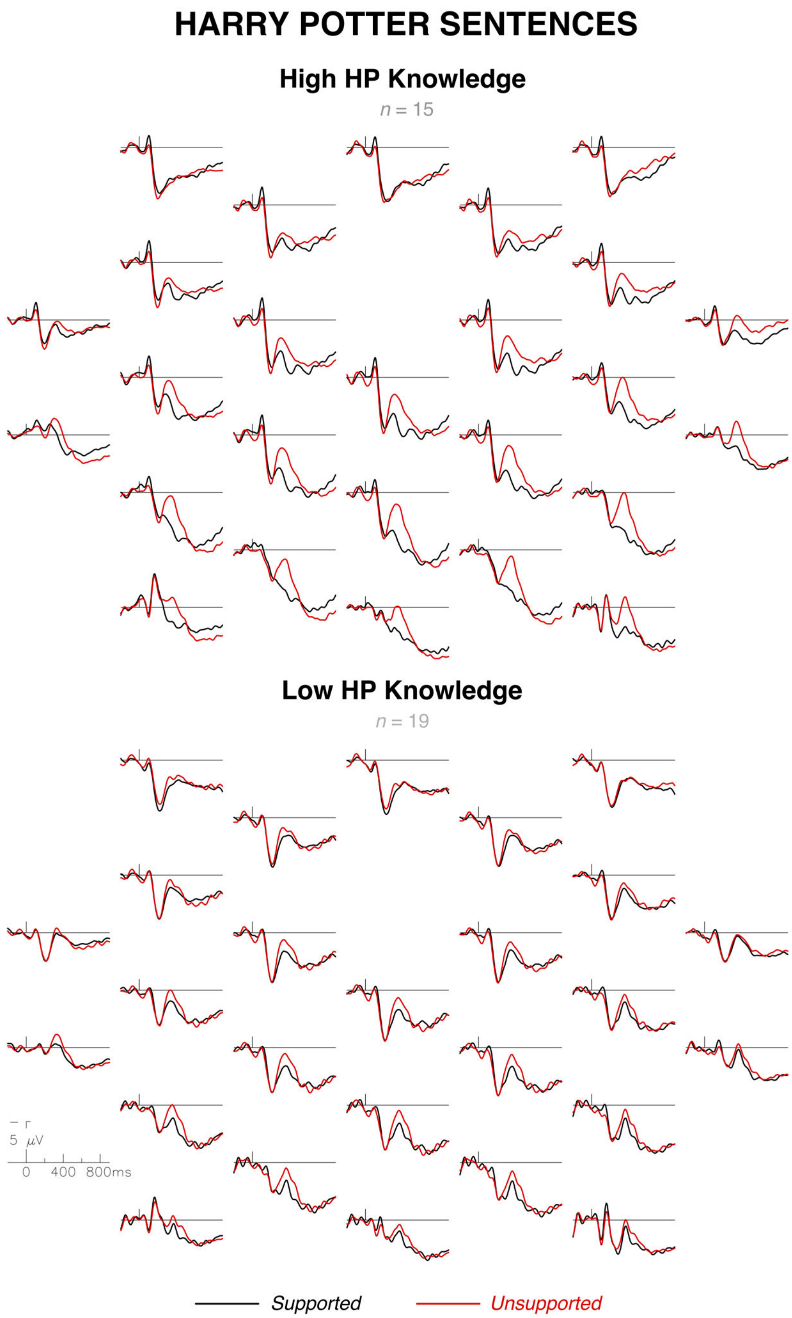


Figure 2.
a. Grand average ERPs for control sentences for high HP knowledge and low HP knowledge groups, assigned by a median (= 6) split on HP quiz score.
b. Grand average ERPs for HP sentences for HP knowledge and low HP knowledge groups, assigned by a median (= 6) split on HP quiz score.

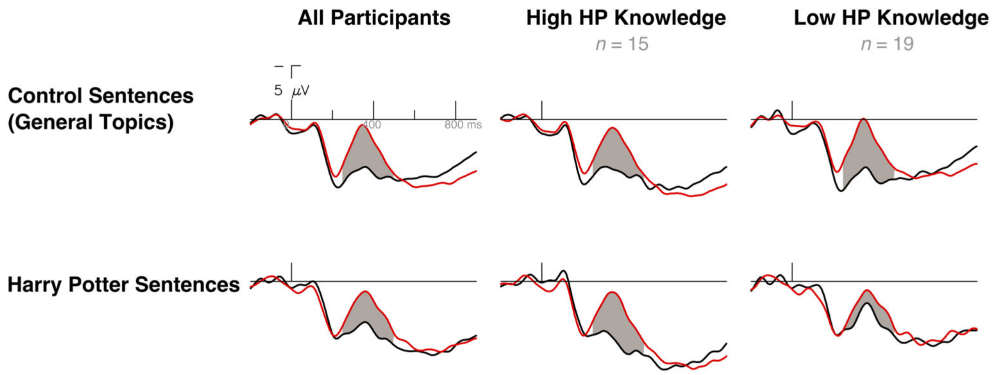


Figure 3.

ERPs from a centro-parietal ROI for supported (black) and unsupported (red) endings to HP sentences. Shaded region from 250 to 500 ms shows N400 effect.

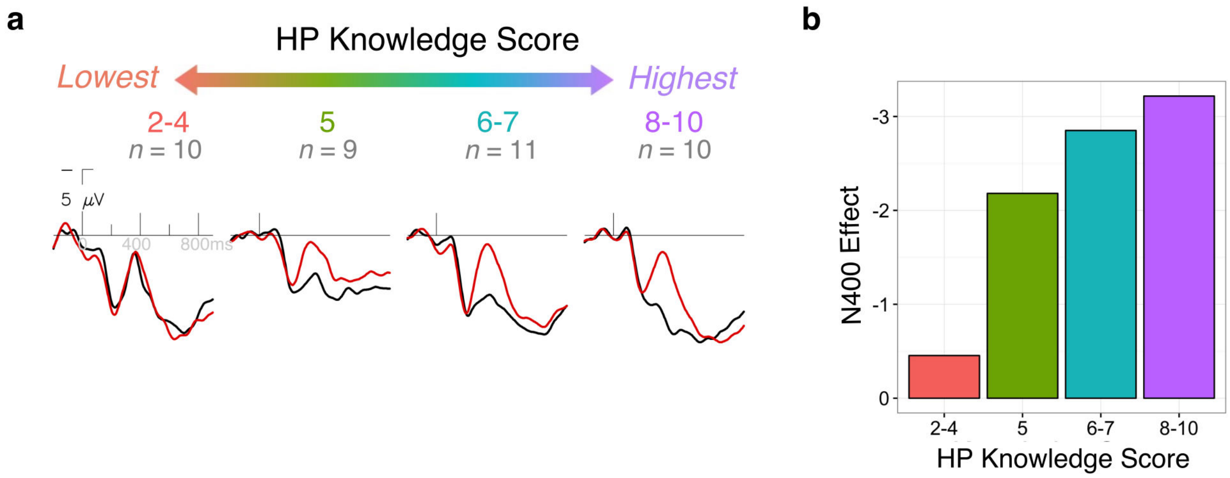


Figure 4.
(a) ERPs from a centro-parietal ROI for supported (black) and unsupported (red) endings to HP sentences by HP knowledge subgroup.
(b) N400 effect (unsupported minus supported) for each HP knowledge subgroup.

Table 1.

Sample experimental stimuli.

Control Sentences		
Sentence frame	**Supported**	**Unsupported**
We had been watching the blue jay for days. The bird laid her eggs in the	nest	yard
The vampire moved in. He bit his victim on the	neck	shoulder
Alicia's first client was a failure. But her second was a	success	triumph

Harry Potter Sentences		
Sentence frame	**Supported**	**Unsupported**
The character Peter Pettigrew changes his shape at times. He takes the form of a	rat	dog
There are two Beaters on every Quidditch team. Their job is to protect their team from	Bludgers	Spellotape
Wizards are able to conjure the Dark Mark. They can use a spell called	Morsmordre	Stupefy

Table 2.

Mean, standard deviation, and range are provided for behavioural measures of individual differences.

	All participants			**High HP group**			**Low HP group**		
	Mean	SD	Range	Mean	SD	Range	Mean	SD	Range
HP Quiz	6.12	(2.33)	[2, 10]	8.67	(1.35)	[7, 10]	4.16	(0.96)	[2, 5]
# of HP Books	2.11	(3.75)	[0, 18]	4.70	(5.22)	[0.04, 18]	0.46	(0.55)	[0, 1.71]
HP Self Score	43.54	(41.32)	[4, 181.5]	75.92	(51.10)	[14, 181.5]	19.84	(11.07)	[4, 47]
ART	0.23	0.12	[0.02, 0.5]	0.29	0.10	[.15, .5]	0.17	0.11	[.02, .42]
MRT	0.22	0.08	[0.08, 0.38]	0.26	0.06	[.15, .38]	0.18	0.08	[.08, .35]
# of Authors Listed	2.60	(1.72)	[0, 5]	3.40	(1.18)	[1, 5]	1.74	(1.69)	[0, 5]
MRH Total	6.67	(3.08)	[0, 15]	7.67	(2.94)	[4, 15]	6.00	(3.06)	[0, 12]
GKQ	19.51	(3.62)	[10, 25]	21.33	(3.09)	[15, 25]	17.74	(3.41)	[10, 23]
CLC	0.35	(0.13)	[.07, .60]	0.41	(0.08)	[0.31, 0.60]	0.29	(0.13)	[0.07, 0.57]
MCLC	0.44	(0.18)	[0, .77]	0.51	(0.14)	[0.13, 0.77]	0.38	(0.20)	[0, 0.70]
PPVT	208.32	(6.70)	[195, 219]	212.07	(5.09)	[200, 219]	205.05	(6.17)	[196, 217]
Sentence Span	2.95	(0.64)	[1.5, 5]	2.77	(0.56)	[1.5, 3.5]	3.08	(0.75)	[2, 5]

Table 3.

Intercorrelations (Pearson's *r*) among behavioral measures of individual differences. *r* values above .31 are significant at $a = .05$; *r* values above .403 are significant at $a = .01$.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 HP Quiz	1	0.67	0.73	0.47	0.51	0.45	0.15	0.43	0.39	0.38	0.48	−0.13	0.54	0.47
2 HP Books	-	1	0.85	0.34	0.32	0.38	0.07	0.23	0.33	0.37	0.18	−0.04	0.38	0.36
3 HP Self Score	-	-	1	0.43	0.34	0.46	0.12	0.25	0.31	0.31	0.21	−0.12	0.46	0.34
4 ART	-	-	-	1	0.54	0.38	0.45	0.42	0.67	0.54	0.30	0.13	0.81	0.64
5 MRT	-	-	-	-	1	0.25	0.22	0.59	0.57	0.55	0.65	0.04	0.69	0.67
6 Authors Listed	-	-	-	-	-	1	0.42	0.49	0.5	0.12	0.23	0.12	0.7	0.43
7 MRH Total	-	-	-	-	-	-	1	0.31	0.42	0.13	0.2	0.03	0.72	0.33
8 GKQ	-	-	-	-	-	-	-	1	0.69	0.5	0.72	0.19	0.62	0.85
9 CLC	-	-	-	-	-	-	-	-	1	0.63	0.5	0.3	0.74	0.9
10 MCLC	-	-	-	-	-	-	-	-	-	1	0.5	0.23	0.46	0.83
11 PPVT	-	-	-	-	-	-	-	-	-	-	1	0.12	0.47	0.67
12 Sentence Span	-	-	-	-	-	-	-	-	-	-	-	1	0.11	0.28
13 Reading Experience	-	-	-	-	-	-	-	-	-	-	-	-	1	0.71
14 General Knowledge	-	-	-	-	-	-	-	-	-	-	-	-	-	1

Table 4.

Whole-head ANOVA results for N400 and late positivity time windows.

	DF	*F*	p-value	ε_{GG}
N400				
List	(1, 38)	0.027	.8699	
Electrode	(25, 950)	26.280	.0000	0.148
Ending Type	(1, 38)	42.625	.0000	
Sentence Type	(1, 38)	0.185	.6692	
List:Electrode	(25, 950)	0.343	.8345	0.148
List:Ending Type	(1, 38)	0.714	.4033	
List:Sentence Type	(1, 38)	0.276	.6021	
Electrode:Ending Type	(25, 950)	35.693	.0000	0.111
Electrode:Sentence Type	(25, 950)	5.347	.0004	0.168
Ending Type:Sentence Type	(1, 38)	0.307	.5828	
List:Electrode:Ending Type	(25, 950)	2.795	.0477	0.111
List:Electrode:Sentence Type	(25, 950)	0.382	.8306	0.168
List:Ending Type:Sentence Type	(1, 38)	9.470	.0039	
Electrode:Ending Type:Sentence Type	(25, 950)	7.704	.0001	0.126
List:Electrode:Ending Type:Sentence Type	(25, 950)	4.150	.0068	0.126
Late positivity				
List	(1, 38)	0.014	.9071	
Electrode	(25, 950)	27.437	.0000	0.115
Ending Type	(1, 38)	3.649	.0637	
Sentence Type	(1, 38)	4.252	.0461	
List:Electrode	(25, 950)	0.164	.9140	0.115
List:Ending Type	(1, 38)	1.819	.1854	
List:Sentence Type	(1, 38)	1.408	.2427	
Electrode:Ending Type	(25, 950)	3.463	.0237	0.106
Electrode:Sentence Type	(25, 950)	3.277	.0161	0.146
Ending Type:Sentence Type	(1, 38)	11.169	.0019	
List:Electrode:Ending Type	(25, 950)	1.456	.2343	0.106
List:Electrode:Sentence Type	(25, 950)	0.895	.4613	0.146
List:Ending Type:Sentence Type	(1, 38)	2.762	.1048	
Electrode:Ending Type:Sentence Type	(25, 950)	8.032	.0000	0.147
List:Electrode:Ending Type:Sentence Type	(25, 950)	3.085	.0211	0.147

Table 5.

ROI analyses using linear mixed-effects models for N400 and late positivity time periods.

	Estimate	SE	DF	T-value	Pr(>|t|)
N400					
Intercept	3.515	0.322	38	10.914	.0000
HP knowledge	0.465	0.323	38	1.442	.1574
Sentence type	0.112	0.125	114	0.894	.3734
Ending type	1.167	0.125	114	9.309	.0000
HP knowledge:sentence type	−0.264	0.126	114	−2.099	.0380
HP knowledge:Ending type	0.170	0.126	114	1.352	.1791
Sentence type:Ending type	0.067	0.125	114	0.531	.5964
HP knowledge:Sentence type:Ending type	−0.255	0.126	114	−2.032	.0445
Late positivity					
Intercept	5.619	0.425	38	13.214	.0000
HP knowledge	0.625	0.426	38	1.466	.1508
Sentence type	−0.393	0.147	114	−2.675	.0086
Ending type	−0.128	0.147	114	−0.869	.3864
HP knowledge:Sentence type	−0.166	0.147	114	−1.131	.2606
HP knowledge:Ending type	0.021	0.147	114	0.142	.8871
Sentence type:Ending type	−0.202	0.147	114	−1.374	.1721
HP knowledge:Sentence type:Ending type	0.021	0.147	114	0.139	.8894

Table 6.

Follow-up ROI analyses using linear mixed-effects for HP and Control sentences during N400 time period.

	Estimate	SE	DF	T-value	Pr(>|t|)
Control sentences					
Intercept	3.627	0.358	38	10.144	.0000
HP knowledge	0.202	0.358	38	0.563	0.577
Ending type	1.233	0.175	38	7.046	.0000
HP knowledge:Ending type	−0.085	0.175	38	−0.487	0.629
HP Sentences					
Intercept	3.403	0.333	38	10.220	.0000
HP knowledge	0.729	0.334	38	2.185	.0351
Ending type	1.100	0.180	38	6.117	.0000
HP knowledge:Ending type	0.425	0.180	38	2.358	.0236