HHS Public Access

Author manuscript

Neuroscience. Author manuscript; available in PMC 2022 October 07.

Published in final edited form as:

Neuroscience. 2022 October 01; 501: 143–158. doi:10.1016/j.neuroscience.2022.07.030.

Evoking the N400 Event-related Potential (ERP) Component Using a Publicly Available Novel Set of Sentences with Semantically Incongruent or Congruent Eggplants (Endings)

Kathryn K. Toffolo[*],

Edward G. Freedman[*],

John J. Foxe[*]

The Frederick J. and Marion A. Schindler Cognitive Neurophysiology Laboratory, The Del Monte Institute for Neuroscience, Department of Neuroscience, University of Rochester School of Medicine and Dentistry, Rochester, NY 14620, USA

Abstract

During speech comprehension, the ongoing context of a sentence is used to predict sentence outcome by limiting subsequent word likelihood. Neurophysiologically, violations of context-dependent predictions result in amplitude modulations of the N400 event-related potential (ERP) component. While N400 is widely used to measure semantic processing and integration, no publicly-available auditory stimulus set is available to standardize approaches across the field. Here, we developed an auditory stimulus set of 442 sentences that utilized the semantic anomaly paradigm, provided cloze probability for all stimuli, and was developed for both children and adults. With 20 neurotypical adults, we validated that this set elicits robust N400′s, as well as two additional semantically-related ERP components: the recognition potential (~250 ms) and the late positivity component (~600 ms). This stimulus set (https://doi.org/10.5061/dryad.9ghx3ffkg) and the 20 high-density (128-channel) electrophysiological datasets (https://doi.org/10.5061/dryad.6wwpzgmx4) are made publicly available to promote data sharing and reuse. Future studies that use this stimulus set to investigate sentential semantic comprehension in both control and clinical populations may benefit from the increased comparability and reproducibility within this field of research.

[*]Corresponding authors. Kathryn_toffolo@urmc.rochester.edu (K. K. Toffolo), Ed_Freedman@urmc.rochester.edu (E. G. Freedman), John_Foxe@urmc.rochester.edu (J. J. Foxe).

AUTHORS' CONTRIBUTIONS

All authors collectively conceived of the study. KKT created the stimulus set, recruited participants, and collected the data. KKT analyzed the data in consultation with EGF and JJF, and produced the figures for this study. KKT wrote the first draft of the manuscript and all authors provided editorial input on subsequent drafts. All authors read and approved the final version of this manuscript.

ETHICS AND CONSENT

The Research Subjects Review Board of the University of Rochester approved all the experimental procedures (STUDY00002036). Each participant provided written informed consent in accordance with the tenets laid out in the Declaration of Helsinki.

COMPETING INTERESTS

The Authors report no biomedical financial interests or potential conflicts of interest.

APPENDIX A.: SUPPLEMENTARY DATA

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroscience.2022.07.030.

Keywords

Late Positivity Component; P600; N400; Recognition Potential; Cloze Probability; EEG

INTRODUCTION

The ability to use prior context to predict a particular sentence ending is thought to facilitate language acquisition and fluid language comprehension (Harris, 1975; Pickering and Garrod, 2004; Leech et al., 2007), although the precise role that prediction plays in language comprehension and its degree of influence continues to be a matter of vigorous debate (Hickok, 2013; Huettig, 2015; Huettig and Mani, 2016; Kuperberg and Jaeger, 2016). In essence, sentence outcome can be predicted because the words already heard limit the words that can sensibly terminate the sentence. The relationship between prior context and the target word is quantified by 'cloze probability' (CP) (Kutas and Hillyard, 1984, 1980), which is a measure of how well a predicted word matches the context. For example, in the sentence '*Finally, the climbers reached the top of the* –' (Pijnacker et al., 2010), one might reasonably predict that the final word is '*mountain*' or '*hill*'. In this example, mountain has a high association with and is semantically congruent with the prior sentence context and would have a high CP. 'Hill' is also a reasonable ending to the sentence, but perhaps a little less likely, and as such would have a lower CP. Conversely, the word 'tulip' would be a very unexpected ending to this sentence, is semantically incongruent and would be associated with an extremely low CP. Prior analysis of electroencephalography (EEG) has identified a broad negative event related potential (ERP) that peaks around 400 ms after the presentation of semantically related or unrelated stimuli (Kutas and Hillyard, 1984, 1980; Osterhout and Holcomb, 1992; Luck, 2005; Lau et al., 2008). This N400 ERP component is evoked by both congruent and incongruent endings, but when incongruent stimuli are presented, the amplitude of the N400 component is substantially increased (Kutas and Hillyard, 1984, 1980; Osterhout and Holcomb, 1992; Luck, 2005; Lau et al., 2008; Block and Baldwin, 2010). A variety of methodological approaches using single words/pictures, sounds, or sentences, have successfully evoked an N400, resulting in many theories of what this component reflects, such as semantic processing or memory retrieval (Kaan, 2007; Lau et al., 2008). Nonetheless, the N400 ERP component is considered an established index of semantic integration (Kutas and Hillyard, 1980; Osterhout and Holcomb, 1992; Brown and Hagoort, 1993; Connolly and Phillips, 1994; Kaan, 2007; Hagoort, 2008; Lau et al., 2008; Brouwer et al., 2017). The processing of congruent stimuli is postulated to be easier to integrate with prior context because they follow ones narrowing of expectations, whereas incongruent stimuli result in larger N400 amplitudes due to the increased difficulty of integrating unexpected words (Kutas and Hillyard, 1984; Connolly and Phillips, 1994; St. George et al., 1994; Federmeier and Kutas, 1999; Van Berkum et al., 1999; Kutas and Federmeier, 2000; Federmeier, 2007; Kaan, 2007; Lau et al., 2008; Fitz and Chang, 2019). As such, the N400 component has proven to be a very valuable tool for language research.

Two other ERP components reported to be sensitive to semantic congruence are the recognition potential (RP) and the late positivity component (LPC). The RP is an N2-like component (Dien et al., 2003; Proverbio and Riva, 2009; Fernandez and Smith Cairns, 2011) that occurs over parietal-occipital scalp regions (Proverbio and Riva, 2009; Martín-Loeches et al., 2004) and is thought to be generated in the left fusiform gyrus or visual word form area (VWFA) (Dien et al., 2003; Martín-Loeches, 2007). The RP shows larger amplitudes in response to congruent stimuli (Martín-Loeches et al., 2004; Proverbio and Riva, 2009) and as such, has been generally related to expectancy (Dien et al., 2003) and contextual semantic processing (Martín-Loeches, 2007). More specifically, the RP has been suggested to represent the point at which an individual recognizes the target word following congruent context (Fernandez and Smith Cairns, 2011). The terms LPC and P600 have tended to be used interchangeably in the literature, with LPC being used most often when the work is specifically addressing responses to semantic contexts (Kaan et al., 2000; Hoeks et al., 2004; Kim and Osterhout, 2005; Luck, 2005; Van Herten et al., 2005; Wang et al., 2009; Gouvea et al., 2010; Friederici, 2011, 2002; Brouwer et al., 2017; Leckey and Federmeier, 2020). The LPC has a more positive amplitude in response to incongruent stimuli (Wang et al., 2009) and has been related to semantic difficulty or modulations (Kaan et al., 2000; Van Herten et al., 2005; Wang et al., 2009; Gouvea et al., 2010; Brouwer et al., 2017). The LPC is thought to reflect the reanalysis of semantic information because it occurs after the N400 (Osterhout and Holcomb, 1992; Kim and Osterhout, 2005; Braeutigam et al., 2008; Friederici, 2011). Studies that have used the P600 nomenclature have shown that it is also sensitive to syntactic context (Osterhout and Holcomb, 1992; Kaan et al., 2000; Frisch et al., 2002; Hahne et al., 2004; Kim and Osterhout, 2005; Gouvea et al., 2010; Schacht et al., 2014; Brouwer et al., 2017), whereas it has been argued that the amplitude of the LPC is related to the integration of both semantic and syntactic information (Ainsworth-Darnell et al., 1998; Friederici, 2002; Van Herten et al., 2005; Wang et al., 2009; Gouvea et al., 2010; Brouwer and Hoeks, 2013; Schacht et al., 2014; Brouwer et al., 2017; Fitz and Chang, 2019). There is no consensus on the interpretation of the P600 ERP as a whole, nor a consensus on its topographic distribution (Kaan et al., 2000; Frisch et al., 2002; Van Herten et al., 2005; Kuperberg, 2007; Friederici, 2011; Brouwer et al., 2017; Leckey and Federmeier, 2020). To summarize, in response to semantically incongruent/unexpected words, a recognition potential (RP ~ 250 ms) is not expected be present, and both the N400 (~400 ms) and late positivity component (LPC ~ 600 ms) should be larger relative to semantically congruent/expected words.

Semantic studies have used various task designs to understand the nuanced linguistic sensitivities of the N400 and LPC. This has resulted in task dependent topographic differences and various interpretations of the relationship that these components have with prediction, semantic processing, and language comprehension (Kotz and Friederici, 2003; Kim and Osterhout, 2005; Van Herten et al., 2005; Kutas and Van Petten, 2006; Van Petten and Luka, 2006; Kuperberg, 2007; Hagoort, 2008; Lau et al., 2008; Friederici, 2011, 2002; Brouwer et al., 2017; Leckey and Federmeier, 2020). It goes without saying that it is important to use diverse methodologies to understand the linguistic sensitivity of these components. However, the lack of a standardized task design can result in conflicting conclusions about the extent of one's semantic ability, especially when investigating atypical populations such as Autism Spectrum Disorder (ASD) (Osterhout and Holcomb, 1992; Kaan et al., 2000; Kamio and Toichi, 2000; Braeutigam et al., 2008; Wang et al., 2009; Gouvea et al., 2010; McCleery et al., 2010; Pijnacker et al., 2010; Friederici, 2011; Ribeiro et al., 2013; Harper-Hill et al., 2014; Fernandes et al., 2016; Brouwer et al., 2017; Coderre et al., 2017; DiStefano et al., 2019). Although a CP validated, written N400 stimulus set for adults is publicly available (Block and Baldwin, 2010), to our knowledge, a CP validated, auditory and written N400 stimulus set meant for children and adults does not exist. The purpose of this study was to create a stimulus set that targeted auditory, sentential semantic comprehension using a semantic anomaly N400 task design. This stimulus set measured the predictability of each sentence via a CP survey and could be used for children as well as adults because sentences were created for individuals of at least 5 years old. Considering that we provided both the written words and non-prosodic audio files for our sentences, this stimulus set could be used to investigate both reading and auditory semantic processing of neurotypical (NT) populations and populations known for disorders that impact language and communication skills, such as ASD. Furthermore, the auditory stimulus set specifically may be useful for researchers wishing to study semantic processing in those with intellectual and developmental disabilities, where reading skills are often poor or absent entirely (e.g. dyslexia, lower functioning ASD) or in a number of rare neurodevelopmental diseases (e.g. Batten disease or Rett syndrome) where the children have severe progressive vision loss or simply cannot fixate on the screen. This stimulus set and the current dataset are made publicly available. Future semantic studies that use this stimulus set may benefit from the increased comparability and reproducibility within this field of research.

EXPERIMENTAL PROCEDURES

Participants

Twenty-four neurotypical adults were recruited and provided written informed consent to participate in this study. Four subjects were excluded from data analysis due to failure to remain alert or to sit still during data collection (n = 2), or due to noisy EEG data (n = 2) (Fig. S1). The remaining participants make up the fully-analyzed dataset. These twenty subjects ranged in age from 18 to 35 (mean age = 25.5 ± 4.36), nine were female, and three were left-handed. Every participant spoke English as their first language, and twelve participants were mono-lingual while eight participants reported being bi- or multi-lingual. Demographic information for all participants, including those removed from further analysis are reported in Table 1.

Stimuli

The semantic anomaly paradigm consisted of 221 sentence pairs with incongruent and congruent endings, for a total of 442 stimuli in this stimulus set. However, twenty sentence pairs were eliminated before analysis because their endings did not match in syllable number, contained hyphenated phrases, cultural references, or upon closer examination, the supposed incongruent endings were in fact congruent. The remaining 402 stimuli (201 congruent and 201 incongruent sentence pairs) ranged from four to eight words in length. Sentences were created using simple words, derived from a set of age-appropriate words contained in the Medical Research Council Psycholinguistic (MRCP) database (https://websites.psychology.uwa.edu.au/school/mrcdatabase/uwa_mrc.htm). For example, from the word “cake”, the congruent sentence “He baked a birthday cake” was created. The words selected from the MRCP database took into consideration the age of acquisition (from a database provided by (Gilhooly and Logie, 1980)) and/or written word frequency (from a normed written word frequency set (Francis and Kucera, 1967)). This was to ensure that each sentence could be readily comprehended by NT individuals aged 5 years or older. The majority of these age-appropriate words were monosyllabic, but a few basic words with 2-syllables (e.g. present) and 3-syllables (e.g. animal) were included, due to their high-written-word frequency or early age of acquisition. Similar to previous studies using the MRCP database (Ross et al., 2007; 2011), we were sensitive to the fact that everyday language use has changed since the creation of these sets, so we carefully selected words that are still in common use.

This stimulus set included sentence pairs where the incongruent endings were just semantic errors. After elimination, this amounts to 132 / 201 stimulus pairs including example stimuli. These incongruent endings were matched to their congruent pair in word type (e.g. noun or verb) and number (e.g. plural or singular). There were also sentence pairs where incongruent endings contained both semantic and syntactic errors. After elimination, this amounts to 69 / 201 stimulus pairs. Semantically incongruent endings were also classified as syntactic errors if the ending deviated from the syntax provided by the sentential context. The type of syntactic errors included: 1. Endings where a plural noun expectation was deviated with a singular noun or vice versa (27 stimulus pairs); 2. Endings where an adjective expectation was deviated with a noun or vice versa (20 stimulus pairs); and 3. Endings where a verb expectation was deviated with a noun or vice versa (22 stimulus pairs). During final analysis, these 3 types of syntactic errors were combined into a single linguistic division (LD) in order to compare the overall response to sentences with both semantic and syntactic errors, and the response to sentences with only semantic errors.

Language comprehension in NT individuals is highly influenced by communicative cues such as prosody, especially when sentence meaning relies on syntactic prosodic cues (Cutler et al., 1997; Frazier et al., 2006; Dahan, 2015; Thorson 2018). The intention of this manuscript was to create a stimulus set that could be used for all populations equally. Therefore, this stimulus set was constructed without prosody to ensure that NT individuals would not have an advantage in language comprehension over other populations known to have difficulty with communication or prosody, such as those with ASD (McCann et al., 2007; DePape et al., 2012; Eigsti et al., 2012; O'Connor, 2012; Wang and Tsao, 2015; Martzoukou et al., 2017) or schizophrenia (Leitman et al., 2005, 2007). To do so, individual words from the word list were recorded from a female speaker, instructed to voice the words with minimal inflection, stress, and intonation (i.e. in a monotonous non-prosodic manner). Words were then compiled into complete sentences using the Audacity Software (Version 3.0.0. Audacity® software is copyright © 1999–2021 Audacity Team. https://audacityteam.org/). These artificially-compiled sentences were manually adjusted to have similar pitch-frequency between each word within a sentence and between all sentences. Concurrently, time gaps of 10 ms were added between words so that all sentences would have similar pacing and that researchers would be able to trigger discretely on each word in future work. Both artificial timing and frequency add to the robotic nature of the stimuli. A future initiative will add the prosodic versions of these sentences to this public stimulus set so that researchers can explore more communicative aspects of language.

Procedure

Participants were fit with a 128-electrode cap (Bio Semi B. V. Amsterdam, the Netherlands) and seated in a sound attenuating, electrically shielded booth (Industrial Acoustics Company, The Bronx, NY) with a computer monitor (Acer Predator Z35 Curved HD, Acer Inc.) and a standard keyboard (Dell Inc.). The task was created with Presentation® Software (Version 18.0, Neurobehavioral Systems, Inc. Berkeley, CA). The task was first explained to the participant during the consent process and then again before the experimental session. Individuals were asked to refrain from excessive movement and to focus on a fixation cross throughout the task in order to reduce movement artifacts. The experimental session began by explaining the task for a third time. All instructions were presented both visually on the screen and auditorily through the headphones (Sennheiser electronic GmbH & Co. KG, USA). Instructions were followed by two practice trials which were the same for every participant. Feedback was given about a participant's response only during practice trials and not during experimental trials. Trials were presented as follows: 1. A fixation cross was on the screen while an auditory sentence stimulus was presented through headphones; 2. A two second pause; and 3. A question (presented both visually and auditorily) asked the participant if the sentence ended as expected, where subjects responded with a right or left arrow key when sentences ended as expected (congruent) or unexpected (incongruent) respectively, to end the trial. A two-second delay was inserted between a subject's response and the start of the next sentence. A total of 442 stimuli were presented to participants in the same order. This was done to ensure that every participant had the same experience throughout the task for every sentence. Two of these stimuli (one congruent and one incongruent) were used for example trials. The remaining 440 were used for the experiment. Stimuli were separated into 11 blocks with optional breaks between each block. Participants could continue onto the next block by pressing the spacebar. After elimination, the responses to 400 out of the 440 stimuli contributed to the analysis of this experiment.

Data preprocessing

Data were digitized online at a rate of 512 Hz, DC to 150 Hz pass-band, and referenced to the common mode sense (CMS) active electrode. EEG data were preprocessed and analyzed offline using in-house scripts leveraging EEGLAB functions (Delorme and Makeig, 2004). Data were filtered using a Chebyshev II spectral filter with a band pass of 0.1–45 Hz. Channels were rejected automatically if recorded data at that site exceeded more than 3 standard deviations from the mean variance and amplitude from all neighboring electrodes. These channels were interpolated using EEGLAB spherical interpolation. Data were then re-referenced to the common average. Prior to analysis, the time from the beginning of a sentence to the onset of the last word was measured for each sentence using Praat software (PRAAT v. 6.1, University of Amsterdam, the Netherlands). These measures were used to adjust the time stamp of each stimulus, so that the data could be precisely aligned to the onset of the last word (i.e. 0 ms), rather than the beginning of the sentence. For all participants, epochs from −200 to 1000 ms were created using a baseline of the 200 ms interval before the onset of the last word. Trials were rejected automatically based on an artifact rejection threshold of 250 μV and if a trial contained amplitudes greater than two standard deviations from the mean amplitude across all channels. Grand average ERP waveforms were generated by first averaging the trials per condition per electrode, and then averaging across participants.

Statistical analysis

JASP (Jeffrey's Amazing Statistic Program Team [2020], Version 0.12.2) was used for statistical analyses. The primary aim of this study was to establish a paradigm that would evoke a robust N400 ERP component. Traditionally, the N400 has been measured over midline central and parietal scalp sites. Here, we performed a simple, two-sided student t-test at scalp site Cz, comparing semantic congruence versus semantic incongruence, in a 10 ms time-window (395–405 ms) to establish whether a significant N400 response was present.

As described in the introduction and per the previous literature, we also expected to see evidence for two additional ERP components (i.e. the RP and LPC (P600)) which have been consistently observed in similar N400 designs. In the case of these latter components, while the general latencies are well-characterized, the literature is not as clear about their precise topographies, so determining *a-priori* which sites to test for their presence was not as clear as in the case of the N400. Here, we adopted the following strategy, employing simple student t-tests to test for conditional effects (congruent endings versus incongruent endings) averaged over pre-determined 10 ms time-windows (RP = 245–255 ms; LPC (P600) = 595–605 ms) across three midline scalp sites (Fz, Cz, and Pz) based on the prior literature (Osterhout and Holcomb, 1992; Luck, 2005; Lau et al., 2008). Given the use of three independent t-tests for each of these components, Bonferroni correction was applied to the resulting p-values to control for potential Type I errors. Other scalp sites (F7, T7, and P7) were investigated *post-hoc* with uncorrected, two-sided student t-tests.

Additional rmANOVA's were conducted at 250 ms, 400 ms, and 600 ms to assess for a main effect of CP (i.e. cloze probability), order (i.e. whether the congruent stimulus was heard prior to the incongruent stimulus and vice versa), linguistic division (i.e. semantic vs syntactic errors), and quartile of the experiment (i.e. time on task) on each ERP component. These rmANOVA were conducted separately for each electrode and each of the 3 time points. Amplitude values were acquired for each condition of each effect (ex. LD has 2 conditions, semantic ending error only and syntactic ending error) per condition (incongruent and congruent) in the same manner as described above. These values were then applied to either a 2 × 4 rmANOVA (i.e. condition × the number of CP or quartile divisions) or 2 × 2 rmANOVA (i.e. condition × the number of order and linguistic divisions). F-scores and p-values for midline electrodes are shown in Tables 2, and 3 for electrodes F7, T7, and P7. Significant values are bolded.

Topography plot statistics were generated using the FieldTrip toolbox (Oostenveld et al., 2011) for MATLAB and displayed using the EEGLAB toolbox. A group level cluster-based permutation test was conducted using two-tailed, independent sample t-statistics with a critical alpha-level of 0.05. This test applied the Monte-Carlo method to estimate significance probability, the triangulation method of the neighbours function for spatial clustering, and a multiple-comparison correction. Single sample clusters were combined using "maxsum" and a 5 % two-sided cutoff criterion was applied to both positive and negative clusters. Topography statistics are presented as the average significance over a 10 ms time window centered at the time point of interest (Fig. S2).

Cloze probability

To further characterize the stimulus set, a RedCap survey was employed to test the CP of all sentence endings. Each sentence in the set was presented with the final word missing (blank) and participants were required to fill in this blank with the first singular word that came to mind. If participants could not think of an answer, they were encouraged to guess rather than leave a blank. Non-answers were not counted towards CP scores; participants were removed from the survey data if they answered fewer than 10 questions out of the 221; and participants were removed if their percent correct was three standard deviations from the mean. After elimination, the survey used the responses from 134 individuals to assess the CP of each sentence. The majority of these stimuli had greater than 80 % CP. The CP distribution of sentences were shown in Fig. S3.

Data availability

This stimulus set (https://doi.org/10.5061/dryad.9ghx3ffkg) and the supporting datasets (https://doi.org/10.5061/dryad.6wwpzgmx4) are available through Dryad for the scientific community to use freely in their experiments. The stimulus set provides the auditory files for all 442 stimuli and a stimulus parameter file that includes stimulus information such as duration, target word onset, derivative divisions (i.e. CP, order, linguistic error, and quarter), and most importantly, the written form of each stimulus so that semantic comprehension via reading can be investigated. Cloze probability survey answer and result files are also within the stimulus set download.

The dataset download provides the 24 datasets in BIDS format via guidelines provided by (Pernet et al., 2019) and all the aforementioned stimulus set files. Participant information is also detailed in the dataset (.tsv and.json). The full dataset includes unfiltered EEG data (.bdf), corresponding event files, and channel rejection files for each participant (.tsv), as well as recording information, electrode positioning, and event file information (.tsv and/or.json). We additionally provide preprocessed ERP derivatives for this study (.mat), the corresponding trial rejection information per derivative for each participant (.tsv), and filtering parameters (.json). Refer to the README.txt files in both the dataset and stimulus set in order to use them appropriately. Use of this dataset, stimulus set, or presenting examples from this stimulus set should include a citation to this paper.

Code availability

The code generated for the analysis of these datasets as well as the Presentation® code is available through Zenodo via Dryad (https://doi.org/10.5061/dryad.6wwpzgmx4). The provided code was utilized to create the preprocessed ERP derivatives as well as figure components.

RESULTS

Accuracy and midline N400 ERP

All participants were able to correctly identify congruent and incongruent endings with at least 97 % accuracy. All responses were included in final analyses. Previous studies of the N400 response focused on midline electrode locations. (Kutas and Hillyard, 1984; Osterhout and Holcomb, 1992; Brown and Hagoort, 1993) To illustrate the N400 along the midline evoked by the current stimulus set, the grand mean ERPs (± SEM) for 20 participants, from electrode locations (Pz, Cz, and Fz), were characterized during sentences with congruent (blue) and incongruent (red) endings (Fig. 1). Data were aligned to the onset of the last word in each sentence (t = 0) and baselined during the 200 ms preceding this last word. As shown in Fig. 1, during congruent sentence endings, there was a positive-going potential recorded at Pz (Fig. 1(A)) that increased near the onset of the final word. A similar, increasing, positive potential was also seen at the central electrode site Cz (Fig. 1(B)). In contrast, during sentences that had incongruent endings, there was a marked negative potential that persisted for ~ 400 ms. At frontal electrode site (Fz), the negative deflection of the ERP after incongruent sentence endings began later than at the more posterior sites (Fig. 1(C)). Here the ERPs in response to congruent and incongruent sentences diverged at approximately 550 ms after final word presentation. This later negativity during incongruent endings persisted for ~ 250 ms. Individual response variation is illustrated to the right of each ERP panel, where individual mean responses over 395–405 ms after final word onset (gray band in each panel) are shown for congruent (Con) and incongruent (Incon) sentences. For both parietal and central electrodes, the majority of participants clearly had more negative N400 responses during incongruent sentences, but there was individual variation. Four of the 20 subjects had little or no increase in the N400 response. Individual responses were even more variable across subjects at the frontal electrode location (Fig. 1(C)). As expected, a significant difference between conditions was present at 400 ms at electrode site Cz (t = 4.432 $p = 7.687e\text{-}5$). A trio of t-tests was conducted at both 250 ms and 600 ms over three midline electrodes (Fz, Cz, Pz) to test for significance of the RP and LPC (P600) components, respectively. Given that multiple t-tests were conducted at three electrodes for both of these components, Bonferroni corrections were applied and the following p-values reflect these corrections. Conditional effects were found at 250 ms for electrode site Pz ($t(1,19) = 4.857$ $p = 1.627e\text{-}4$), but not at Fz ($t(1,19) = 0.611$ $p = 1.000$) or Cz ($t(1,19) = 2.757$ $p = 0.119$). Conditional effects were also found at 600 ms for electrode site Cz ($t(1,19) = 3.585$ $p = 0.011$), but not at Fz ($t(1,19) = 2.225$ $p = 0.451$) or Pz ($t(1,19) = 1.850$ $p = 1.000$).

Topography and Spatial-temporal clustering

In order to identify other regions of interest, the grand mean topography for the 20 included participants was measured over a series of five, 10 ms time windows, aligned to the onset of the last word in each sentence (Fig. 2). The scalp potentials in response to congruent sentence endings are shown in the top row (red = positive; blue = negative). There was clear frontal negativity and central-parietal positivity visible in the 400 ms time window. At later intervals (700 ms and 800 ms), this frontal activity reversed polarity, and the central parietal positivity was more centralized (row ‘a’, columns four and five). Activity across the scalp in response to incongruent sentence endings is shown in row ‘b’. There was a frontal negativity visible in the 400 ms time window that was sustained until 800 ms, where the activity was more positive. The 400 ms time window also displayed bilateral temporal positivities (row ‘b’ of column 2). At later time intervals (600–700 ms), strong positive foci were located over left temporo-parietal and parietal-occipital regions (row ‘b’ of columns three and four). The difference between congruent and incongruent responses is illustrated in row ‘c’. The 400 ms time window of row ‘d’ illustrates a negative, central-parietal difference between conditions (row ‘d’ of column 2). Here, there was also positivity in fronto-temporal regions of the scalp. Subsequently, positivity was distributed more temporo-parietally and occipitally at later time intervals (600–700 ms) (row ‘d’ of columns three and four).

A spatial–temporal clustering t-test analysis was conducted on the data in order to assess for significance and identify other significant regions and timeframes of interest. Presented in Fig. 3 are the positive (red) and negative (blue) t-values from 0 to 1000 ms respectively. Between 250 and 600 ms, there were several negative clusters denoting significant differences between conditions (incongruent vs congruent) at central-parietal electrodes. More rightward central scalp regions sustained this significant negative difference through 800 ms. There were also significant positive clusters at left frontal, temporal, and temporo-parietal electrodes between 250 and 600 ms. Additionally, left frontal electrodes showed a significant positive difference between conditions in earlier time periods, between 150 and 300 ms, while more temporo-parietal and parietal-occipital regions showed a significant positive difference at later time periods, 600–850 ms. Derived from this spatial temporal cluster analysis was a visual guide used for selection of significant regions of interest, Fig. S2. Depicted in Fig. S2 is the topographic distribution of p-values averaged over five 10 ms time windows from this spatial-temporal cluster analysis.

Semantic sensitivity over left temporal, parietal, and fronto-temporal electrodes

The grand average ERPs for the left temporal, parietal, and fronto-temporal regions of interest (identified for post hoc analysis via grand mean topography and cluster analysis (Figs. 2, 3, and S2)), are characterized in Figs. 3–6 respectively. Like midline electrodes (Cz) and (Pz) (Fig. 1(A, B)), the conditional response at left temporal (T7) and temporo-parietal (TP7) electrodes peaked 400 ms after final word onset, although with reversed polarity (Figs. 4(C) and 5(A)). In response to congruent endings, these electrodes had a small early negative deflection beginning around 200 ms which persisted for ~ 200 ms. At the end of this early deflection, ~400 ms after final word onset, there was a substantial negative potential that spanned the remainder of the epoch. This negative potential was mimicked in the incongruent condition following a positive-going potential that began around 150 ms and peaked at approximately at 400 ms (Figs. 4 (C) and 5(A)). Conditional responses in left parietal (P7) and parietal-occipital (PO7) electrodes were similar to those from temporo-parietal electrodes (Fig. 5 (B, C)). However, the onset of divergence between conditional responses began later for more occipital sites (Fig. 5(C)). Two-sided student t-tests of *post-hoc* electrodes sites (F7, T7, P7) showed an effect of condition at 600 ms for both T7 ($t(1,19) = -2.712$ $p = 0.010$) and P7 ($t(1,19) = -3.472$ $p = 0.001$), but not for F7 ($t(1,19) = -1.029$ $p = 0.310$).

The conditional responses in the left fronto-temporal regions of the scalp (Fig. 6(A, B)) were similar to those in temporal and temporo-parietal scalp regions (Figs. 4(C) and 5(A)). Specifically, in response to incongruent endings, there was a positive going deflection that began ~ 175 ms after final word onset and peaked at 400 ms, followed by a negative deflection that spanned until 800 ms (Fig. 6 (A, B)). However, in response to congruent sentence endings, fronto-temporal scalp regions had a much more robust early deflection that began ~ 100 ms and peaked ~ 250 ms after final word onset. This early peak was then followed by a negative going potential that started at 400 ms and continued for the remainder of the epoch, similar to the congruent response in temporal electrodes. Two-sided student t-tests of *post-hoc* electrodes sites showed an effect of condition at 250 ms for both F7 ($t(1,19) = -3.191$ $p = 0.003$) and T7 ($t(1,19) = -4.968$ $p = 1.467e\text{-}5$), but not P7 ($t(1,19) = -1.485$ $p = 0.146$). In order to more accurately characterize this early peak, the grand mean topography for the 20 included participants was measured over a 10 ms time window, centered at 250 ms after final word onset (Fig. 6(D)). The scalp potentials in response to congruent and incongruent sentence endings are shown on top row (red = positive; blue = negative). There was substantial frontal negativity in both conditions. In contrast, in response to congruent endings, there was positive activity in central-parietal locations, while in response to incongruent endings, there was a leftward positivity over fronto-temporal regions of the scalp. Both of these central-parietal and left fronto-temporal scalp regions were significantly different between conditions (Fig. 3).

Effects of cloze probability

To investigate for an effect of CP on conditional responses, the grand mean ERPs for midline electrodes (Pz, Cz, Fz) were characterized after separating sentences by their cloze probability (CP) score: sentences with 96–100 % CP (54 sentence pairs; Fig. 7 (A)); sentences with 90–<96 % CP (56 sentence pairs; Fig. 7(B)); sentences with 80–<90 % CP (46 sentence pairs; Fig. 7(C)); and sentences with < 80 % CP (43 sentence pairs; Fig. 7(D)). The distribution of CP in this stimulus set is shown in Fig. S3. Each CP division (Fig. 7(A–D)) for these electrodes was similar to their respective grand average ERPs depicted in Fig. 1. Furthermore, the response to each condition was similar between all CP divisions. For midline electrodes (Pz) and (Cz), in response to incongruent sentence endings, all divisions had a negative-going potential that began ~ 200 ms after final word onset, persisted for ~ 400 ms, and peaked approximately at 400 ms. In response to congruent endings, although the amplitudes varied, all divisions had a positive potential that began at the onset of the final word to ~ 600 ms. At frontal electrode (Fz), regardless of CP division, there was a negative going deflection in response to both congruent and incongruent sentence endings. An rmANOVA conducted at three separate time points (250 ms, 400 ms, and 600 ms) for midline electrodes (Fz, Cz, Pz) showed no effect of CP or Condition*CP (Table 2). The effects of CP at frontal, temporal and parietal electrode sites (F7, T7, P7) are characterized in Fig. S4. There was only an effect of CP and Condition*CP at 600 ms for parietal electrode (P7).

DISCUSSION

The purpose of this study was to develop and make publicly available, a standardized and validated auditory semantic anomaly stimulus set that evoked the classic N400 ERP component, and to provide a high-density EEG dataset from neurotypical adults (N = 20) for use by researchers interested in the neurophysiology of semantic and syntactic processing. This stimulus set did indeed evoke a reliable N400 response across the 20 NT adults in both congruent and incongruent sentences. Consistent with prior literature (Kutas and Hillyard, 1984, 1980; Osterhout and Holcomb, 1992; Luck, 2005; Lau et al., 2008; Block and Baldwin, 2010), relative to congruent sentence endings, incongruent sentence endings produced a larger amplitude (a difference of 1.43 μV at Cz and 1.46 μV at Pz) deflection 400 ms after the onset of the final word over central-parietal electrodes. These data reinforce the interpretation that the N400 represents an index of semantic processing and integration (Kutas and Hillyard, 1980; Brown and Hagoort, 1993; Connolly and Phillips, 1994; Kaan, 2007; Lau et al., 2008; Brouwer et al., 2017). Left-lateralized, semantically-driven responses 400 ms after final word onset were also observed over temporal scalp sites. Significant semantically-related activity was also observed over left anterior electrodes. There was an early negative deflection peaking ~ 250 ms post-stimulus onset in response to only congruent sentence endings, consistent with the RP literature (Dien et al., 2003; Martín-Loeches et al., 2004; Martín-Loeches, 2007; Proverbio and Riva, 2009; Gouvea et al., 2010; Fernandez and Smith Cairns, 2011; Brouwer et al., 2017). However, the topography of the RP in the prior literature was not over fronto-temporal scalp regions, but rather over the VWFA (Dien et al., 2003; Martín-Loeches, 2007). This difference is almost certainly due to the sensory modality in which the stimuli were presented. Unlike some prior studies that used visual words (Dien et al., 2003; Martín-Loeches et al., 2004) or pictures (Proverbio and Riva, 2009), stimuli here were presented auditorily in order to evoke responses in the auditory cortex and areas commonly associated with language comprehension. These topographic differences suggest that the RP topography is dependent on the sensory modality of the stimuli. There was also significant semantically-related positive activity over temporal and parietal-occipital scalp regions, peaking 600–1000 ms post-stimulus onset, that was consistent with LPC (P600) literature (Luck, 2005; Braeutigam et al., 2008; Wang et al., 2009). We would like to remind the reader that the terms LPC and P600 have been used interchangeably, with LPC being used most often when the work is specifically addressing responses to semantic contexts (Kaan et al., 2000; Hoeks et al., 2004; Kim and Osterhout, 2005; Luck, 2005; Van Herten et al., 2005; Wang et al., 2009; Gouvea et al., 2010; Friederici, 2011, 2002; Brouwer et al., 2017; Leckey and Federmeier, 2020). For our current purposes, we will refer to this component as "LPC (P600)" in an effort to encompass the terminology of prior studies and avoid confusion.

Because the peak amplitude of the N400 component has been shown to be modulated by expectation (Kutas and Hillyard, 1984; Connolly and Phillips, 1994; St. George et al., 1994; Federmeier and Kutas, 1999; Van Berkum et al., 1999; Kutas and Federmeier, 2000; Federmeier, 2007; Kaan, 2007; Lau et al., 2008; Fitz and Chang, 2019), we investigated the effect of CP on the amplitude of condition specific modulations of the N400. It was hypothesized that as word expectation became less dependent on context (i.e. lower CP), the integration of subsequent words would be more difficult. Data from the current study did not support this hypothesis. For midline, frontal, temporal, and parietal electrode sites, there were no main effects of CP, and there was no CP by Condition interaction at 400 ms. There were however, effects of CP and CP*Condition at later time points over parietal electrode sites, which suggested that the LPC (P600) was more sensitive to CP differences in this stimulus set. This lack of an effect at 400 ms may also be explained by the 'high' CP of the majority of our stimuli (Bloom and Fischler, 1980; Block and Baldwin, 2010; Khachatryan et al., 2014).

An effect of presentation order (i.e. hearing the congruent sentence before its incongruent counterpart, and vice versa) was investigated because the task design included two iterations of each sentence, where the ending was either congruent or incongruent. It was expected that semantic violations of word expectation in newer contexts would be harder to integrate than violations in repeated contexts, but the data did not support this hypothesis (Figs. S5 and S6; Tables 2 and 3). However, there was an effect of Condition*Order at electrode (Pz) that corresponded with more positive amplitudes in response to incongruent sentence endings over the 600–900 ms timeframe, when the incongruent pair was presented first (Fig. S5). This suggests that the LPC (P600) was more sensitive to the presentation order of these stimuli.

Because prior literature has suggested that the LPC (P600) component was sensitive to both semantic and syntactic errors (Ainsworth-Darnell et al., 1998; Friederici, 2002; Van Herten et al., 2005; Wang et al., 2009; Gouvea et al., 2010; Brouwer and Hoeks, 2013; Schacht et al., 2014; Brouwer et al., 2017; Fitz and Chang, 2019), effects of syntactic errors in this stimulus set were also investigated. For this analysis, responses to sentences were separated into 2 linguistic divisions (LD): 1. Sentence pairs with only semantic errors; and 2. Sentence pairs with both semantic and syntactic errors. This analysis found an effect of LD at 400 ms at frontal electrode site Fz. There were no effects of Condition*LD for any electrode at any time point (Figs. S7 and S8; Tables 2 and 3). This suggests that semantic errors were more salient than syntactic errors in this stimulus set, which reinforces that this task was primarily a semantic task for neurotypical adults.

The analyses for an effect of CP, order, and linguistic error had relatively minor effects for this NT sample, which may not be the case for other populations. Future studies that utilize these analyses could provide insight into expectation, memory, or linguistic comprehension differences between NT populations and other populations of interest. Furthermore, semantic studies that make use of this stimulus set may increase the comparability and reproducibility within this field of research.

A number of limitations of this study should be mentioned. First, although inconsistent, there was an effect of Quarter at electrode F7 and Condition*Quarter at electrode Pz over the LPC (P600) (Figs. S9 and S10; Tables 2 and 3). To reduce any effects related to particular quarters of the experiment, future studies should follow our break schedule, and consider either enforcing an extended halfway break or reducing the number of stimuli employed. Second, stimuli were presented in a pseudorandomized order. Although the grand average mean for each channel would not be affected by this, pseudorandomization could present bias in other analyses such as an effect of CP, stimulus order, linguistic error, and quarter. Therefore, differences uncovered here could have been influenced by when in the task the stimuli were presented. To have a more complete analysis of the data, future studies that use this stimulus set should properly randomize the stimuli. Third, both congruent and incongruent sentence pairs were presented in the task. Although this could introduce repetition effects, after analyzing for an effect of stimulus order on the ERPs, there was only an effect of Condition*Order at electrode Pz at 600 ms (Figs. S5 and S6; Tables 2 and 3). Moreover, providing both sentence pairs in our study provides a key control for cloze probability and allows the entirety of this stimulus set to be analyzed. Fourth, this study was primarily concerned with the first language of participants. As such, the participants in this dataset were reflective of the general population, including both left handed and bi/multilingual individuals. Due to insufficient power, an effect of handedness or bilingualism on these ERPs could not be investigated. Essentially, differences in ERPs could not be evaluated beyond individual differences. However, ERP's were constructed for both monolingual and bi/multilingual participants to show that both groups exhibit a robust N400 and LPC (P600) at major electrode locations (Fig. S11).

Future research may find that this N400 stimulus set will be especially useful for studying reading/auditory semantic abilities of atypical populations, such as individuals with ASD. Primarily because this stimulus set used a linguistic, semantic anomaly paradigm, which targeted higher-level semantic comprehension rather than single word comprehension (decoding (Fernandez and Smith Cairns, 2011: Fernandes et al., 2016)). Additionally, our stimuli are non-prosodic auditory sentences. Even though the prosody in our sentences would have been superfluous to a degree, this assures that NT individuals don't not have an advantage in language comprehension over other populations with atypical auditory processing or difficulty with communication/prosody, such as those with ASD (McCann et al., 2007; DePape et al., 2012; Eigsti et al., 2012; O'Connor, 2012; Wang and Tsao, 2015; Martzoukou et al., 2017). This stimulus set was also designed for language levels of individuals > 5 years, which could provide seamless comparability across development, and the predictability of each sentence was measured via CP. Furthermore, this stimulus set elicited the LPC (P600), which was reportedly different in individuals with ASD (Braeutigam et al., 2008; Coderre et al., 2017; DiStefano et al., 2019). Future use of this stimulus set may increase the comparability and reproducibility within reading/auditory ASD semantic processing research.

In conclusion, this auditory stimulus set successfully elicited an N400 response as well as other semantically related ERP components: the LPC (P600) and RP. This stimulus set can be used for a wide range of ages and populations, and its availability could contribute to more comparable, consistent results between future studies that use the N400 ERP component to derive information about sentential semantic integration.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The authors acknowledge the contribution of Mr. Oren Bazer for assistance with data collection, and Dr. Kathryn-Mary Wakim and Mr. David Richardson for their contributions to the analysis pipeline.

FUNDING INFORMATION

This work was supported by the Ernest J. Del Monte Institute for Neuroscience Pilot Program via a grant from the Harry T. Mangurian, Jr. Foundation (to EGF and JJF). Participant recruitment and phenotyping were conducted through the Human Clinical Phenotyping and Recruitment Core, and neurophysiological recordings were conducted through the Translational Neuroimaging and Neurophysiology Core of the University of Rochester Intellectual and Developmental Disabilities Research Center (UR-IDDRC) supported by a center grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (P50 HD103536 – to JJF).

Abbreviations:

ASD	Autism Spectrum Disorder
EEG	electroencephalography
ERP	event-related potential
LD	linguistic division
LPC	late positivity component
MRCP	Medical Research Council Psycholinguistic
NT	neurotypical
RP	recognition potential
VWFA	visual word form area

REFERENCES

Ainsworth-Darnell K, Shulman HG, Boland JE (1998) Dissociating Brain Responses to Syntactic and Semantic Anomalies: Evidence from Event-Related Potentials. J Mem Lang 38:112–130. 10.1006/imla.1997.2537.

Block CK, Baldwin CL (2010) Cloze probability and completion norms for 498 sentences: Behavioral and neural validation using event-related potentials. Behav Res Methods 42:665–670. 10.3758/BRM.42.3.665. [PubMed: 20805588]

Bloom PA, Fischler I (1980) Completion norms for 329 sentence contexts. Mem Cognit 8:631–642. 10.3758/BF03213783.

Braeutigam S, Swithenby SJ, Bailey AJ (2008) Contextual integration the unusual way: A magnetoencephalographic study of responses to semantic violation in individuals with autism spectrum disorders. Eur J Neurosci 27:1026–1036. 10.1111/j.1460-9568.2008.06064.x. [PubMed: 18333970]

Brouwer H, Crocker MW, Venhuizen NJ, Hoeks JCJ (2017) A Neurocomputational Model of the N400 and the P600 in Language Processing. Cogn Sci 41:1318–1352. 10.1111/cogs.12461. [PubMed: 28000963]

Brouwer H, Hoeks JCJ (2013) A time and place for language comprehension: Mapping the N400 and the P600 to a minimal cortical network. Front Hum Neurosci 7:1–12. 10.3389/fnhum.2013.00758. [PubMed: 23355817]

Brown C, Hagoort P (1993) The processing nature of the N400: Evidence from masked priming. J Cogn Neurosci 5:34–44. 10.1162/iocn.1993.5.1.34. [PubMed: 23972118]

Coderre EL, Chernenok M, Gordon B, Ledoux K (2017) Linguistic and Non-Linguistic Semantic Processing in Individuals with Autism Spectrum Disorders: An ERP Study. J Autism Dev Disord 47:795–812. 10.1007/s10803-016-2985-0. [PubMed: 28083778]

Connolly JF, Phillips NA (1994) Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. J Cogn Neurosci 6:256–266. 10.1162/iocn.1994.6.3.256. [PubMed: 23964975]

Cutler A, Dahan D, van Donselaar W (1997) Spoken Laiigimge : A Literature Review :: Lang. Speech 40:141–201.

Dahan D (2015) Prosody and language comprehension. Wiley Interdiscip Rev Cogn Sci 6:441–452. 10.1002/wcs.1355. [PubMed: 26267554]

Delorme A, Makeig S (2004) EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21. 10.1016/j.ineumeth.2003.10.009. [PubMed: 15102499]

DePape AMR, Chen A, Hall GBC, Trainor LJ (2012) Use of prosody and information structure in high functioning adults with Autism in relation to language ability. Front Psychol 3:1–13. 10.3389/fpsyg.2012.00072. [PubMed: 22279440]

Dien J, Frishkoff GA, Cerbone A, Tucker DM (2003) Parametric analysis of event-related potentials in semantic comprehension: Evidence for parallel brain mechanisms. Cogn Brain Res 15:137–153. 10.1016/S0926-6410(02)00147-7.

DiStefano C, Senturk D, Jeste SS (2019) ERP evidence of semantic processing in children with ASD. Dev Cogn Neurosci 36. 10.1016/j.dcn.2019.100640100640 100640. [PubMed: 30974225]

Eigsti IM, Schuh J, Mencl E, Schultz RT, Paul R (2012) The neural underpinnings of prosody in autism. Child Neuropsychol 18:600–617. 10.1080/09297049.2011.639757. [PubMed: 22176162]

Federmeier KD (2007) Thinking ahead: The role and roots of prediction in language comprehension. Psychophysiology. 10.1111/j.1469-8986.2007.00531.x.

Federmeier KD, Kutas M (1999) A Rose by Any Other Name: Long-Term Memory Structure and Sentence Processing. J Mem Lang 41:469–495. 10.1006/imla.1999.2660.

Fernandes FDM, De La Higuera Amato CA, Cardoso C, Navas ALGP, Molini-Avejonas DR (2016) Reading in Autism Spectrum Disorders: A Literature Review. Folia Phoniatr Logop 67:169–177. 10.1159/000442086.

Fernandez EM, Smith Cairns H (2011) Fundamentals of Psycholinguisitics. Sereal Untuk: Blackwell Publishing, West Sussex.

Fitz H, Chang F (2019) Language ERPs reflect learning through prediction error propagation. Cogn Psychol 111:15–52. 10.1016/j.cogpsvch.2019.03.002. [PubMed: 30921626]

Francis WN, Kucera H (1967) Computational Analysis of Present-Day American English. Providence, RI: Brown University Press.

Frazier L, Carlson K, Clifton C (2006) Prosodic phrasing is central to language comprehension. Trends Cogn Sci 10:244–249. 10.1016/j.tics.2006.04.002. [PubMed: 16651019]

Friederici AD (2002) Towards a neural basis of auditory sentence processing. Trends Cogn Sci 6:78–84. [PubMed: 15866191]

Friederici AD (2011) The brain basis of language processing: From structure to function. Physiol Rev 91:1357–1392. 10.1152/physrev.00006.2011. [PubMed: 22013214]

Frisch S, Schlesewsky M, Saddy D, Alpermann A (2002) The P600 as an indicator of syntactic ambiguity. Cognition 85:83–92. 10.1016/S0010-0277(02)00126-9.

Gilhooly KJ, Logie RH (1980) Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. Behav Res Methods Instrum 12:395–427. 10.3758/BF03201693.

Gouvea AC, Phillips C, Kazanina N, Poeppel D (2010) The linguistic processes underlying the P600. Lang Cogn Process 25:149–188. 10.1080/01690960902965951.

Hagoort P (2008) The fractionation of spoken language understanding by measuring electrical and magnetic brain signals. Phil Trans R Soc B 363:1055–1069. 10.1098/rstb.2007.2159. [PubMed: 17890190]

Hahne A, Eckstein K, Friederici AD (2004) Brain signatures of syntactic and semantic processes during children's language development. J Cogn Neurosci 16:1302–1318. 10.1162/0898929041920504. [PubMed: 15453981]

Harper-Hill K, Copland D, Arnott W (2014) Pathways to meaning: Written and spoken word priming in children with ASD versus typically developing peers. Res Autism Spectr Disord 8:1351–1363. 10.1016/j.rasd.2014.07.004.

Harris P (1975) Inferences and semantic development. J Child Lang 2:143–152. 10.1017/S0305000900000933.

Hickok G (2013) Predictive coding? Yes, but from what source? Behav Brain Sci. 10.1017/S0140525X12002750.

Hoeks JCJ, Stowe LA, Doedens G (2004) Seeing words in context: The interaction of lexical and sentence level information during reading. Cogn Brain Res 19:59–73. 10.1016/j.cogbrainres.2003.10.022.

Huettig F (2015) Four central questions about prediction in language processing. Brain Res 1626:118–135. 10.1016/j.brainres.2015.02.014. [PubMed: 25708148]

Huettig F, Mani N (2016) Is prediction necessary to understand language? Probably not. Lang Cogn Neurosci 31:19–31. 10.1080/23273798.2015.1072223.

Kaan E (2007) Event-Related Potentials and Language Processing: A Brief Overview. Lang Linguist Compass 1:571–591. 10.1111/j.1749-818x.2007.00037.x.

Kaan E, Harris A, Gibson E, Holcomb P (2000) The P600 as an index of syntactic integration difficulty. Lang Cogn Process 15:159–201. 10.1080/016909600386084.

Kamio Y, Toichi M (2000) Dual access to semantics in autism: Is pictorial access superior to verbal access? J Child Psychol Psychiatry Allied Discip 41:859–867. 10.1017/S0021963099006137.

Khachatryan E, Van Vliet M, De Deyne S, Storms G, Manvelyan H, Van Hulle MM (2014) Amplitude of N400 component unaffected by lexical priming for moderately constraining sentences. 4th Int. Work Cogn Inf Process - Proc CIP 2014. 10.1109/CIP.2014.6844516.

Kim A, Osterhout L (2005) The independence of combinatory semantic processing: Evidence from event-related potentials. J Mem Lang 52:205–225. 10.1016/Uml.2004.10.002.

Kotz SA, Friederici AD (2003) Electrophysiology of normal and pathological language processing. J Neurolinguistics 16:43–58. 10.1016/S0911-6044(02)00008-8.

Kuperberg GR (2007) Neural mechanisms of language comprehension: Challenges to syntax. Brain Res 1146:23–49. 10.1016/j.brainres.2006.12.063. [PubMed: 17400197]

Kuperberg GR, Jaeger TF (2016) What do we mean by prediction in language comprehension? Lang Cogn Neurosci 31:32–59. 10.1080/23273798.2015.1102299. [PubMed: 27135040]

Kutas M, Van Petten C (2006) Chapter 4: Psycholinguistics Electrified: Event-Related Brain Potential Investigations. Handbook of Psycholinguistics. San Diego: Academic Press. 10.1016/B978-0-12-369374-7.X5000-77.

Kutas M, Federmeier KD (2000) Electrophysiology reveals semantic memory use in language comprehension. Trends Cogn Sci. 10.1016/S1364-6613(00)01560-6.

Kutas M, Hillyard SA (1980) Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity. Science (80-). 207:203–205.

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

Lau EF, Phillips C, Poeppel D (2008) A cortical network for semantics: (De)constructing the N400. Nat Rev Neurosci 9:920–933. 10.1038/nrn2532. [PubMed: 19020511]

Leckey M, Federmeier KD (2020) The P3b and P600(s): Positive contributions to language comprehension. Psychophysiology 57:1–15. 10.1111/psyp.13351.

Leech R, Aydelott J, Symons G, Carnevale J, Dick F (2007) The development of sentence interpretation: Effects of perceptual, attentional and semantic interference. Dev Sci 10:794–813. 10.1111/j.1467-7687.2007.00628.x. [PubMed: 17973797]

Leitman DI, Foxe JJ, Butler PD, Saperstein A, Revheim N, Javitt DC (2005) Sensory contributions to impaired prosodic processing in schizophrenia. Biol Psychiatry 58:56–61. 10.1016/j.biopsych.2005.02.034. [PubMed: 15992523]

Leitman DI, Hoptman MJ, Foxe JJ, Saccente E, Wylie GR, Nierenberg J, Jalbrzikowski M, Lim KO, Javitt DC (2007) The neural substrates of impaired prosodic detection in schizophrenia and its sensorial antecedents. Am J Psychiatry 164:474–482. 10.1176/aip.2007.164.3.474. [PubMed: 17329473]

Luck S (2005) An Introduction to the Event-Related Potential Technique. Cambridge, Massachusetts: MIT Press.

Martín-Loeches M (2007) The gate for reading: Reflections on the recognition potential. Brain Res Rev 53:89–97. 10.1016/j.brainresrev.2006.07.001. [PubMed: 16938350]

Martín-Loeches M, Hinojosa JA, Casado P, Muñoz F, Fernández-Frías C (2004) Electrophysiological evidence of an early effect of sentence context in reading. Biol Psychol 65:265–280. 10.1016/j.biopsvcho.2003.07.002. [PubMed: 14757312]

Martzoukou M, Papadopoulou D, Kosmidis MH (2017) The Comprehension of Syntactic and Affective Prosody by Adults with Autism Spectrum Disorder Without Accompanying Cognitive Deficits. J Psycholinguist Res 46:1573–1595. 10.1007/s10936-017-9500-4. [PubMed: 28647830]

McCann J, Peppé S, Gibbon FE, O'Hare A, Rutherford M (2007) Prosody and its relationship to language in school-aged children with high-functioning autism. Int J Lang Commun Disord 42:682–702. 10.1080/13682820601170102. [PubMed: 17885824]

McCleery JP, Ceponiene R, Burner KM, Townsend J, Kinnear M, Schreibman L (2010) Neural correlates of verbal and nonverbal semantic integration in children with autism spectrum disorders. J Child Psychol Psychiatry Allied Discip 51:277–286. 10.1111/j.1469-7610.2009.02157.x.

O'Connor K (2012) Auditory processing in autism spectrum disorder: A review. Neurosci Biobehav Rev 36:836–854. 10.1016/j.neubiorev.2011.11.008. [PubMed: 22155284]

Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput Intell Neurosci 2011. 10.1155/2011/156869.

Osterhout L, Holcomb PJ (1992) Event-related brain potentials elicited by syntactic anomaly. J Mem Lang 31:785–806. 10.1016/0749-596X(92)90039-Z.

Pernet CR, Appelhoff S, Gorgolewski KJ, Flandin G, Phillips C, Delorme A, Oostenveld R (2019) EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. Sci Data 6:6–10. 10.1038/s41597-019-0104-8. [PubMed: 30890711]

Pickering MJ, Garrod S (2004) Toward a mechanistic psychology of dialogue. Behav Brain Sci 27:169–190. 10.1017/s0140525×04000056. [PubMed: 15595235]

Pijnacker J, Geurts B, van Lambalgen M, Buitelaar J, Hagoort P (2010) Exceptions and anomalies: An ERP study on context sensitivity in autism. Neuropsychologia 48:2940–2951. 10.1016/j.neuropsychologia.2010.06.003. [PubMed: 20542048]

Proverbio AM, Riva F (2009) RP and N400 ERP components reflect semantic violations in visual processing of human actions. Neurosci Lett 459:142–146. 10.1016/j.neulet.2009.05.012. [PubMed: 19427368]

Ribeiro TC, Valasek CA, Minati L, Boggio PS (2013) Altered semantic integration in autism beyond language: A cross-modal event-related potentials study. NeuroReport 24:414–418. 10.1097/WNR.0b013e328361315e. [PubMed: 23629689]

Ross LA, Saint-Amour D, Leavitt VM, Javitt DC, Foxe JJ (2007) Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. Cereb Cortex 17:1147–1153. 10.1093/cercor/bhl024. [PubMed: 16785256]

Ross LA, Molholm S, Blanco D, Gomez-Ramirez M, Saint-Amour D, Foxe JJ (2011) The development of multisensory speech perception continues into the late childhood years. Eur J Neurosci 33:2329–2337. 10.1111/j.1460-9568.2011.07685.x. [PubMed: 21615556]

Schacht A, Sommer W, Shmuilovich O, Martíenz PC, Martín-Loeches M (2014) Differential task effects on N400 and P600 elicited by semantic and syntactic violations. PLoS ONE 9:1–7. 10.1371/iournal.pone.0091226.

St. George M, Mannes S, Hoffman JE (1994) Global semantic expectancy and language comprehension. J Cogn Neurosci 6:70–83. 10.1162/iocn.1994.6.1.70. [PubMed: 23962331]

Van Berkum JJA, Hagoort P, Brown CM (1999) Semantic integration in sentences and discourse: Evidence from the N400. J Cogn Neurosci 11:657–671. 10.1162/089892999563724. [PubMed: 10601747]

Van Herten M, Kolk HHJ, Chwilla DJ (2005) An ERP study of P600 effects elicited by semantic anomalies. Cogn Brain Res 22:241–255. 10.1016/j.cogbrainres.2004.09.002.

Van Petten C, Luka BJ (2006) Neural localization of semantic context effects in electromagnetic and hemodynamic studies. Brain Lang 97:279–293. 10.1016/j.bandl.2005.11.003. [PubMed: 16343606]

Wang S, Dong X, Ren Y, Yang Y (2009) The development of semantic priming effect in childhood: An event-related potential study. NeuroReport 20:574–578. 10.1097/WNR.0b013e328329f215. [PubMed: 19300294]

Wang JE, Tsao FM (2015) Emotional prosody perception and its association with pragmatic language in school-aged children with high-function autism. Res Dev Disabil 37:162–170. 10.1016/j.ridd.2014.11.013. [PubMed: 25463248]

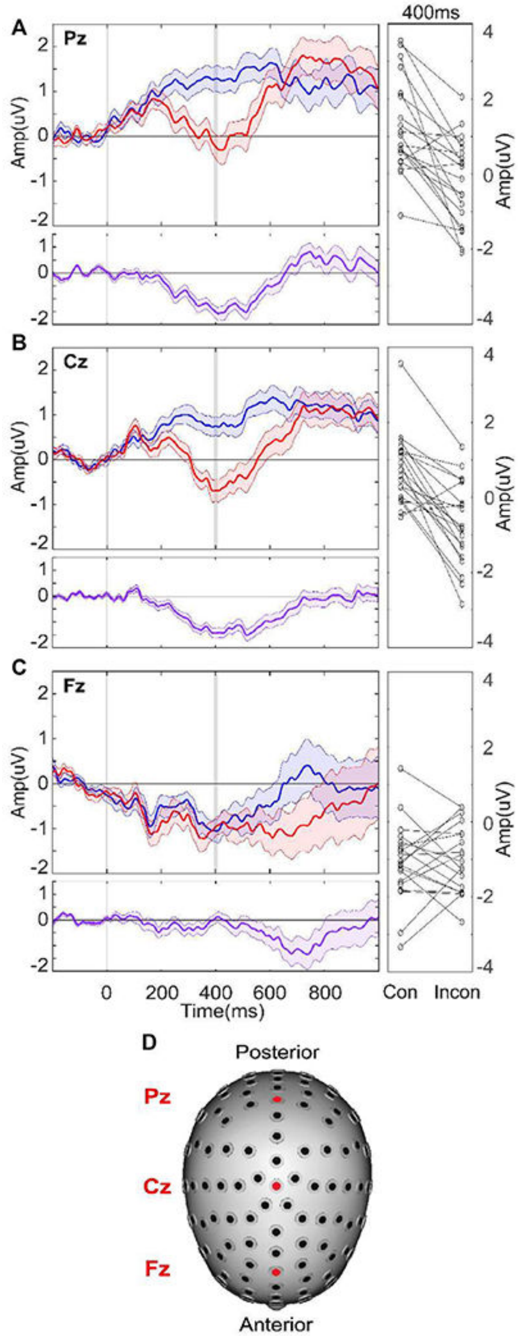

Fig. 1. ERP Plots for Midline Electrodes.

ERP plots for (**A**) Pz, (**B**) Cz, and (**C**) Fz, show amplitude changes over 1000 ms, per condition (Blue, congruent. Red, Incongruent). Difference plots are in purple (incongruent – congruent amplitudes). The shading around each line represents the s.e.m. To the right of each electrode is the individual participant amplitudes at 400 ms (avg. 395–405 ms represented by the gray bar) per condition (congruent vs incongruent). (**D**) Head model showing electrode locations.

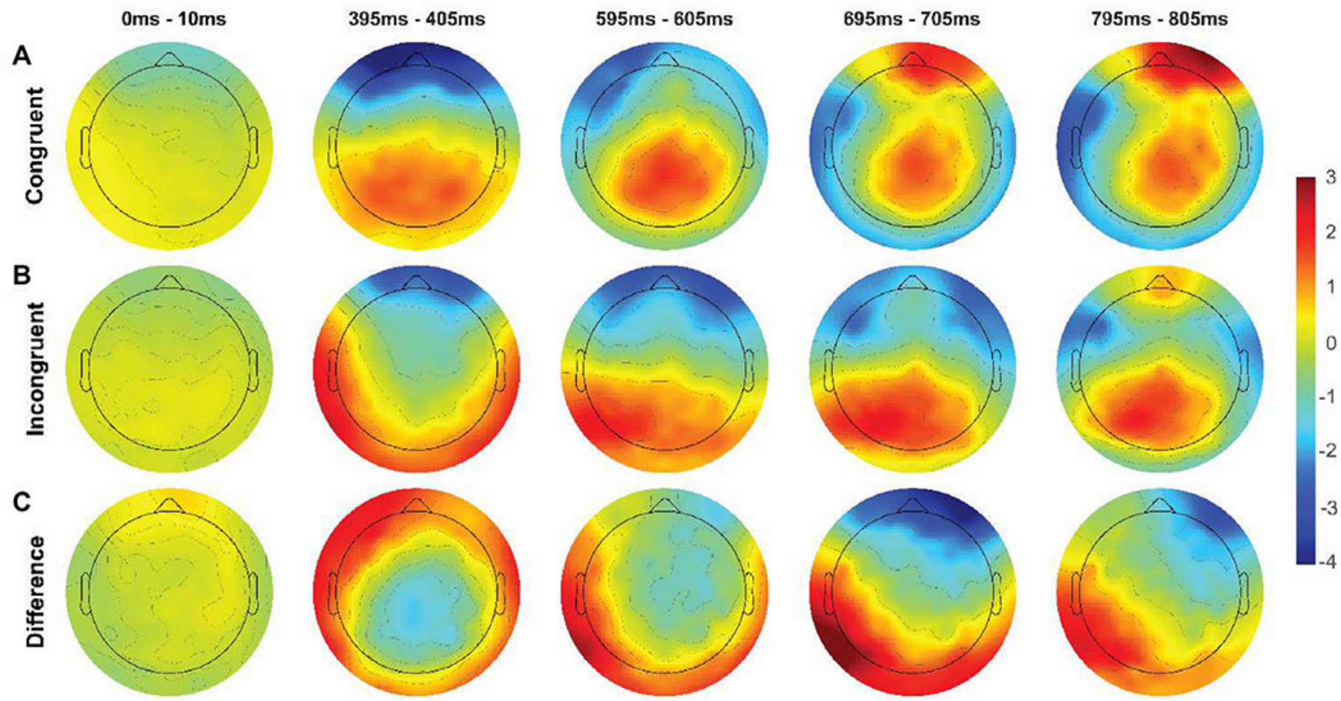

Fig. 2. Grand average scalp topography over 5 time points.

The average distribution of EEG amplitudes for the 20 included participants based on the average reference. Focused on 5 major time points (0 ms, 400 ms, 600 ms, 700 ms, and 800 ms) averaged over 10 ms. Topographic amplitudes in response to (**A**) Congruent stimuli, (**B**) Incongruent stimuli. (**C**) The difference between conditions (Incongruent – Congruent amplitudes).

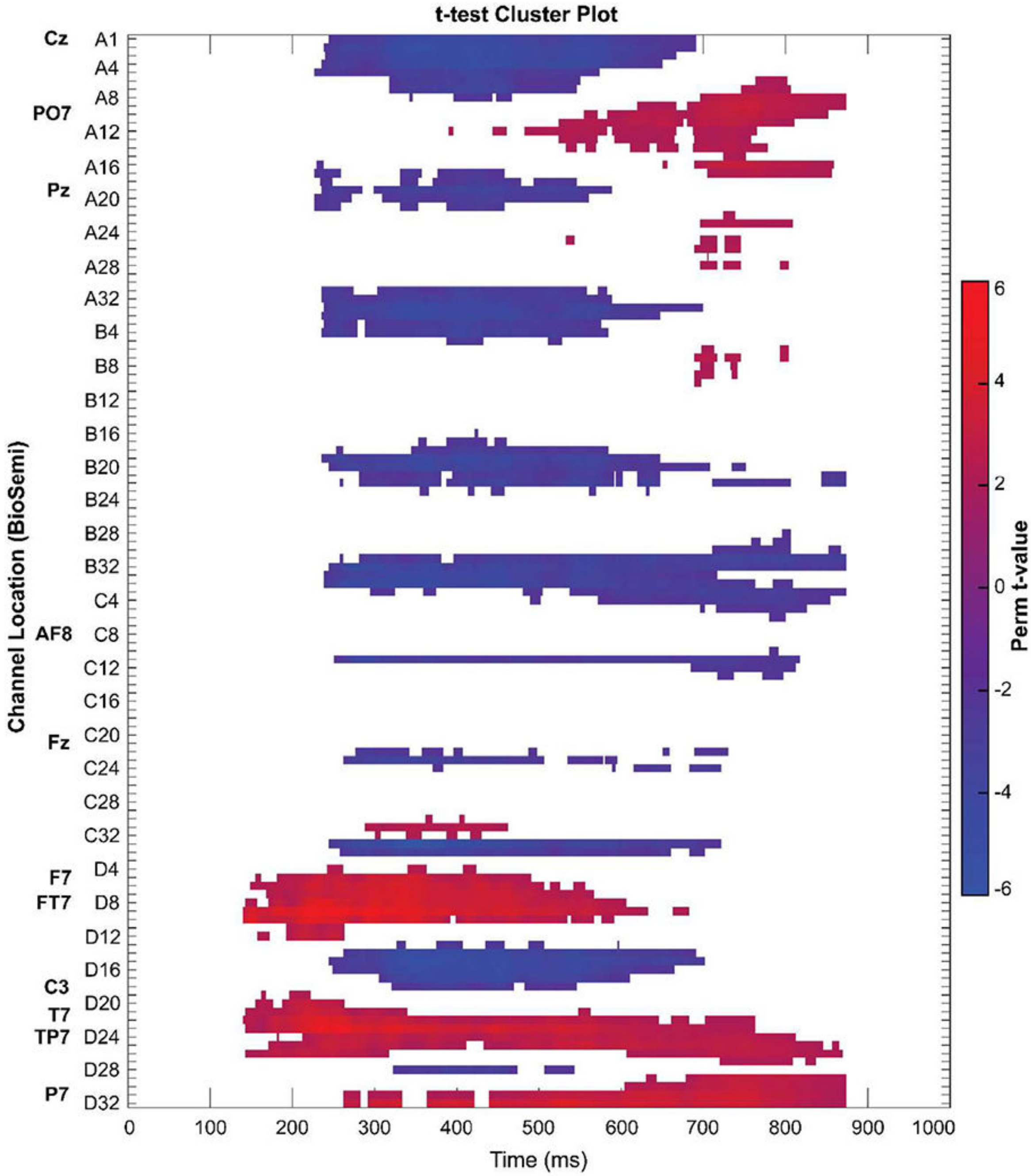

Fig. 3. Cluster plot of t-values for all 128 channels after conducting two-tailed independent sample t-tests.

T-tests used the Monte-Carlo method, the triangulation method for spatial clustering, and a multiple-comparison correction. A 5% two-sided cutoff criterion was applied to both positive and negative clusters. Channels are identified by BioSemi cap coordinates and the electrodes mentioned in this dataset are noted with the standard 10–20 electrode positions.

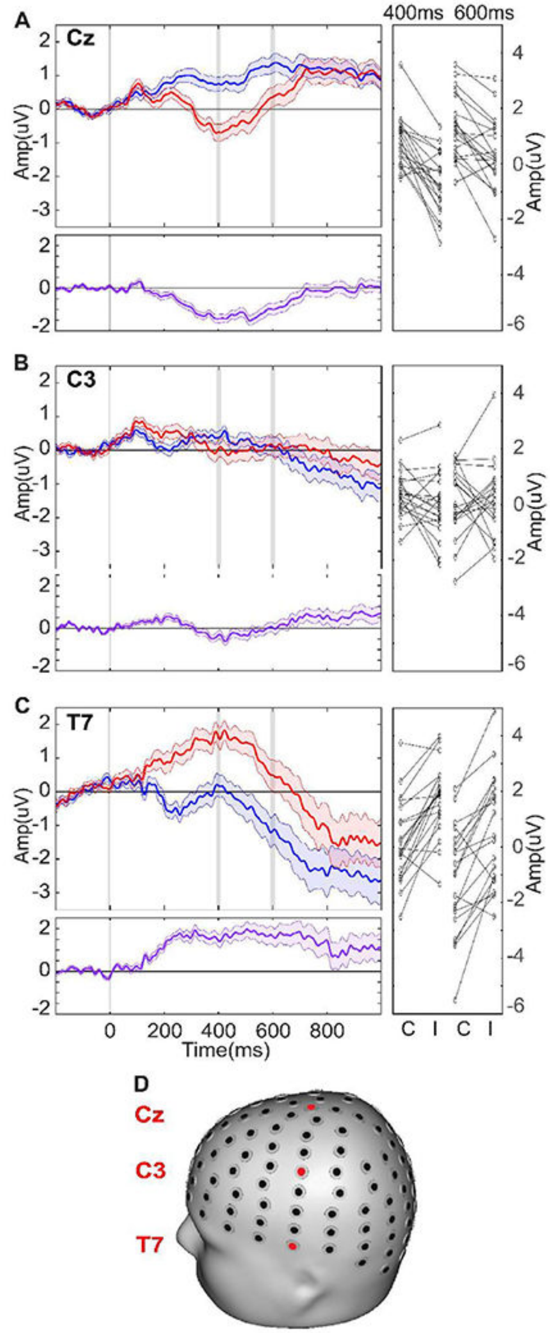

Fig. 4. ERP Plots for Temporal Electrodes.

ERP plots for (**A**) Cz, (**B**) C3, and (**C**) T7, show amplitude changes over 1000 ms, per condition (Blue, congruent. Red, Incongruent). Difference plots are in purple (incongruent – congruent amplitudes). The shading around each line represents the s.e.m. To the right of each electrode is the individual participant amplitudes at 400 ms and 600 ms (avg. 395–405 ms and 595–605 ms respectively, represented by the gray bars) per condition (congruent vs incongruent). (**D**) Head model showing electrode locations.

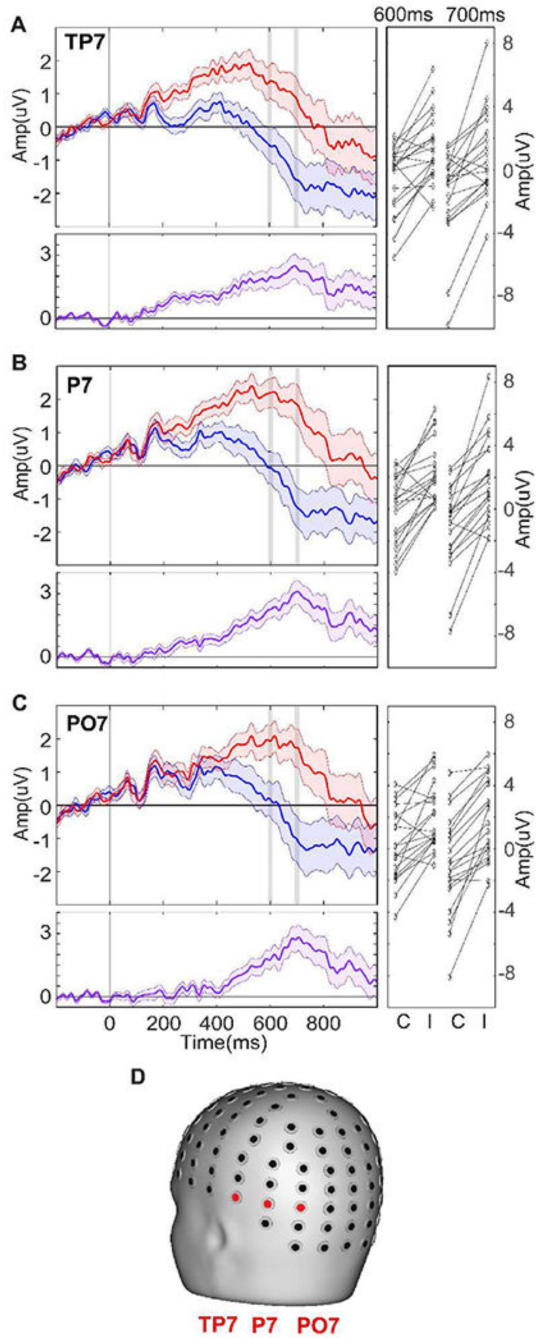

Fig. 5. ERP Plots for Temporo-Occipital Electrodes.

ERP plots for (**A**) TP7, (**B**) P7, and (**C**) PO7, show amplitude changes over 1000 ms, per condition (Blue, congruent. Red, Incongruent). Difference plots are in purple (incongruent – congruent amplitudes). The shading around each line represents the s.e.m. To the right of each electrode is the individual participant amplitudes at 600 ms and 700 ms (avg. 595–605 ms and 695–705 ms respectively, represented by the gray bars) per condition (congruent vs incongruent). (**D**) Head model showing electrode locations.

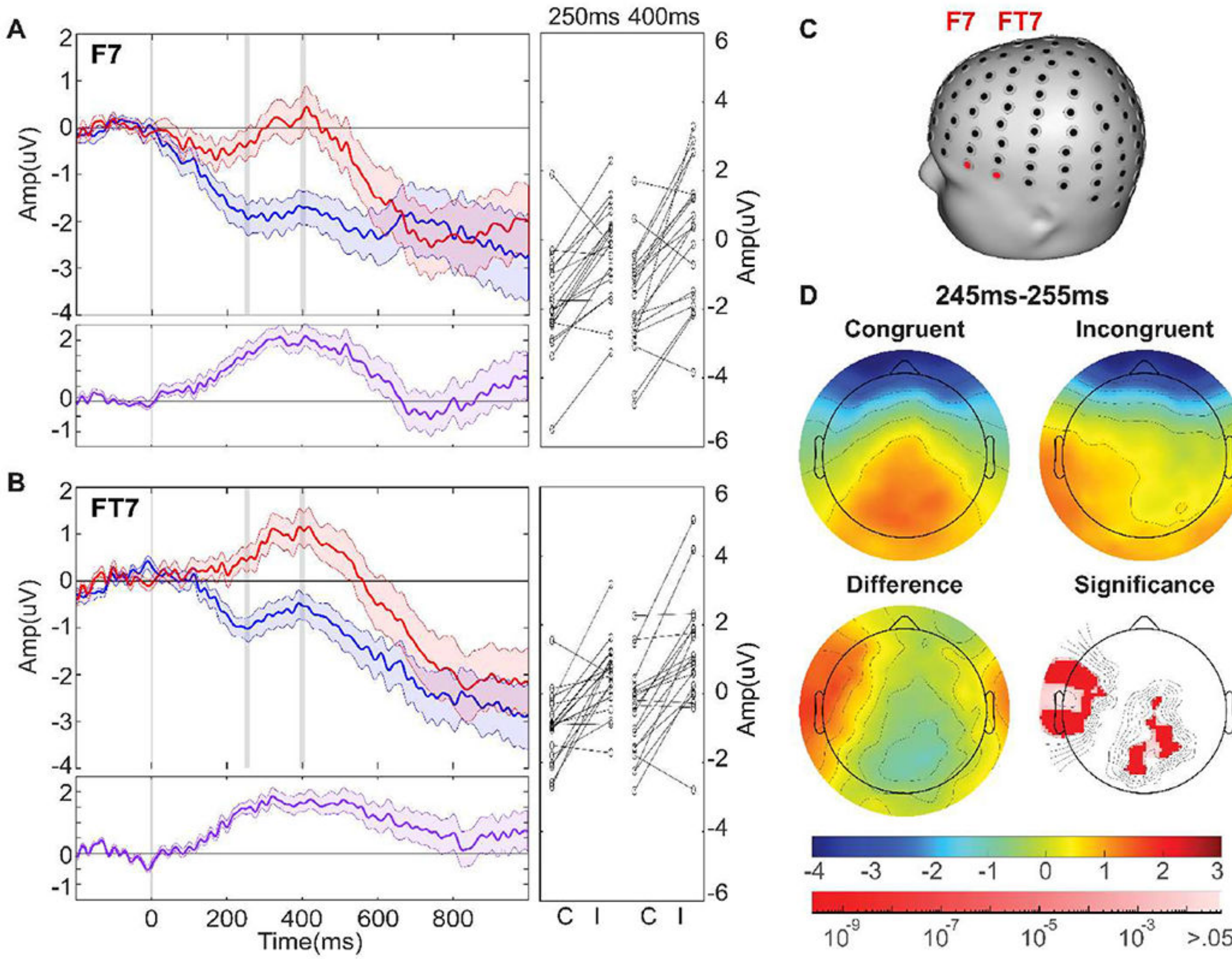

Fig. 6. ERP Plots for Electrodes over Fronto-Temporal electrodes.

ERP plots for (**A**) F7 and (**B**) FT7, show amplitude changes over 1000 ms, per condition (Blue, congruent. Red, Incongruent). Difference plots are in purple (incongruent – congruent amplitudes). The shading around each line represents the s.e.m. To the right of each electrode is the individual participant amplitudes at 250 ms and 400 ms (avg. 245–255 ms and 395–405 ms respectively, represented by the gray bars) per condition (congruent vs incongruent). (**C**) Head model showing electrode locations. (**D**) Topography averaged between 225 and 250 ms for congruent stimuli, incongruent stimuli, difference between conditions (incongruent – congruent amplitudes). Although averaged over 10 ms, p-value cluster plots for this time frame were provided as a visual guide (white means p-value is > 0.05). Refer to Fig. 3 for spatial–temporal clustering analysis.

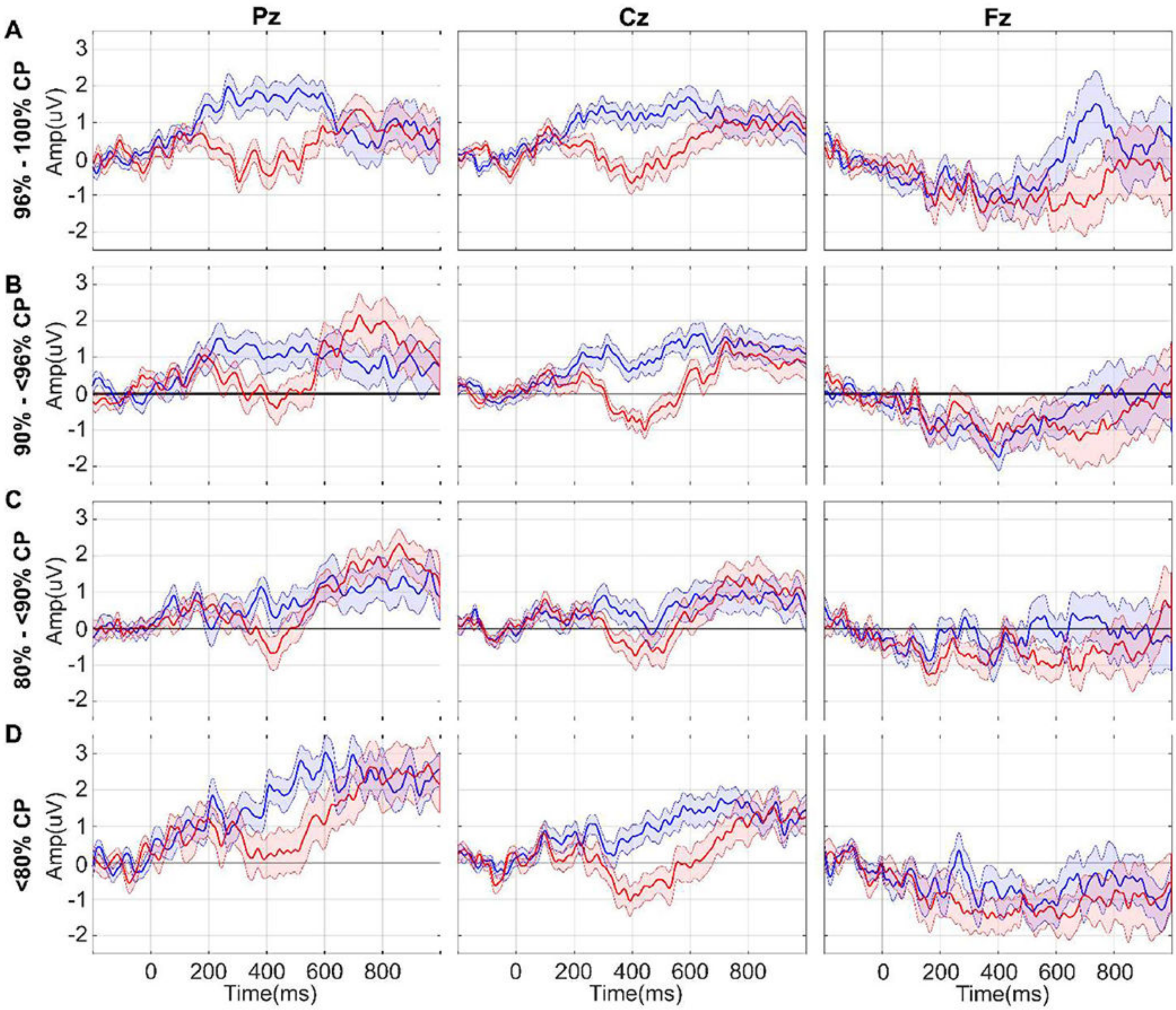

Fig. 7. ERP Plots of Midline Electrodes Pz, Cz, and Fz separated by differing levels of Cloze Probability.

The ERP plots show amplitude changes over 1000 ms, per condition (Blue, congruent. Red, Incongruent). The shading around each line represents the s.e.m. (**A**) Sentences with 96–100 % cloze probability (54 congruent and 55 incongruent stimuli). (**B**) Sentences with 90–<96 % cloze probability (56 sentences pairs). (**C**) Sentences with 80–<90 % cloze probability (47 sentence pairs). (**D**) Sentences with < 80 % cloze probability (43 congruent and 42 incongruent stimuli). Statistics for these electrodes are on Table 2.

Table 1.

Demographic data of the participants.

Includes gender, age, handedness, primary language, the number of languages known, and whether the participant was included in the final analysis

Subject	*Gender*	*Age*	*Handedness*	Language			*Final Dataset*
				Primary/First	*Known*		
01	Female	24	Right	English	Mono-lingual	1	Included
02	Female	26	Left	English	Bi-lingual	2	Included
03	Male	26	Left	English	Bi-lingual	2	Included
04	Male	19	Right	English	Multi-lingual	7	Included
05	Male	20	Right	English	Bi-lingual	2	Excluded (fell asleep)
06	Female	24	Right	English	Mono-lingual	1	Included
07	Male	23	Right	English	Mono-lingual	1	Included
08	Female	24	Right	English	Mono-lingual	1	Included
09	Male	28	Right	English	Mono-lingual	1	Included
10	Female	23	Right	English	Mono-lingual	1	Excluded (excessive noise)
11	Male	28	Left	English	Mono-lingual	1	Included
12	Female	29	Right	English	Bi-lingual	2	Included
13	Male	32	Right	English	Mono-lingual	1	Included
14	Male	27	Right	English	Mono-lingual	1	Included
15	Female	21	Right	English	Bi-lingual	2	Excluded (excessive noise)
16	Female	22	Right	English	Mono-lingual	1	Included
17	Male	26	Right	English	Multi-lingual	3	Included
18	Male	40	Left	English	Mono-lingual	1	Excluded (excessive moving)
19	Female	22	Right	English	Mono-lingual	1	Included
20	Female	21	Right	English	Multi-lingual	3	Included
21	Male	32	Right	English	Mono-lingual	1	Included
22	Male	24	Right	English	Bi-lingual	2	Included
23	Male	35	Right	English	Mono-lingual	1	Included
24	Female	18	Right	English	Bi-lingual	2	Included

Table 2.

Alternative main effects for midline electrodes Fz, Cz, and Pz at four time intervals.

Amplitudes for the congruent and incongruent condition were acquired for each electrode, effect (i.e. cloze probability (CP), order, language division (LD), and quarter (Q) of the experiment), time interval, and participant by averaging the amplitude values from 5 ms before the time point to 5 ms after. Using these values, the main effects of condition and alternative affects were assessed using the repeated measure ANOVA. Provided are the F statistic (F) and p-values from this test. Bolded values represent a significant difference between conditions (*p* < 0.05, two-tailed). Italicized values represent where the assumption of sphericity (*p* < 0.05, two-tailed) was violated using the Mauchly's test of sphericity

	Time Interval		
	250 ms	***400 ms***	***600 ms***
Main Effect	**Electrode Fz**		
Close Probability	F = 1.088	F = 0.677	F = 0.625
	p = 0.361	*p* = 0.570	*p* = 0.602
Order	**F = 4.552**	F = 0.017	F = 0.141
	***p* = 0.046**	*p* = 0.898	*p* = 0.712
Language Division	F = 2.854	**F = 4.387**	F = 1.578
	p = 0.108	***p* = 0.050**	*p* = 0.224
Quarter	F = 0.369	F = 0.120	F = 0.183
	p = 0.776	*p* = 0.948	*p* = 0.907
Condition*CP	*F = 1.873*	F = 1.544	F = 0.823
	p = 0.144	*p* = 0.213	*p* = 0.486
Condition*Order	F = 0.061	F = 0.846	F = 1.578
	p = 0.807	*p* = 0.369	*p* = 0.224
Condition*LD	F = 0.003	F = 0.290	F = 0.087
	p = 0.959	*p* = 0.597	*p* = 0.771
Condition*Q	F = 0.183	F = 0.356	*F = 0.568*
	p = 0.907	*p* = 0.785	*p = 0.639*
	Electrode Cz		
Close Probability	F = 1.403	*F = 1.1665*	F = 2.288
	p = 0.252	*p = 0.331*	*p* = 0.088
Order	F = 0.917	F = 0.423	F = 0.316
	p = 0.250	*p* = 0.523	*p* = 0.518
Language Division	F = 1.364	F = 1.247	F = 0.102
	p = 0.247	*p* = 0.278	*p* = 0.752
Quarter	F = 0.655	F = 0.530	F = 0.494
	p = 0.583	*p* = 0.663	*p* = 0.688
Condition*CP	F = 1.822	F = 1.596	F = 1.894
	p = 0.153	*p* = 0.200	*p* = 0.141
Condition*Order	F = 0.008	F = 0.088	F = 1.312
	p = 0.931	*p* = 0.770	*p* = 0.266
Condition*LD	F = 0.113	F = 0.040	F = 0.396

	Time Interval		
	250 ms	***400 ms***	***600 ms***
	p = 0.740	*p* = 0.844	*p* = 0.537
Condition*Q	F = 0.433	F = 0.431	F = 0.471
	p = 0.730	*p* = 0.732	*p* = 0.704
	Electrode Pz		
Close Probability	*F = 1.165*	F = 0.762	F = 1.886
	p = 0.331	*p* = 0.520	*p* = 0.142
Order	F = 2.628	F = 0.548	F = 1.041
	p = 0.121	*p* = 0.466	*p* = 0.320
Language Division	F = 1.868	F = 2.320	F = 0.137
	p = 0.188	*p* = 0.144	*p* = 0.715
Quarter	F = 0.578	F = 0.777	F = 1.250
	p = 0.632	*p* = 0.512	*p* = 0.300
Condition*CP	F = 0.797	F = 1.049	**F = 2.781**
	p = 0.501	*p* = 0.378	***p* = 0.049**
Condition*Order	F = 2.446	F = 1.521	**F = 9.046**
	p = 0.134	*p* = 0.232	***p* = 0.007**
Condition*LD	F = 1.350E-5	F = 2.589E-4	F = 0.077
	p = 0.006	*p* = 0.122	*p* = 0.784
Condition*Q	F = 0.913	*F = 0.700*	**F = 2.793**
	p = 0.441	*p = 0.556*	***p* = 0.048**

Table 3.

Main effects for electrodes F7, T7, and P7 at four time intervals.

Amplitudes for the congruent and incongruent condition were acquired for each electrode, effect (i.e. cloze probability (CP), order, language division (LD), and quarter (Q) of the experiment), time interval, and participant by averaging the amplitude values from 5 ms before the time point to 5 ms after. Using these values, the main effects of condition and alternative affects were assessed using the repeated measure ANOVA. Provided are the F statistic (F) and p-values from this test. Bolded values represent a significant difference between conditions (p < 0.05, two-tailed). Italicized values represent where the assumption of sphericity (p < 0.05, two-tailed) was violated using the Mauchly's test of sphericity

	Time Interval		
	250 ms	***400 ms***	***600 ms***
Main Effect	**Electrode F7**		
Close Probability	F = 0.213	F = 0.282	F = 0.662
	p = 0.887	*p* = 0.838	*p* = 0.579
Order	F = 0.038	F = 0.024	F = 0.024
	p = 0.847	*p* = 0.878	*p* = 0.878
Language Division	F = 0.352	F = 2.897	F = 0.158
	p = 0.560	*p* = 0.105	*p* = 0.696
Quarter	F = 0.774	F = 1.061	**F = 3.645**
	p = 0.514	*p* = 0.373	***p* = 0.018**
Condition*CP	F = 0.868	F = 0.797	F = 0.673
	p = 0.463	*p* = 0.501	*p* = 0.572
Condition*Order	F = 0.590	F = 0.250	F = 3.293
	p = 0.452	*p* = 0.623	*p* = 0.085
Condition*LD	F = 2.040	F = 0.510	F = 2.000
	p = 0.169	*p* = 0.484	*p* = 0.173
Condition*Q	F = 0.039	F = 0.153	F = 0.503
	p = 0.990	*p* = 0.927	*p* = 0.682
	Electrode T7		
Close Probability	F = 0.716	*F = 1.093*	F = 1.099
	p = 0.546	*p* = 0.360	*p* = 0.357
Order	F = 1.268	F = 1.995	F = 0.039
	p = 0.274	*p* = 0.174	*p* = 0.845
Language Division	F = 0.027	F = 0.612	F = 0.393
	p = 0.872	*p* = 0.444	*p* = 0.538
Quarter	F = 1.247	F = 1.289	F = 0.347
	p = 0.301	*p* = 0.287	*p* = 0.791
Condition*CP	*F = 0.955*	*F = 1.251*	*F = 1.461*
	p = 0.420	*p = 0.300*	*p = 0.235*
Condition*Order	F = 2.049	F = 0.347	F = 0.028
	p = 0.169	*p* = 0.563	*p* = 0.870
Condition*LD	F = 0.301	F = 4.757E-9	F = 0.091

	Time Interval		
	250 ms	***400 ms***	***600 ms***
	p = 0.590	*p* = 1.000	*p* = 0.766
Condition*Q	F = 0.168	F = 1.096	F = 0.602
	p = 0.918	*p* = 0.358	*p* = 0.617
	Electrode P7		
Close Probability	*F = 0.407*	F = 0.780	**F = 10.003**
	p = 0.748	*p* = 0.510	***p* = 2.150E-5**
Order	F = 0.390	F = 0.035	F = 0.259
	p = 0.540	*p* = 0.853	*p* = 0.062
Language Division	F = 1.118	F = 1.870	F = 0.024
	p = 0.304	*p* = 0.187	*p* = 0.879
Quarter	F = 0.353	F = 0.874	F = 0.882
	p = 0.787	*p* = 0.460	*p* = 0.456
Condition*CP	F = 1.502	F = 0.037	**F = 6.275**
	p = 0.224	*p* = 0.990	***p* = 9.864E-4**
Condition*Order	F = 0.477	F = 1.162	F = 1.088
	p = 0.498	*p* = 0.295	*p* = 0.310
Condition*LD	F = 0.306	F = 2.963	F = 0.547
	p = 0.587	*p* = 0.101	*p* = 0.468
Condition*Q	*F = 0.214*	F = 0.884	F = 0.855
	p = 0.886	*p* = 0.455	*p* = 0.470