Tell me sweet little lies: An event-related potentials study on the processing of social lies

Article *in* Cognitive Affective & Behavioral Neuroscience · March 2016

DOI: 10.3758/s13415-016-0418-3

3 authors:

Eva M Moreno
Complutense University of Madrid
23 PUBLICATIONS **873** CITATIONS
SEE PROFILE

Pilar Casado
Instituto de Salud Carlos III
53 PUBLICATIONS **617** CITATIONS
SEE PROFILE

Manuel Martín-Loeches
Complutense University of Madrid
127 PUBLICATIONS **2,084** CITATIONS
SEE PROFILE

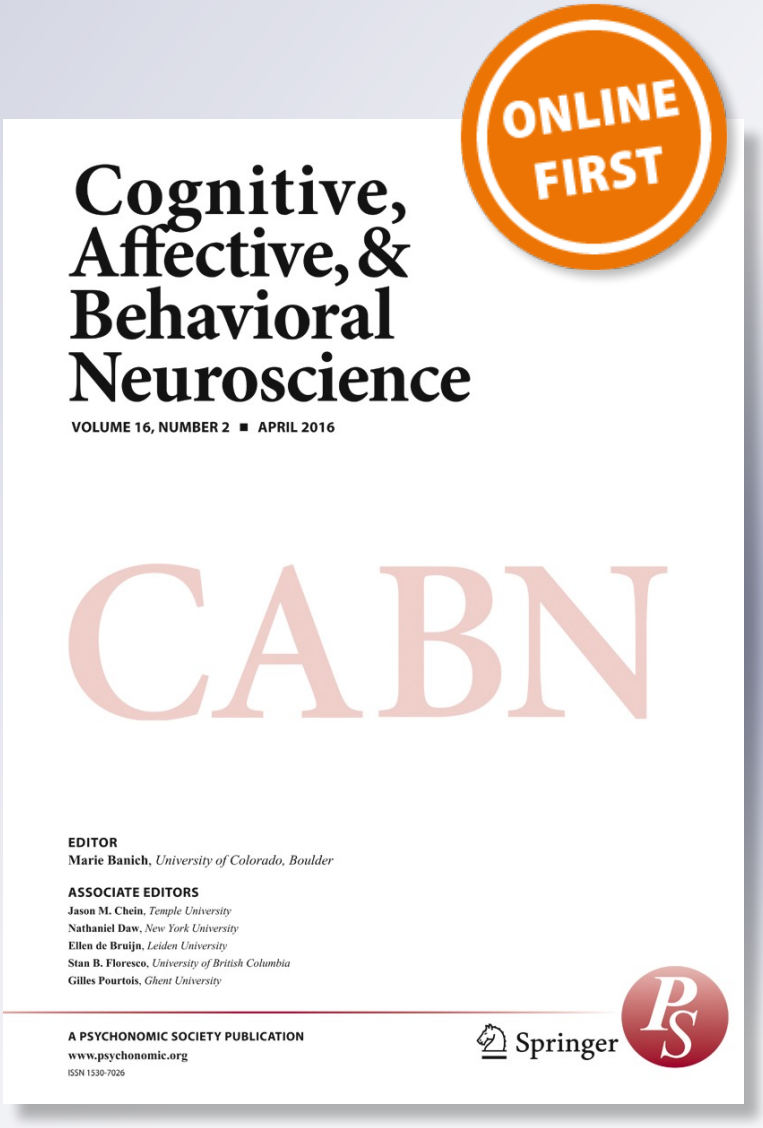
Tell me sweet little lies: An event-related potentials study on the processing of social lies

Eva M. Moreno, Pilar Casado & Manuel Martín-Loeches

Cognitive, Affective, & Behavioral Neuroscience

ISSN 1530-7026

Cogn Affect Behav Neurosci
DOI 10.3758/s13415-016-0418-3



Springer

Cogn Affect Behav Neurosci
DOI 10.3758/s13415-016-0418-3

Tell me sweet little lies: An event-related potentials study on the processing of social lies

Eva M. Moreno[1] · **Pilar Casado**[2] · **Manuel Martín-Loeches**[2]

Abstract In reading tasks, words that convey a false statement elicit an enhanced N400 brainwave response, relative to words that convey a true statement. N400 amplitude reductions are generally linked to the online expectancy of upcoming words in discourse. White lies, contrary to false statements, may not be unexpected in social scenarios. We used the event-related potential (ERP) technique to determine whether there is an impact of social context on sentence processing. We measured ERP responses to target words that either conveyed a social "white" lie or a socially impolite blunt truth, relative to semantic violations. Word expectancy was controlled for by equating the cloze probabilities of white lying and blunt true targets, as measured in previous paper-and-pencil tests. We obtained a classic semantic violation effect (a larger N400 for semantic incongruities relative to sense making statements). White lies, in contrast to false statements, did not enhance the amplitude of the N400 component. Interestingly, blunt true statements yielded both a late frontal positivity and an N400 response in those scenarios particularly biased to white lying. Thus, white lies do not interfere with online semantic processing, and they do not engage further reanalysis processes, which are typically indexed by subsequent late positivity ERP effects. Instead, an N400 and a late frontal positivity obtained in response to blunt true statements indicate that they were treated as unexpected events. In conclusion, unwritten rules of social communicative behavior influence the electrical brain response to locally coherent but socially inappropriate statements.

Keywords Event-related potentials · Language · Social lies · N400 · Frontal late positivity

People lie, on average, once or twice per day (DePaulo, Kirkendol, Kashy, Wyer, & Epstein, 1996; but see Serota, Levine, & Boster, 2010). Lying is generally considered an antisocial behavior, and most human cultures have some prohibition against lying. However, it sometimes serves a prosocial function depending on the context in which communication takes place as well as on the speaker's motivation to lie. In human social interactions, a particular type of lies commonly named "white" lies are often uttered. They consist of trivial, diplomatic, or well-intentioned untruths told in order to be polite or to stop someone from being upset by the truth. In fact, quite early in our social development, we are able to make moral judgments about lie telling, taking into account its expected social consequences. In a public situation where telling a truth is likely to have a negative social consequence (e.g., hurt feelings), children as young as 7 to 11 years of age rate lie telling more positively than truth telling (Ma, Xu, Heyman, & Lee, 2011).

The high temporal resolution of event-related potentials (ERPs), make them ideally suited to study how social lying is processed. ERPs allow a direct measure of neural responses without the need of any additional task or behavioral response such as making a grammatical, semantic, or moral judgment. Thus far, ERP and functional magnetic resonance imaging (fMRI) studies have explored lying from the "liar" perspective (Proverbio, Vanutelli, & Adorni, 2013). Only a few ERP studies have explored lying from the receiver of the lie

✉ Eva M. Moreno
emmoreno@ucm.es

[1] Human Brain Mapping Unit at Instituto Pluridisciplinar & Psychology Department, Universidad Complutense de Madrid, Paseo Juan XXIII 1, 28040 Madrid, Spain

[2] Center for Human Evolution and Behavior, UCM-ISCIII, Madrid, Spain

Published online: 23 March 2016

perspective. In this regard, and resembling the classical N400 effect elicited by semantic incongruities embedded in a sentence (Kutas & Hillyard, 1980), violations of true facts from our world knowledge (e.g., fact-related lies) elicit an N400 effect, that is, a larger negative-going voltage post target word onset, in the 400 ms range, relative to target words conveying the truth. For example, Dutch native speakers reading the sentence "The Dutch trains are *white* and very crowded" generate a large N400 at the critical word *white*, since it is a well-known fact among Dutch people that their trains are *yellow* instead (Hagoort, Hald, Bastiaansen, & Petersson, 2004; Hald, Steenbeek-Planting, & Hagoort, 2007). This result indicates that the processing of false statements incurs in a difficulty of semantic processing. Thus, statements that clash with world knowledge stored in long-term memory elicit an N400 response. However, the most recent views on what the N400 indexes highlight its role as a sign that online predictions are being made about upcoming words in discourse (Boudewyn, Long, & Swaab, 2015; Otten, Nieuwland, & Van Berkum, 2007; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; Wicha, Moreno, & Kutas, 2004). Thereafter, the N400 is not as sensitive to truth-value computations as it is an online predictor for upcoming information based on world knowledge (Nieuwland, 2015; Nieuwland & Kuperberg, 2008).

ERP measures also reveal that a wider discourse context (e.g., a fictitious one) can overrule the consequences of processing a priori unexpected statements. For example, reading that a peanut was "in love" becomes paradoxically easier to process than reading that it was "salted" when embedded in a supportive discourse context about an animated peanut (Nieuwland & Van Berkum, 2006). Thus, the N400 effect is sometimes neutralized based on a broader discourse context. Intuitively, a part of our world knowledge includes rules of social communicative behavior. Some statements are socially sanctioned despite being inaccurate based on prior context (white lies). Others can be true based on prior discourse context but are socially inappropriate statements and perhaps unexpected to be told. Online expectancy might go beyond the lexical level (a word that is expected to be told) to the social level (what kinds of words are expected or unexpected to be told in a social scenario).

In the ERP literature on language processing, two types of late positivities (frontal versus parietally distributed) have been linked to different kinds of unexpected events (Van Petten & Luka, 2012). Frontal late positivities arise for plausible but unexpected words in highly constraining contexts (Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007). This frontal effect is linked to the cost of a disconfirmed prediction. Its brain topography, maximal over frontal sites, suggests that it arises from different brain regions and thus reflects different functional processes than those attributed to parietal P600 effects (Thornhill & Van Petten, 2012; Van Petten & Luka, 2012).

Parietal positivities, instead, have been linked to the processing of anomalies that prompt sentence reanalysis processes. For example, the comprehension of ironic versus literal meanings has revealed an increase in this P600 ERP component (Regel, Coulson, & Gunter, 2010; Regel, Gunter, & Friederici, 2011; Regel, Meyer, & Gunter, 2014). To date, the P600 functional interpretation remains controversial. Originally obtained for syntactic violations and ambiguities, it is also observed following semantic abnormalities (see Kuperberg, 2007, for a review). A recent study distinguishes between a widespread P600 reflecting reanalysis or repair processes that follow syntactic anomalies and a more constrained centro-parietal P600 to pragmatic anomalies, such as the processing of ironic messages (Regel et al., 2014). However, a P600 effect most prominent over frontal electrode sites has also been reported for the processing of ironic versus literal statements (Spotorno, Cheylus, Van Der Henst, & Noveck, 2013). Finally, when extralinguistic information (i.e., prosody) is available, brain responses show a distinction between white lies and true compliments (Rigoulot, Fish, & Pell, 2014). Specifically, listeners generate greater P600 amplitudes over right frontal areas of the scalp in response to sincere versus insincere compliments.

In the current study, we aimed to determine whether there is an impact of social knowledge on sentence comprehension, beyond the online expectancy for a particular word item. We specifically examined how the brain responds to visually presented target words that either conveyed a "white" lie or a socially impolite blunt truth. To that end, we presented an initial paragraph depicting a social situation (e.g., having guests for dinner) in which the truth was unpleasant (e.g., the host is not a good cook and dinner got burned). In that social scenario, the host requests an opinion (e.g., she asks her guests how much they liked dinner) and then a word-by-word sentence is presented to the reader in which one of the guests tells a white lie (e.g., the meat sauce was *tasty*) or a socially impolite truth (e.g., the meat sauce was *overcooked*). Based on prior paper-and-pencil tests, these alternative endings (i.e., those marked in italics in previous examples) were matched in their low likelihood of being used to complete the sentence. They are thus expected to elicit an equally large N400 response. Nonetheless, by using these contrasting words as experimental targets, the local semantic and the social contexts became at odds. Whereas the local semantic context favors the processing of a word such as *overcooked* (i.e., a word closer to reality according to the information provided earlier in the discourse, such as *burned*), our social knowledge favors the encountering of a word such as *tasty* or a similar complimenting adjective, given the context of a social interaction where compliments are rather appropriate. An overt semantic violation was also included in our experimental design (e.g., the meat sauce was *romantic*) in order to obtain a

robust N400 effect, against which any potential N400 enhancements could be contrasted.

As outlined earlier, in contrast to fact-related lies, prosocial lies are morally acceptable and might not be unexpected in social scenarios. Our study thus uses ERP measures to examine how processing unfolds for verbal information that is factually inaccurate but socially sanctioned (white lies) compared to factually true but impolite information. If white lies are processed as the factual lies in the Dutch white trains study (Hagoort et al., 2004; Hald et al., 2007) they would raise the N400 ERP component. However, if our social knowledge anticipates a white-lying communicative behavior, they would fail to increase the amplitude of the N400 component. Alternatively, they could be processed as ironic messages mismatching reality, thus giving rise to late latency ERP effects. On the other hand, factually true but impolite statements might clash with anticipated verbal outcomes and thus elicit a larger N400 relative to white-lying statements, a frontal late positivity, or both.

Based on the finding that men and women tend to differ in semantic processes indexed by the N400 component (Daltrozzo, Wioland, & Kotchoubey, 2007; Wirth et al., 2007) and particularly on the social or pragmatic aspects of language processing (i.e., empathy; van den Brink et al., 2012), for this initial study we limited the sample to female participants.

Method

Participants

Twenty-seven female native Spanish speakers (mean age = 22.6 years, range = 18–30 years) volunteered to participate in the study. All participants gave written informed consent. Twenty-five participants reported being right-handed and two being left-handed. The average handedness score (Oldfield, 1971) for the right-handed participants was +78.3 (range, 43 to +100). All participants reported normal or corrected-to-normal vision and none had a history of neurological or psychiatric disorders. The study was approved by the local ethics committee.

Materials

Target words used in our study were selected from a cloze-probability (CP) norming study. The purpose of the study was twofold: (1) to measure the likelihood of producing a white lie or an inconvenient truth as a sentence continuation in a paper-and-pencil test and (2) to obtain equally expected target words for each condition, as it is well established that there is an inverse correlation between CP and N400 amplitude, with lower CP words eliciting larger N400 amplitudes (Kutas & Hillyard, 1984). The final purpose was to avoid potential confounds of N400 effects simply due to higher or lower word expectancies rather than to the quality of social appropriateness. In addition, we used a wide range of social scenarios (N = 93) in order to avoid target word repetitions, also known to decrease N400 amplitudes (see Kutas & Federmeier, 2000, for a review).

An initial set of experimental stimuli was created. It consisted of 100 two- or three-sentence-long paragraphs depicting a social scenario followed by an incomplete sentence up to the point at which either a white lie or a blunt truth was most likely to continue the sentence. The norming procedure examined the cloze probability values (CP) of plausible word continuations to these sentences. Sixty students completed one of two lists (30 participants per list), each of which comprised 50 experimental scenarios. Participants were asked to read each paragraph and the following incomplete sentence, and then write down the word they would generally expect to find continuing the sentence fragment.

We selected a final set of 93 experimental scenarios such that the critical words for inclusion in the ERP experiment (either as the white lying or the blunt truth-telling condition) were matched in CP value. For example, for the experimental item: "*My grandfather was losing his abilities day by day, and it was hard for him to move by himself. My mother told him: I find yourself quite …,"* the word *well* had the highest CP value (39 %). Nonetheless in the ERP study we used either the word *agile*, as a white-lying target (7.1 % CP) or the word *weaken* (7.1 %), as a blunt-true target. By using this approach, mean CP values for white-lie (WL) and blunt-truth (BT) target words did not differ on average, 6.5 %, $t(184) = 0.236$, $p = .814$. See Table 1 for details on target words cloze probability, frequency of use, and length.

In order to better characterize our scenarios, besides this classical cloze probability measurements for a particular word entry we coded our sentences for their probability to induce any kind of word that would render a socially appropriate statement, regardless of the particular word item used. For that purpose, we overall classified our participants' responses as white lies (59 %) or blunt-truth (19 %) continuations. Remaining responses (22 %) were rendered unclassifiable in socially appropriateness terms (e.g., pronouns, articles, and determiners). Among the final pool of 93 sentences, a subset of 34 sentences were classified as highly white lie (WL)-inducing sentences (77 to 100 %), 27 sentences had a 53 to 73 % probability for WL sentence continuations, and 32 sentences had a 47 % or lower WL probability. We categorized this variable as "tendency to white lying" in paper-and-pencil tests.

Additionally, we conducted a survey with 40 participants over the 93 paragraph–sentence pairs to be used in the ERP experiment to verify whether the final sentence was considered socially appropriate or inappropriate. The majority of

Table 1 Cloze probabilities, frequency of use, and number of letters of target words in the ERP study

	Cloze probability of target word (%)			Frequency of use (occurrences per million)			Length (# letters)		
	Mean	Range	*SD*	Mean	Range	*SD*	Mean	Range	*SD*
All sentences ($N = 93$)									
White lie	6.7	0–46.7	7.8	396.8	0–5454	898.5	7.2	2–14	2.3
Blunt truth	6.4	0–63.3	11.1	532.5	0–20028	2171.9	7.4	2–13	5.4
Semantic violation	0.0	–	–	273.0	0–3350	601.5	7.1	3–12	1.8
Sentences highly biased to white lying ($N = 34$)									
White lie	7.1	0–46.7	9.1	409.9	0–5454	1013.3	7.0	2–12	2.2
Blunt truth	2.7	0–20.0	3.9	937.6	2–20028	3468.6	6.2	2–12	2.2
Semantic violation	0.0	–	–	288.0	5–2240	461.5	6.9	4–10	1.5
Sentences moderately biased to white lying ($N = 27$)									
White lie	5.9	0–26.7	6.0	262.1	0–3228	623.8	8.1	4–14	2.6
Blunt truth	5.8	0–23.3	5.9	366.4	0–3570	759.0	7.7	2–13	2.9
Semantic violation	0.0	–	–	248.1	0–3350	647.4	7.3	4–11	1.9
Sentences lowly biased to white lying ($N = 32$)									
White lie	7.0	0–33.3	7.8	496.7	0–3744	974.9	6.7	4–11	2.0
Blunt truth	10.7	0–63.3	16.9	242.1	0–3521	646.4	7.3	4–12	2.3
Semantic violation	0.0	–	–	278.1	0–3948	703.5	7.3	3–12	1.9

white-lying statements (95.7 %) were considered socially appropriate. Most blunt-true (BL) statements (94.6 %) were considered socially inappropriate according to this survey.

During the ERP experiment, after reading the main paragraph, the following sentence randomly included a WL, a BT, or a semantic violation (SV; zero CP value). All target words were open-class words (i.e., verbs, adjectives, nouns). Table 2 shows examples of experimental stimuli. Three experimental lists were constructed such that each sentence was assigned to a list in order to avoid sentence repetition effects. Frequency of use (Sebastián-Gallés, Martí, Carreiras, & Cuetos, 2000) and length of target words was matched within lists and across conditions. Cloze probability was significantly different between WL and SV continuations (6.7 % and 0, respectively), $t(184) = 8.281$, $p = .001$, and between BT and SV continuations (6.4 % and 0, respectively), $t(184) = 5.536$, $p = .001$, but, as mentioned earlier, mean CP values for selected white lie (WL) and blunt truth (BT) target words did not differ (6.5 %), $t(184) = 0.236$, $p = .814$. Stimuli were randomized within each list such that there were no more than five consecutive items of the same experimental condition. Each list was randomly assigned to participants. During the ERP recording session participants just read for comprehension, without any additional task, as we were not interested in biasing them toward a moral judgment on whether lying or telling the truth was more or less convenient. We were interested in how their brains processed both white lies and true statements naturally, in the absence of any additional task other than reading for comprehension. However, in order to ensure their attention to the reading materials, 10 two-choice common questions were randomly included along the experimental session. For example, in the following experimental item: *It is the third time that Nacho gets a necktie as his birthday present and he almost never wears a necktie. When he takes off the wrapping paper and sees the gift, he says: "Thank you, I am enthusiastic/disappointed/shaved about this gift,"* the probe question was: *Does Nacho usually wear a necktie, Yes or No?* Participants responded aloud and the experimenter marked her response as correct or incorrect.

Experimental procedure

After signing informed consent, participants were fitted with encephalogram (EEG) electrodes while they filled out handedness, vision, and health questionnaires. They were seated approximately 100 cm in front of a 19-in. computer monitor. The session began with a short set of practice stimuli to acclimate the participants to the reading task. Paragraphs were fully presented in the screen in a black 36-point lowercase Arial font on a white background. Participants were instructed to read the paragraph and press a button to initiate the continuing sentence. The final sentence was then presented one word at a time in the center of the screen. Each word was presented for 300 ms with an interwords interval of 300 ms. Figure 1 shows the sequence of events. Lists were divided into three blocks of 31 paragraphs each, to allow participants some rest. The participants proceeded from one block to the next at their own pace. The whole reading session lasted about 20 to 30 minutes, including breaks. Participants were highly

Table 2 Examples of stimuli presented in the ERP study (translated from Spanish)

Paragraph	Sentence beginning	Target word[1]	Cloze[2]	Exptal condition	Sentence ending
Ana doesn't know how to cook and the meal she prepared for her guests got burned. As they finish having dinner, she asks: So, what do you think of dinner? One of her guests says:	The meat sauce was …	*tasty*	10.3	white lie	… and it was creamy.
		overcooked	10.3	blunt truth	
		romantic	0	nonsense	
Sally has a complex about her terrible American accent when she speaks Spanish. The other day she asked Jaime whether he understood her and he said:	Your accent is very …	*attractive*	3.3	white lie	… to try to understand you.
		noticeable	3.3	blunt truth	
		yellow	0	nonsense	
The relationship didn't last and Esther was never very much involved. After a few weeks she had another boyfriend. When her ex-boyfriend asked her how she was doing, she said:	You know that I never …	*will forget you*	33.3	white lie	… deep inside.
		loved you	26.7	blunt truth	
		rented you	0	nonsense	
Inés has put on too much make-up and she looks older. When she asks her friend how does the make-up look on her, her friend says:	The make-up highlights your …	*cheeks*	6.7	white lie	… and it matches your skin tone.
		wrinkles	6.7	blunt truth	
		flowers	0	nonsense	
Jaime has decided to put on some hair gel and comb his hair back because he thinks it looks good on him. When he asked his friend Juanjo what he thought of his new look, Juanjo said:	That new hair style highlights your …	*cheeks*	3.3	white lie	… and your forehead.
		receding hairline	6.7	blunt truth	
		months	0	nonsense	
Sandra has put on some weight this summer due to the medication she is on. When she asks her friend Sonia whether a tight-fitting dress suits her well, her friend says:	That dress makes you look …	*thinner*	20.0	white lie	… than before.
		more fat	16.7	blunt truth	
		written	0	nonsense	

[1] Target words always consisted of a single word in Spanish

[2] Cloze probability percentage as determined by a paper-and-pencil norming study

accurate in answering the probe questions during the experimental session (99 % hits).

EEG recording and analyses

EEGs were recorded from 31 tin electrodes mounted in an electrode cap (Electro-Cap International, Eaton, Ohio, USA). Electrode impedances were kept below 5 KΩ. Electrodes were referenced online to the left mastoid, amplified with Brain Amps amplifiers (Brain Products, Munich, Germany) at a sampling rate of 250 Hz with a bandpass of 0.01 to 40 Hz, and rereferenced off-line to the mastoids average.[1] Bipolar horizontal and vertical electrooculograms (EOGs) were recorded for artifact rejection and blink correction purposes using the Gratton, Coles, and Donchin (1983) method. Data were processed using BrainVision Analyzer software (Brain Products, Munich). After visual inspection of individual data files, the following artifact threshold criteria were set: maximal allowed voltage step, 50 μV; minimal and maximal allowed amplitude, ±100 μV; lowest allowed activity (max–min), 5 μV for a 1,500-ms interval length. Once any threshold was met in the continuous EEG file, data recorded at that point were marked and discarded, together with data recorded during the 200 ms before and after the detected artifact. This was done to avoid including any residual artifacts in subsequent computations of ERP averages. EEG raw data from all subjects were scanned and marked using the same criteria. As a result, 14.6 % of trials were discarded and an average of 26.5 trials remained per experimental condition.

A Butterworth zero phase filter was applied to the EEG data (low cutoff at 0.1 Hz, time constant = 1.6 s, 24 dB/oct; high cutoff at 20 Hz, 24 dB/oct). The continuous EEG was segmented into 1,000-ms epochs starting 100 ms before the onset of the target word. Artifact-free subject weighted average waveforms were calculated for each word continuation (WL, BT, SV) separately, after subtraction of the prestimulus baseline.

Based on visual inspection and previous ERP literature, mean amplitude values in the N400 (300–500 ms) and P600 (550–800 ms) latency ranges were submitted to repeated-measures analyses of variance (ANOVAs) involving the within-subjects factors sentence continuation (three levels: WL, BT, SV) in four quadrants (left anterior: FP1, F7, F3, FT7, FC3; right anterior: FP2, F8, F4, FC4, FT8; left posterior: CP3, TP7, P7, P3, O1; right posterior: CP4, TP8, P4, P8, O2), as well as midline sites (Fz, FCz, Cz, CPz, Pz). All *p* values are reported with the Greenhouse–Geisser correction for repeated measures with more than 1 degree of freedom. Effect sizes were computed using the partial eta-square (η_p^2) method. Relevant pairwise comparisons are reported using Bonferroni correction for multiple comparisons. All statistical analyses were carried out using IBM SPSS Statistics (Version 22).

[1] Electrode sites included: Fp1/z/2, F7/3/z/4/8, FT7/8, FC3/z/4, T7/8, C3/z/4, TP7/8, CP3/z/4, P7/3/z/4/8, O1/z/2, and right mastoid.

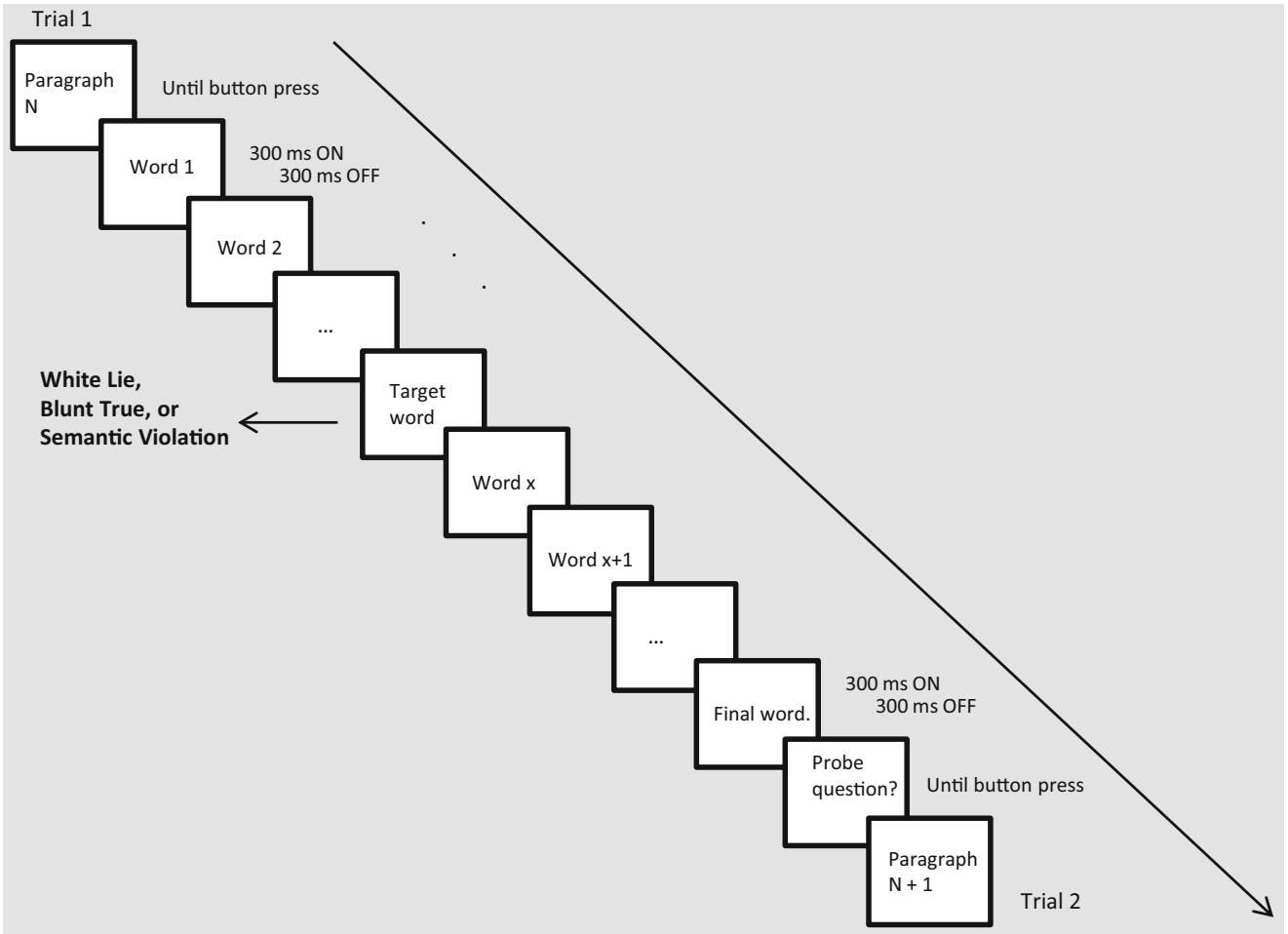


Fig. 1 Scheme of the experimental design

Results

Figure 2 displays the grand average ERPs time-locked to the onset of the critical word: semantic violation (red lines), blunt truth (blue lines), and white lie (black lines), at nine representative electrodes. All sentence continuations evoked a negativity in the N400 latency range. This negativity was larger for SV relative to WL and BT conditions. No subsequent parietal positive deflections were observed. However, over frontal electrodes the BT condition showed a larger positivity, starting at around 500 ms.

Mean amplitude ANOVA in the 300-ms to 500-ms latency range revealed a main effect of sentence continuation, $F(2, 52) = 27.04$, $p < .001$, $\eta_p^2 = .51$. Planned comparisons showed that SVs elicited more negative going brainwaves than BT (-1.36 vs. 1.32 μV, $p < .001$) and WL continuations (-1.36 vs. 1.63 μV, $p < .001$). No differences emerged in the amplitude of the N400 response between WL and BT conditions ($p = 0.1$). The interaction between sentence continuation and quadrant, $F(8, 208) = 7.01$, $p < .001$, $\eta_p^2 = .21$, was further explored. The larger N400 for SV than WL continuations was significant at all quadrants and at midline sites (all ps < .001). The larger N400 for SV than BT continuations was also significant at all regions (all ps < .005). No significant differences emerged in N400 amplitude between WL and BT continuations within each region (all ps > .30).

The mean amplitude ANOVA on the 550-ms to 800-ms time range also revealed a sentence continuation effect, $F(2, 52) = 4.10$, $p = .028$, $\eta_p^2 = .14$; a larger positivity for BT relative to WL sentence continuations (4.13 vs. 2.51 μV, $p = .006$). The interaction with quadrant was significant, $F(8, 208) = 4.67$, $p = .007$, $\eta_p^2 = .15$, and follow-up ANOVAs were carried out at each region. In the two anterior (right and left) and the left posterior regions, BT sentence continuations elicited a larger positivity than WL continuations (all ps < .018). In addition, BT continuations elicited a larger positivity than SVs at the right frontal region (3.59 vs. 1.70 μV, $p = .027$).

Analyses including tendency to white lying in paper-and-pencil tests as a factor

Despite the fact that cloze probability was matched across WL and BT target words, 34 out of the 93 scenarios were classified

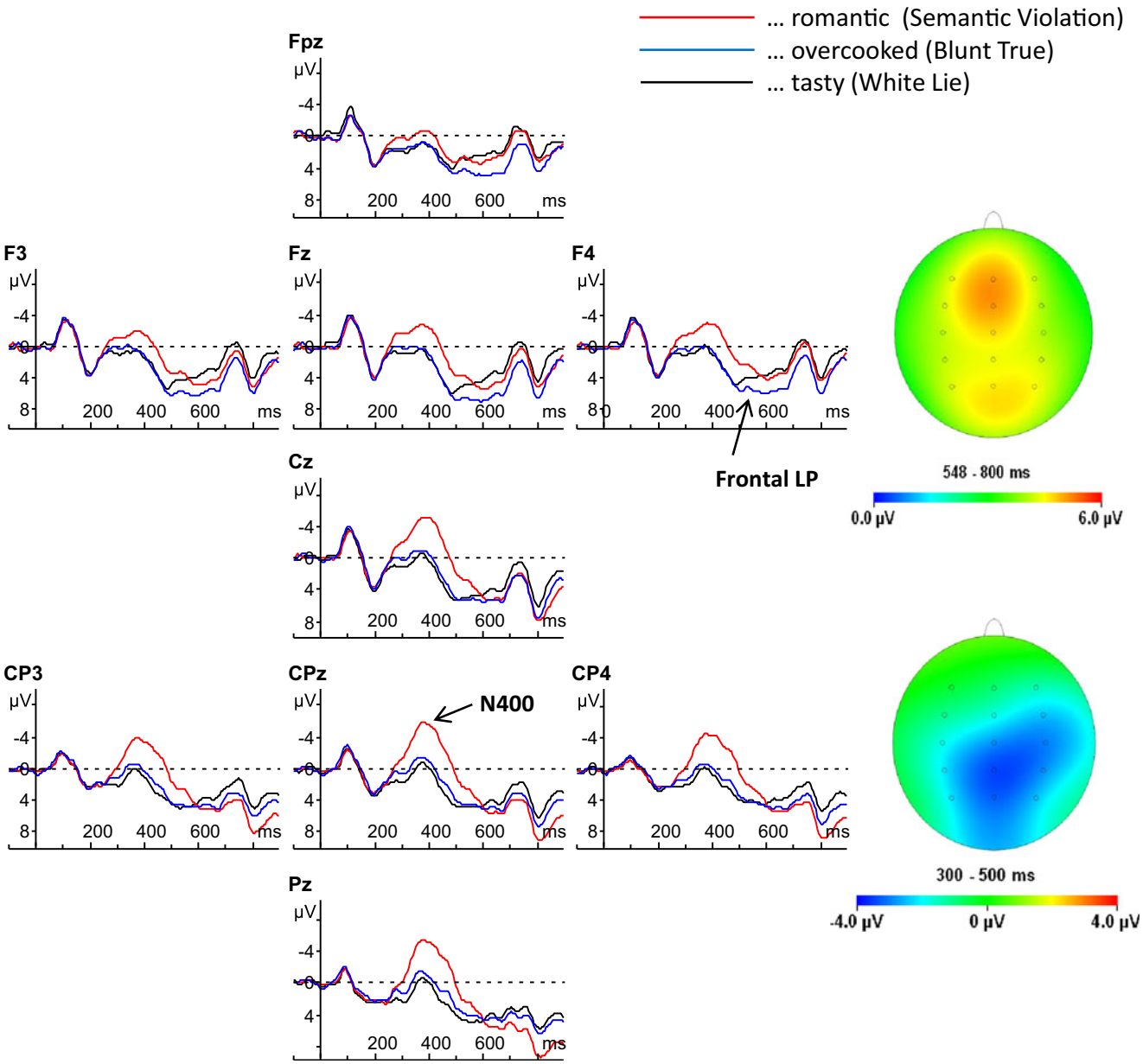


Fig. 2 ERP responses to target words that conveyed a: semantic violation (red), blunt true statement (blue), or white lie (black). Responses are plotted at a representative selection of nine electrodes. Frontal sites are at the top; parietal sites at the bottom of the figure. Semantic violations elicited an N400 response most prominent at centro-parietal electrode sites (see voltage map). Over frontal sites, blunt true statements generated a late positivity (see voltage map). (Color figure online)

as having a strong bias to produce white-lying sentence continuations (77 % to 100 %). For this subset of sentences, the target word in the WL condition had in fact a higher CP (7.1 %) than the target word used for the BT condition (2.7 %), $t(66) = 0.2533$, $p = .014$. In 27 sentences that showed a moderate (53 % to 73 %) probability for white-lying continuations, the cloze probability of BT and WL continuations was matched (5.8 % and 5.9 %, respectively), $t(52) = 0.0.08$, $p = .937$. Finally, in the remaining 32 sentences, with 47 % or lower probability for WL continuations, there were no CP differences between BT and WL targets (10.7 % and 7 %, respectively), $t(62) = -1.135$, $p = .261$.

Figure 3 shows the grand average ERPs time-locked to the onset of the critical word for each of these subset of sentences:

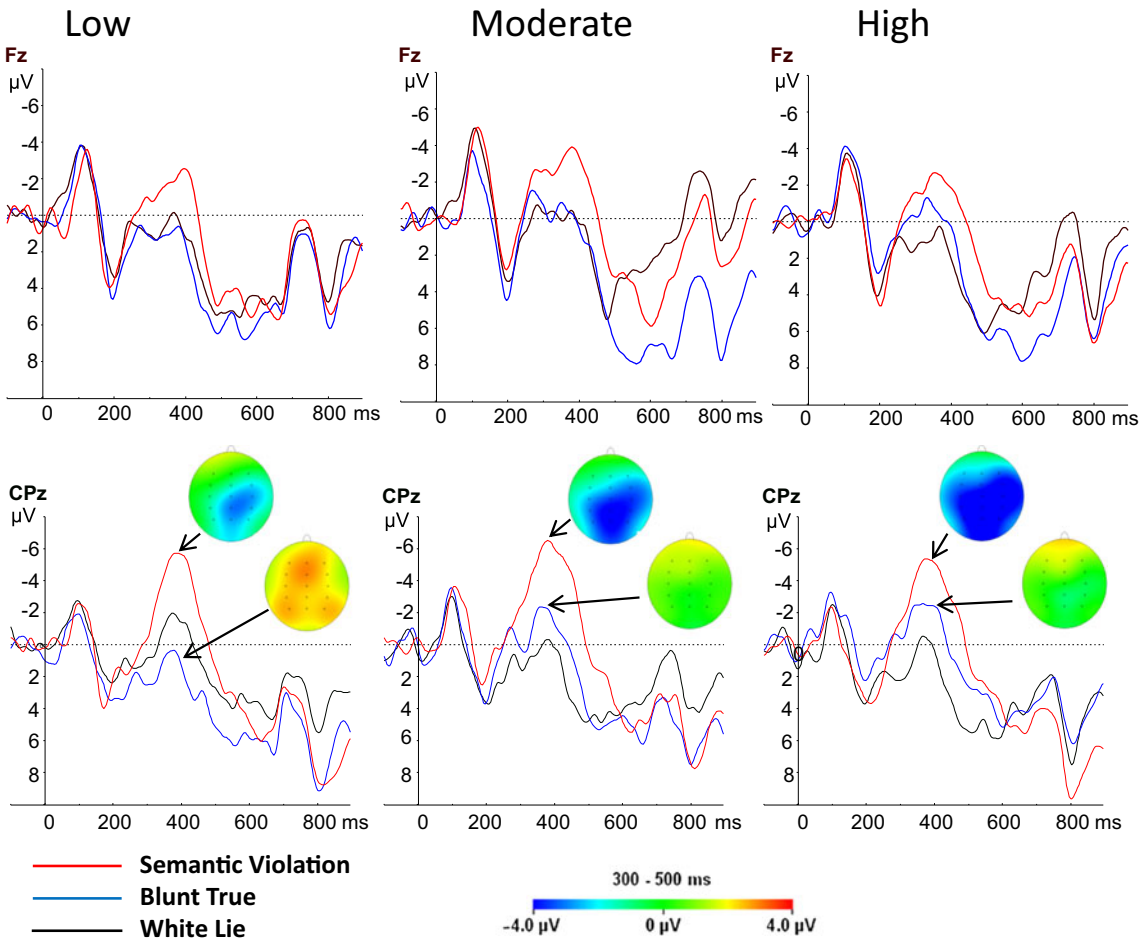


Fig. 3 Electrophysiological responses to our reading materials as a function of their tendency to induce a white lying sentence continuation (low, moderate, or high) in a paper-and-pencil test, at Fz and CPz electrodes. As we saw earlier in Fig. 2, semantic violations (red) always elicited the highest N400 response. Critically, now blunt true statements (blue) also elicited an N400 response when the tendency to white lying was medium or high (see CPz electrode). Despite the fact that the N400 was smaller in amplitude than the one elicited by semantic violations, its scalp distribution was similar (see voltage maps). (Color figure online)

semantic violation (red lines), blunt truth (blue lines), and white lie (black lines), at a frontal electrode (Fz) and at a N400-representative centro-parietal electrode (CPz). For sentences highly and moderately biased for white lying, not only semantic violations but also blunt-true statements seemed to evoke a larger negativity in the N400 latency range, relative to white-lying conditions.

To verify this impression, we conducted statistical analyses including the variable tendency to white lying as a factor, with three values (high, moderate, and low). The mean amplitude ANOVA in the 300-ms to 500-ms latency range confirmed a main effect of sentence continuation, $F(2, 52) = 26.44$, $p < .001$, $\eta_p^2 = .50$, and interactions between sentence continuation and quadrant, $F(8, 208) = 7.23$, $p = .001$, $\eta_p^2 = .22$. A three-way interaction arose between sentence continuation, tendency to white lying, and quadrant, $F(16, 416) = 2.85$, $p = .009$, $\eta_p^2 = .01$. Overall, planned comparisons revealed that SV were more negative than WL and BT continuations (SV: -1.39, BT: 1.24, and WL: 1.53 μV, $p < .001$). However, separate analyses conducted at each region, revealed a significant interaction between sentence continuation and tendency to white lying, $F(4, 104) = 3.73$, $p = .017$, $\eta_p^2 = .13$, over the right posterior region, where the N400 component typically reaches its maximal amplitude. In the left posterior region, the interaction only approached significance ($p = .093$), whereas in frontal and midline quadrants the interaction was not significant ($p \geq .12$).

We conducted separate ANOVAS and pairwise comparisons on the right posterior region. Sentences with a low tendency to white lying showed a larger N400 for SV relative to BT statements ($p = .001$). Sentences with a moderate tendency to white lying showed a three-way split pattern, with SVs showing larger N400s than BT and WL statements ($ps = .001$) and BTs larger N400 than WLs ($p = .040$). Finally, for sentences highly biased to white lying, both SV and BT statements showed larger N400s than WL statements ($p = .001$ and $p = .038$, respectively). The N400 response for SV and BT statements was not significantly different ($p = .19$) for this subset of sentences.

Discussion

As expected, a classic N400 incongruity effect was obtained for semantic violations. However, and in contrast to the N400 effect obtained for the processing of false statements (Hagoort et al., 2004; Hald et al., 2007), lies in our study (social lies, such as, "The meat sauce was *tasty*") did not provoke an N400 enhancement in spite of their low cloze probability as lexical ending targets. It is critical to distinguish between lexical constraints (i.e., the anticipation of a word target based on traditional cloze probability norms) and social constraints. Since white lies in our study were embedded in a broader social context, the N400 elicited by them was even smaller than that to a word that bears a message closer to truth but socially inappropriate (e.g., prior paragraph: dinner got burned … target sentence: the meat sauce was *overcooked*). Thus, white lying (such as: it was *tasty*) does not interfere with semantic processing within a social discourse context, and it even becomes easier to process than semantically closer-to-truth impolite statements. In line with recent views on what the N400 indexes (online anticipation processes; see a review by Kutas & Federmeier, 2011), white lies may even be actively anticipated. In this regard, the social constraints to lie that we manipulated in our study seem to overrule rather pure lexical constraints.

In addition, and in contrast to ironic messages (Regel et al., 2010, 2011), white lies do not evoke any subsequent activity in the P600 time window, neither parietal nor frontally distributed. The visual presentation of our sentences entails a lack of prosodic cues about an ironic interpretation. Thus, in line with current views on the functional significance of P600 parietal effects (Brouwer, Fitz, & Hoeks, 2012; Kuperberg, 2007) white lying, in contrast to irony, does not appear to incur in sentence reanalysis or reinterpretation processes. In addition, the absence of frontally distributed late positivities indicates that white lies do not to incur in a cost for alternative predicted outcomes (Federmeier et al., 2007).

On occasions, the violation of communicative social norms (i.e., the use of blunt-truth target words) was in fact what elicited an N400 response. This occurred for those sentences that were moderately and highly biased to white-lying continuations in paper-and-pencil tests. Word cloze probability was a confounding factor for sentences highly biased to white lying. Thus, blunt-true continuations elicited a larger N400 response, but they also had a lower CP value relative to WL continuations. By contrast, this CP confound was not present for the moderately biased sentences. For this subset of sentences, both WL and BT target continuations had a similar CP value. Nonetheless an N400 enhancement for BT relative to WL targets was found. Therefore, BT statements elicited a larger N400 than WL statements beyond their cloze probability value. These results reinforce the view that the N400 is indicative of a difficulty of processing for words that cannot possibly be anticipated based on an established social world-knowledge.

Interestingly, considering the whole verbal materials, we also observed a frontal late positivity in response to blunt-truth statements. Implausible sentence continuations (e.g., semantic violations) do not show this frontal effect (Federmeier, Kutas, & Schul, 2010; Moreno & Rivera, 2014). According to what these frontally distributed late positivities index (DeLong, Quante, & Kutas, 2014), blunt-true statements were treated as plausible but unexpected events. The uniqueness of this frontal effect in response to blunt true statements indicates that brains responded particularly to socially inappropriate statements.

In sum, not only are we able from early childhood to make acceptable moral judgments about lie telling in social situations (Ma et al., 2011), but based on electrophysiological responses our study reveals that the processing of white lies lacks of any semantic (N400) or interpretative (P600) difficulty. White lies are processed neither as false nor as ironic messages. Their immersion in a social context overrules the neural consequences that have previously been linked to the processing of factually untrue statements. Indeed, closer to truth statements became more difficult to process, as indexed by an enhanced N400 for those sentences strongly/moderately biased toward white lying, as well as by a frontal late positivity in all cases.

Our study favors the inclusion of a pragmatic aspect such as social impact on language comprehension tasks.

The conclusions of our study are limited to female readers. Future studies are needed to better disentangle the partial contributions of cloze probability (expectancy) and social adequacy factors to the amplitude of the long-standing N400 effect.

Conclusion

Theories on the pragmatics of language comprehension ought to consider that our knowledge stored in long-term memory most likely includes rules of social communicative behavior. Our study shows that electrical brainwaves are not only sensible to whether words in context are lexically (un)expected, they are also sensible to how convenient or inconvenient a verbal statement is in a situational context, as in those scenarios in which saying a sweet little lie is preferable than saying an inconvenient truth.

Acknowledgments This research was supported by Spanish Ministry of Science and Innovation grants PSI2008-04961 and PSI2014-60682-R to EMM, and PSI2013-43107-P to PC and MML. We thank all the study participants for their cooperation.

References

Boudewyn, M. A., Long, D. L., & Swaab, T. Y. (2015). Graded expectations: Predictive processing and the adjustment of expectations during spoken language comprehension. *Cognitive, Affective, & Behavioral Neuroscience, 15*(3), 607–624.

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.

Daltrozzo, J., Wioland, N., & Kotchoubey, B. (2007). Sex differences in two event-related potentials components related to semantic priming. *Archives of Sexual Behavior, 36*(4), 555–568.

DeLong, K. A., Quante, L., & Kutas, M. (2014). Predictability, plausibility, and two late ERP positivities during written sentence comprehension. *Neuropsychologia, 61,* 150–162.

DePaulo, B. M., Kirkendol, S. E., Kashy, D. A., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70*(5), 979–995.

Federmeier, K. D., 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.

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research, 1146,* 75–84.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55*(4), 468–484.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304*(5669), 438–441.

Hald, L. A., Steenbeek-Planting, E. G., & Hagoort, P. (2007). The interaction of discourse context and world knowledge in online sentence comprehension. Evidence from the N400. *Brain Research, 1146,* 210–218.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, 1146,* 23–49.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science, 4*(12), 463–470.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62,* 621–647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203–205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*(5947), 161–163.

Ma, F., Xu, F., Heyman, G. D., & Lee, K. (2011). Chinese children's evaluations of white lies: Weighing the consequences for recipients. *Journal of Experimental Child Psychology, 108*(2), 308–321.

Moreno, E. M., & Rivera, I. C. (2014). Setbacks, pleasant surprises and the simply unexpected: Brainwave responses in a language comprehension task. *Social Cognitive and Affective Neuroscience, 9*(7), 991–999.

Nieuwland, M. S. (2015). The truth before and after: Brain potentials reveal automatic activation of event knowledge during sentence comprehension. *Journal of Cognitive Neuroscience, 27*(11), 2215–2228.

Nieuwland, M. S., & Kuperberg, G. R. (2008). When the truth is not too hard to handle: An event-related potential study on the pragmatics of negation. *Psychological Science, 19*(12), 1213–1218.

Nieuwland, M. S., & Van Berkum, J. J. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience, 18*(7), 1098–1111.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113.

Otten, M., Nieuwland, M. S., & Van Berkum, J. J. (2007). Great expectations: Specific lexical anticipation influences the processing of spoken language. *BMC Neuroscience, 8,* 89.

Proverbio, A. M., Vanutelli, M. E., & Adorni, R. (2013). Can you catch a liar? How negative emotions affect brain responses when lying or telling the truth. *PLoS ONE, 8*(3), e59383.

Regel, S., Coulson, S., & Gunter, T. C. (2010). The communicative style of a speaker can affect language comprehension? ERP evidence from the comprehension of irony. *Brain Research, 1311,* 121–135.

Regel, S., Gunter, T. C., & Friederici, A. D. (2011). Isn't it ironic? An electrophysiological exploration of figurative language processing. *Journal of Cognitive Neuroscience, 23*(2), 277–293.

Regel, S., Meyer, L., & Gunter, T. C. (2014). Distinguishing neurocognitive processes reflected by P600 effects: Evidence from ERPs and neural oscillations. *PLoS ONE, 9*(5), e96840.

Rigoulot, S., Fish, K., & Pell, M. D. (2014). Neural correlates of inferring speaker sincerity from white lies: An event-related potential source localization study. *Brain Research, 1565,* 48–62.

Sebastián-Gallés, N., Martí, M. A., Carreiras, M., & Cuetos, F. (2000). *Lexesp: Una base de datos informatizada del español* (Lexesp: An Spanish computarized data base). Barcelona: Edicions de la Universitat de Barcelona.

Serota, K. B., Levine, T. R., & Boster, F. J. (2010). The prevalence of lying in America: Three studies of self-reported lies. *Human Communication Research, 36*(1), 2–25.

Spotorno, N., Cheylus, A., Van Der Henst, J. B., & Noveck, I. A. (2013). What´s behind a P600? Integration operations during irony processing. *Plos One, 8*(6). doi:10.1371/journal.pone.0066839

Thornhill, D. E., & Van Petten, C. (2012). Lexical versus conceptual anticipation during sentence processing: Frontal positivity and N400 ERP components. *International Journal of Psychophysiology, 83*(3), 382–392.

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*(3), 443–467.

van den Brink, D., Van Berkum, J. J., Bastiaansen, M. C., Tesink, C. M., Kos, M., Buitelaar, J. K., … Hagoort, P. (2012). Empathy matters: ERP evidence for inter-individual differences in social language processing. *Social Cognitive and Affective Neuroscience, 7*(2), 173–183.

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology, 83*(2), 176–190.

Wicha, N. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience, 16*(7), 1272–1288.

Wirth, M., Horn, H., Koenig, T., Stein, M., Federspiel, A., Meier, B., & Strik, W. (2007). Sex differences in semantic processing: Event-related brain potentials distinguish between lower and higher order semantic analysis during word reading. *Cerebral Cortex, 17*(9), 1987–1997.