Brain & Language 254 (2024) 105438

Contents lists available at ScienceDirect

Brain and Language

journal homepage: www.elsevier.com/locate/b&l

Registered Report Stage II

An electrophysiological investigation of referential communication

Veena D. Dwivedi [a,*], Janahan Selvanayagam [b,1]

[a] Department of Psychology/Centre for Neuroscience, Brock University, 1812 Sir Isaac Brock Way, St. Catharines, ON L2S 3A1, Canada
[b] Centre for Neuroscience, Brock University, 1812 Sir Isaac Brock Way, St. Catharines, ON L2S 3A, Canada

ARTICLE INFO

Keywords:
ERP
Language
N400
P200
Semantic
Reference
Sentence
Definite article, demonstrative determiner
Presupposition

ABSTRACT

A key aspect of linguistic communication involves semantic reference to objects. Presently, we investigate neural responses at objects when reference is disrupted, e.g., *"The connoisseur tasted *that wine"... vs. "...*that roof..."* Without any previous linguistic context or visual gesture, use of the demonstrative determiner *"that"* renders interpretation at the noun as incoherent. This incoherence is not based on knowledge of how the world plausibly works but instead is based on grammatical rules of reference. Whereas Event-Related Potential (ERP) responses to sentences such as *"The connoisseur tasted the wine ..."* vs. *"the roof"* would result in an N400 effect, it is unclear what to expect for doubly incoherent *"...*that roof..."*. Results revealed an N400 effect, as expected, preceded by a P200 component (instead of predicted P600 effect). These independent ERP components at the doubly violated condition support the notion that semantic interpretation can be partitioned into grammatical vs. contextual constructs.

1. Introduction

Reference is a key aspect of human communication. That is, both speakers and hearers (and/or writers and readers) assume a shared common ground of objects during communication (Stalnaker, 1978; Heim, 1982). Languages use numerous grammatical devices for the purposes of referring to both animate and inanimate objects (Roberts, 2002). For example, English uses names ("*Pat is an excellent art teacher*"), pronouns ("*She loves her easel*") and demonstrative determiners ("*That easel has been in her studio forever*") as tools of reference (King, 2006). These objects (names, pronouns, nouns with demonstrative determiners) are assumed to exist for both speaker and hearer; that is, they are "presupposed" to exist.

Reference can be disrupted[2] if a term is used that is unknown or unfamiliar to the hearer/reader in the context of utterance. For example, if *"Pat"* is mentioned in conversation and she is unknown to the listener, communication is disrupted. Next, starting a conversation with a pronoun, as in *"She loves her easel"*, without any previous context to indicate who the pronoun "*she*" would refer to, also results in incoherence. Similarly, in the absence of context or any visual/gestural cue of demonstration, using the determiner "*that*" in *"That easel has..."* would be difficult to understand, resulting in incoherence.[3] We note that the latter two devices for reference, pronouns and determiners, are closed class elements (also known as grammatical or function words) that contribute to a sentence's grammatical meaning. Thus, semantic context can refer to a shared discourse or text between individuals; and if/when an anaphoric term (e.g., *s/he, they*; *that* in *that easel; that car*) is used that has no explicit previous mention (i.e., no antecedent) –reference is disrupted. On a wider scale, semantic context can also refer to a shared understanding of how the world plausibly works. This shared understanding of how the world works is dependent on experience (and is independent of grammatical function words and/or structure). For example, "*The connoisseur tasted the roof*" is a nonsensical sentence vs. "*The connoisseur tasted the wine*". In the former sentence, *"roof"* does not violate grammatical expectations (it is a concrete noun in direct object position, as expected) nor is it violation of presupposed existence. Instead, the sentence is incoherent because *"roof"* does not fit the immediate sentence context. In other words, it violates our shared understanding of what a connoisseur might taste. The word is contextually implausible. For coherence to ensue, massive contextual adjustments

* Corresponding author.
E-mail address: vdwivedi@brocku.ca (V.D. Dwivedi).
1 Department of Experimental Psychology, University of Oxford, Tinsley Building, 13 Mansfield Rd, Oxford OX13SR, UK.

2 also known as presuppositional failure in linguistic theory.

3 On a related note, proper names (e.g., *Pat, Canada)* directly refer to specific entities in the world; in contrast, pronouns (e.g., *"she", "it", "they"*, etc.) are anaphoric, as are demonstrative determiners (e.g., *"that", "those"*) which refer to previously mentioned entities in discourse/text.

https://doi.org/10.1016/j.bandl.2024.105438
Received 18 April 2023; Received in revised form 13 June 2024; Accepted 19 June 2024
Available online 28 June 2024
0093-934X/© 2024 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

would need to be made (e.g., where *"the roof"* would be the name of a brand of wine[4]).

In ERP language studies, both kinds of aforementioned incoherence are typically labeled as 'semantic' violations; however, it is clear from the descriptions above that the violations arise from independent modules of knowledge of language. The lexical-experience based violation is derived from conceptual word-level semantics or world knowledge. The violation of referential meaning is derived from compositional semantic knowledge, i.e., grammatical knowledge.

In the present work, we asked the question: what is the neural response to sentences that are doubly incoherent according to both experience in the world, as well as grammatical devices of reference?

We expect that when sentences are incoherent due to experience in the world (e.g., *"The connoisseur tasted the roof"* vs. *"The connoisseur tasted the wine"*), an N400 response would be elicited. Since Kutas & Hillyard (1980; 1983), the N400 component has been hailed as the neural signature of (lexical) 'semantic anomaly'. This negative-going waveform generally peaks around 400 ms post-stimulus onset (Kutas & Federmeier, 2011), as in *"Jai spread the warm toast with #socks/butter."*[5].

Predictions for sentences that exhibit anomaly at both the lexical level (*"…tasted the #roof"*) and grammatical level (*"…tasted *that #roof"*) are less clear. However, based on previous work in our lab where reference was disrupted due to linguistic semantic context, we predict a P600 response (discussed below), in addition to an N400 for such double violations.

In previous work (Dwivedi et al., 2006; see also replication in Dwivedi et al., 2010) we examined ERP responses to incoherent sentences embedded in context. In that 2x2 study, the semantically anomalous condition (1c below) included the pronoun *"it"* which did not have an appropriate referent in the previous sentence (see Karttunen, 1976). That is, we compared two-sentence discourses such as in (1) below:

(1) a. John is considering writing a novel. It might end quite abruptly.
b. John is reading a novel. It might end quite abruptly.
c. John is considering writing a novel. #It ends quite abruptly.
d. John is reading a novel. It ends quite abruptly.

In (1a), the use of a modal auxiliary "*might*" in "*It might end…*" indicates that the hypothetical nature of the first sentence "*John is considering….*" is continued in the second sentence. As a result, the pronoun "*it*" can successfully co-refer with its antecedent, (hypothetical) "*a novel*" in the first sentence. In contrast, in (1c), the lack of a modal auxiliary in *"It ends …."* means that the sentence (and therefore the pronoun) is not hypothetical—in contrast to its antecedent. This results in referential disruption or anomaly–since now the anaphoric pronoun "*it*" is asserted to exist but is linked with the antecedent "*a novel*" which was hypothetical. Results showed that this intuitive contrast for sentence (1c) (vs. its control 1d) did result in an ERP contrast: a frontal P600 effect (Osterhout & Holcomb, 1992; Hagoort et al., 1993; Kaan et al., 2000; Kaan & Swaab, 2003a, 2003b); no such P600 effect was in evidence for sentence (1a) vs. its control (1b). This positive-going waveform, usually peaking in the 600 ms range had been typically associated with difficulty in structural integration of a word into a sentence (e.g., *"The spoilt child *__throw__ the toys; The broker persuaded *__to__ conceal the transaction…"*). We interpreted our finding as a 'semantic' P600 effect, (see also, Aurnhammer et al., 2021; Bornkessel-Schlesewsky & Schlesewsky. 2008; van Herten et al., 2006; Kuperberg, 2007). We note that the latter two devices for reference, pronouns, and determiners, are closed class elements (also known as grammatical or function words) that contribute to a sentence's grammatical meaning, (among others), where in our case, the structure that was difficult to integrate was the (semantic) discourse representation structure (Dwivedi, 1996; Roberts, 1989). Moreover, given that we found frontal positivity, which is associated with revision of structure (Kaan et al., 2000; Kuperberg et al., 2020), we speculated that the frontal positivity was a re-interpretation of the noun phrase "*a novel*". That is, rather than as a non-specific, hypothetical novel (that does not yet exist), perhaps the noun phrase was re-interpreted as a specific novel, e.g., "*John is considering writing a (specific) novel. It ends quite abruptly (given what we know about John and his tendencies)*".[6] Given that the meaning violation arose via grammatical rules governing the interpretation of pronouns and other closed-class elements such as modal auxiliaries (*might, should, would*, etc.) in discourse structure (Dwivedi, 1996; Kaplan, 1978; Karttunen, 1976; Roberts, 1996; Heim & Kratzer, 1998)—we called this a violation of compositional semantics, or a semantic P600 effect (see also van Herten et al., 2006, van Berkum, 2009; among others for other claims regarding semantic P600 results).

Given these previous findings, we hypothesize a similar semantic P600 effect in the current experiment, when reference is again incoherent—albeit now due to anaphoric nature of the determiner *"that"*. Thus, in addition to N400 ERP responses expected at semantically anomalous critical words such as *roof* (vs. *wine*), as in (i) "*The connoisseur tasted the # roof on the tour"* vs. (ii) *"The connoisseur tasted the wine…"* we compared responses at these conditions to sentences using demonstrative determiners such as (iii) *"The connoisseur tasted *that #roof…"* vs. (iv) *"The connoisseur tasted *that wine…"* For the double violation condition **that #roof…* we expected an N400-P600 effect at the critical word. Following the logic above, for the condition **that wine…* a P600 effect was also expected at semantically congruent "*wine*".

On the other hand, we note that in our previous work, we examined pronouns embedded in sentences in discourse, in contrast to the current experiment which used single sentences only. In addition, we note that the current experiment is exploratory since little is known whether empirical effects are found (if any, and if so, under what conditions) when determiners, such as the demonstrative *"that"*,[7] are used with nouns that lack appropriate referents in context, vs. definite determiner *"the"*[8] (see Murphy, 1984; Anderson & Holcomb, 2005 for experimental work on *"the")*. Most cognitive neuroscientific investigations of anaphoric elements have focused on pronouns with and without appropriate antecedents (see, among others, Filik et al., 2008; Hammer et al., 2005; Osterhout and Mobley, 1995; Nieuwland and van Berkum,

[4] For example, the Niagara winery Creekside has an excellent brand of wine called "Red Tractor". If we were at the winery and someone said *"The connoisseur tasted the Red Tractor (before the Backyard Bubbly)…"* this would be perfectly coherent.

[5] Note the symbol "#" is borrowed from semantic theory, marking *infelicity*. That is, this sentence is considered grammatically well-formed, but does not make sense from a contextual point of view (Katz & Fodor, 1963; Roberts, 1989; see also Magidor, 2019). In contrast, the asterisk "*" indicates a grammatical or structural violation (Chomsky, 1957).

[6] This re-interpretation is called presupposition accommodation in semantic theory (Stalnaker, 1978; Heim, 1982; Karttunen, 1976; see also von Fintel, 2008; Rullmann, 2003).

[7] The theoretical semantic framework assumed in this experimental work is that of formal semantics (in the tradition of Stalnaker, 1978; Heim, 1982; also see Heim & Kratzer, 1998) in contrast to cognitive semantic frameworks. For a review of recent work on deictic and spatial demonstratives, especially in the cognitive semantic framework see the recent volume by Diessel et al. (2021). Furthermore, we note that in assuming this framework, we are assuming that a lack of a referent for the noun phrase due to demonstrative determiner is a violation of grammatical expectation (vs. that of pragmatics) since the meaning of *"that Noun"* typically presupposes the existence of the object in context. Further clarification regarding the semantic/pragmatic interface goes beyond the scope of this paper, and so we will not discuss further (especially given that the empirical predictions would remain unchanged).

[8] We note here that *"the"* is a weak definite determiner and therefore does not require an anaphoric search (Roberts, 2003; Schwarz, 2009).

2008). We build on those findings here by examining potential interpretive costs at the direct object position when it is preceded by the demonstrative vs. definite determiner. Differing empirical effects are expected at the direct object "*roof*" vs. "*wine*" when preceded by "*that*" vs. "*the*" because noun phrase interpretation is compositional. That is, the interpretation of the sentence *"The apples on the table are delicious"* differs from *"All apples on the table are delicious"* due to meaning differences associated with *"the apples"* vs. *"all apples"* (Chierchia & McConnell-Ginet, 1990).

Thus, in this within-participants study, we examined two independent variables, direct object type (Plausible vs. Implausible) and determiner type (Demonstrative vs. Definite). See Table 1 for a list of the four conditions.

2. Methods

2.1. Participants

37 Brock University undergraduates were recruited and either paid for their participation or received partial course credit. All participants were native, monolingual speakers of English, had normal or corrected-to-normal vision and were right-handed, as assessed by the handedness inventory. No participants reported any neurological impairment, history of neurological trauma, or use of neuroleptics

Four participants with comprehension question accuracy for filler items (discussed below) at less than 85 % were excluded from analysis leaving 33 eligible participants (25 females; mean age = 19.6 years; ranging from 18 to 25 years).

This study received ethics approval from the Brock University Bioscience Research Ethics Board (BREB) prior to the commencement of the experiment (REB 13-282). Written, informed consent was received from all participants prior to their participation in the experiment.

2.2. Materials

Stimuli used here are described in Dwivedi & Selvanayagam (2021). In brief, these stimuli consisted of 160 critical items in four conditions (see Table 1) counterbalanced across four lists and 170 filler items. Critical items were simple active sentences with an animate subject (e.g., *the connoisseur*), an active past-tense verb (e.g., *tasted*), a determiner (definite: *the* vs demonstrative: *that*), an inanimate direct object (plausible: e.g., *wine* vs implausible: e.g., *roof*) and a prepositional phrase (e.g., *on the tour*).[9] Direct objects were not repeated, word length and frequency for direct objects in plausible vs implausible conditions were controlled for.

Table 1
Experimental conditions with example stimuli. Critical words are in bold and underlined.

Determiner type	Direct object type	Example Sentence
Definite determiner	**Implausible object**	(i) The connoisseur tasted *the* #<u>**roof**</u> on the tour.
	Plausible object	(ii) The connoisseur tasted *the* <u>**wine**</u> on the tour.
Demonstrative determiner	**Implausible object**	(iii) The connoisseur tasted **that* #<u>**roof**</u> on the tour.
	Plausible object	(iv) The connoisseur tasted **that* <u>**wine**</u> on the tour.

No tasks were associated with critical trials. 170 filler sentences were included in order to reduce predictability. Comprehension questions were asked at 125 filler items (38 % of all trials) consisting of superficial Yes/No or True/False questions. An example filler item and corresponding question is given below:

(2) After her surgery Anita slept for two days.

Q: Anita had a vacation. 1) True 2) False.

2.2.1. Offline plausibility ratings

We evaluated the plausibility of our critical materials by conducting a web-based norming study using Qualtrics software, Version (March 2020) of the Qualtrics Research Suite (Qualtrics, 2020). Participants were asked to rate sentences in terms of plausibility on a scale from 0 (very implausible) to 5 (neutral) to 10 (very plausible), in steps of 0.1. There was no time pressure as 16 sentences were presented on each webpage, for a total of ten pages. The 160 critical items were presented in eight pseudorandomized, counterbalanced lists such that half of the critical items were presented in each list and each participant only saw each item once. 80 filler items were presented in all lists, for a total 160 items in each list.

109 participants completed the study, of which 66 met the eligibility criteria described above (as outlined in Section "Participants").[10] Twenty-five participants were excluded for having a mean plausibility rating lower than seven on filler items (all of which were perfectly plausible). Data from the remaining 41 participants (36 females; mean age = 18.73; ranging from 18 to 22) were used to calculate plausibility ratings.

The mean rating for sentences with plausible objects and definite determiners (… *the wine*) was 8.06 ($SE = 0.16$), whereas mean ratings for plausible objects with demonstrative determiners (… *that wine*) was 7.80 ($SE = 0.16$). Mean ratings for implausible objects with definite vs. demonstrative determiners (…*the roof* vs. …**that roof*) were 1.88 ($SE = 0.23$) and 1.79 ($SE = 0.23$), respectively.

An ANOVA conducted on the mean plausibility ratings with the independent variables of object type (congruent vs. incongruent), determiner type (definite vs. demonstrative) was conducted. Significant main effects of object type, $F(1, 40) = 453.9$, $MSE = 3.36$, $p < 0.001$, $\eta_p^2 = 0.919$, and determiner type, $F(1, 40) = 5.48$, $MSE = 0.23$, $p = 0.024$, $\eta_p^2 = 0.121$ were observed. No significant interaction of object and determiner type was observed, $F(1, 40) = 0.88$, $MSE = 0.37$, $p = 0.355$, $\eta_p^2 = 0.021$. Overall, these results confirm the intended contrasts.

2.3. Electrophysiological measures

Electroencephalographic recordings were made using a 64-channel Active Two BioSemi system (BioSemi, Amsterdam). Data were sampled at a rate of 512 Hz and digitized with a 24-bit analog-to-digital converter. Two infinite impulse response filters were applied at 12 db/octave: a bandpass filter from 0.1 to 100 Hz used to remove high and low frequency noise and a bandstop filter from 59 to 61 Hz used to remove 60 Hz noise. All electrodes were re-referenced to the averaged mastoids for analysis. Prior to segmentation, eye movements artifacts and blinks were filtered from the data using a spatial ocular artifact correction algorithm (Pflieger, 2001) available in the EMSE v5.5.1 software (Cortech Solutions, 2013). Due to equipment malfunction, data from electrode Fp1 was lost in some participants. A spatial interpolation filter (Cortech Solutions, 2013) was applied for this electrode, for all participants.

[9] We note that although *"that"* can also occur as a complementizer, as in "*Anita thinks that Mary arrived.*", the critical stimuli used verbs that select for direct object nouns, not sentential complements—thereby avoiding any potential lexical ambiguity. A full list of stimuli may be found at link: https://gitlab.com/dwivedilab/erp_reference.

[10] It is a practice in our lab to not exclude students enrolled in courses requiring experimental participation for course credit. Data collected from ineligible students are eliminated later for data analysis (presently, 38 bilingual/ multilingual students, 4 students diagnosed with neurological disorders, and 1 student under 18 years of age were ineligible. Thus, a total of 43 students were excluded; as such, their data were not included for analyses).

Epochs were created from an interval 200 ms prior to stimulus onset to 1200 ms after stimulus onset.

2.4. Procedure

Participants were tested individually in one session of approximately three hours. In each session, participants completed a short questionnaire regarding reading habits, a handedness inventory (Briggs & Nebes, 1975) and the Positive and Negative Affect Schedule (PANAS, Watson et al., 1988) before the application of the electrodes.[11] Following a practice session of eight trials, each participant completed the experimental trials in six blocks of 55 trials, with rest periods between each block. Each participant saw one of four pseudorandomized, counterbalanced lists consisting of 330 items. The pseudorandomized lists were created using the Mix utility (van Casteren & Davis, 2006) such that the first three items and last two items of each block were always filler sentences; no more than two critical items were presented sequentially and items from the same condition were never presented sequentially. Sentences were presented in the centre of the computer monitor in light grey, 18-point Courier New font on a black background. See Fig. 1 for a sample trial procedure.

Each trial sentence began with the participant being prompted to press a button on the response pad, then the word "Blink" was presented for 1000 ms, followed by a fixation cross (+) for 500 ms. After a variable inter-trial interval lasting between 200–400 ms, the sentence was presented one word at a time with a stimulus onset asynchrony (SOA) of 600 ms and an inter-stimulus interval (ISI) of 200 ms. 125 filler items were followed by comprehension questions after the last word of the sentence, to which participants were asked to press a "1" or "2" key corresponding to answers on the screen using the response pad. Response time and accuracy was recorded for each response. The next trial began following another inter-trial interval lasting between 500–1000 ms.

3. Results

3.1. Behavioural analyses

Comprehension rates for questions at filler conditions were 95.2 % ($SD = 2.75$ %), indicating that participants were indeed paying attention to sentence materials.

3.2. Electrophysiological analyses

The grand average ERPs, time-locked at the position of the critical word *#roof* vs. *wine* are shown for all four conditions, i.e., *the/*that wine* and *the/*that #roof*. Fig. 2A shows a clear N400 effect for incongruent "*roof*" vs. "*wine*", occurring with typical distribution (see Fig. 2C): maximal over centroparietal sites with a slight right lateralization (Kutas & Federmeier, 2011). Contrary to our a priori hypothesis, no P600 effect emerged at the dual violation condition **that roof*.

That is, a positive deflection did not follow the N400 effect but instead preceded it (see Fig. 2B). A P200 effect was elicited before the N400 component for the double violation condition **that #roof* but not *the #roof* or **that wine*. Moreover, the N400 effect at the double violation condition **that #roof* was attenuated (see Fig. 2A and 2C). In order to ensure that N400 amplitudes were not influenced by the immediately preceding positive-going P200 deflection, we renormalized N400 amplitudes relative to the post-stimulus interval of 100–300 ms (the P200 latency window) after onset of *roof/wine*[12] (see also Hagoort, 2003, and Carreiras, Vergara & Barber, 2005 for similar analyses). Results below are reported using the renormalized N400 amplitude (see Fig. 3).[13]

Next, we conducted single trial, linear mixed effect regression analyses using the R statistical programming language (v4.2.2) with packages lme4 (v1.1.34, for linear mixed effects regression model fitting) and EM means (v1.8.8, for Bonferroni corrected pairwise contrasts). Statistical analyses reported below were completed using custom R code, and figures were generated using custom Python code. All materials (stimuli, data, and scripts) associated with this experiment are available at https://gitlab.com/dwivedilab/erp_reference.

Given the clear evidence of P200 and N400 effects, ERP analyses were conducted in standard time windows generally associated with these components, namely at 100–300 ms and 300–500 ms, respectively (Kiel et al., 2014). Analyses were conducted using electrode regions of interest (ROIs) as described in Dwivedi & Gibson (2017) (see also Selvanayagam et al., 2019; Gisladottir et al., 2015). Medial regions of interest consisted of: Anterior Medial (Fz, F3, F4, FC1, FC2), Central (Cz, C1, C2, C3, C4) and Posterior Medial (Pz, P3, P4, PO3, PO4); and lateral regions of interest consisted of Left Anterior (FT7, F5, F7, T7), Right Anterior (FT8, F6, F8, T8), Left Posterior (CP5, P7, PO7, O1) and Right Posterior (CP6, P8, PO8, O2). Separate linear mixed effects regressions were conducted for medial and lateral ROIs separately for each time window (P200 effect: 100–300 ms, N400 effect: 300–500 ms). We evaluated the effects of the linguistic factors: Object Type (2 levels: plausible, implausible), Determiner Type (2 levels: definite, demonstrative) and any interactions with non-linguistic effects of Anteriority (medial, 3 levels: Anterior, Central, Posterior; lateral, 2 levels: Anterior, Posterior) and for lateral models, Hemisphere (2 levels: Left, Right). Additionally, we evaluated the effects of the random slope and intercept terms for participant, electrode, and item. Thus, the medial model was fit with the formula: mean voltage ~ object * determiner * anteriority + (1 + object + determiner | electrode) + (1 + object + determiner | pptid) + (1 + object + determiner | itemid) and the lateral model with: mean voltage ~ object * determiner * anteriority * hemisphere + (1 + object + determiner | electrode) + (1 + object + determiner | pptid) + (1 + object + determiner | itemid). Our linear mixed modelling approach here follows the traditional approach of stepwise regression wherein predictors are iteratively removed and compared to evaluate significance (see for example Bornkessel-Schlesewsky et al., 2022). We evaluated the significance of these models using a chi-squared likelihood ratio test by contrasting with reduced models which excluded interactions and main effects involving fixed terms, as well as random slope and intercept terms, in a step-wise manner.

3.3. Critical word position, wine vs #roof

3.3.1. 100–300 ms

A significant P200 effect was revealed for the double violation condition **that #roof* vs. **that wine* (and **the #roof*). The highest order interaction, of Object, Determiner, and Anteriority (and Hemisphere for lateral) models significantly improved fit as compared to lower order models for medial, $\chi(8) = 17.87$, $p = 0.022$, and lateral ROIs, $\chi(8) = 22.89$, $p = 0.029$. Pairwise contrasts revealed a significant increase in mean voltage in the dual violation condition (**that #roof* vs **that wine*, $\Delta = 0.70–1.03\ \mu V$) 100–300 ms following critical word onset. This effect was largely observed at medial posterior sites (p's < 0.05) and was marginal ($.045 < p < .058$) at lateral posterior sites as well as at medial anterior and central sites but not at lateral anterior sites (see Fig. 2B). No such effects were observed in the definite conditions (*the #roof* vs. *the*

[11] This questionnaire was employed to ask questions orthogonal to the current paper and is not discussed further. For a thorough account of that question and results, see Dwivedi & Selvanayagam (2021).

[12] That is, we subtracted the mean voltage in the 100-300ms time window from the 300-500ms time-window to compute a re-normalized N400 amplitude.

[13] For a discussion of statistical analyses for the N400 effect observed in Fig. 2C, see Supplementary Material S-1.

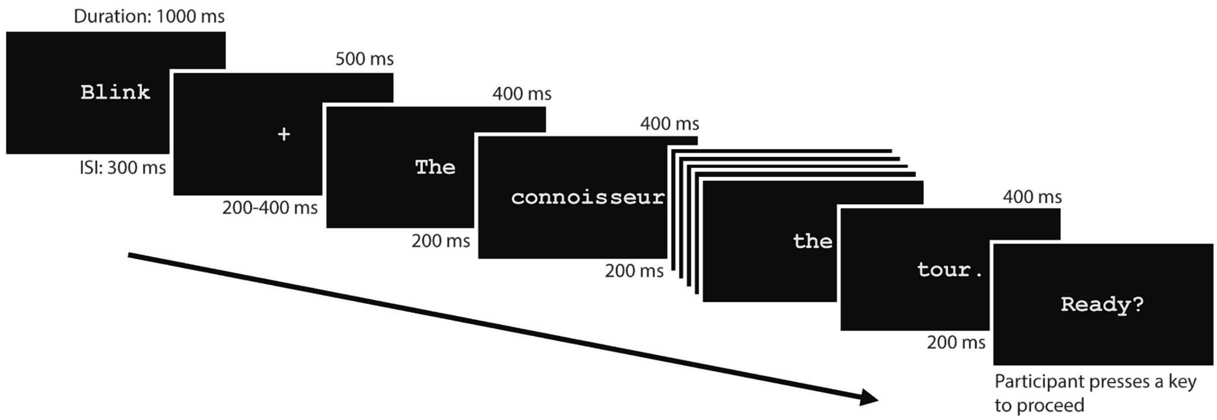

Fig. 1. Condensed sample trial for the current paradigm. Time values above the screen represent the duration of stimulus presentation, and time values below represent inter-stimulus intervals. The “Ready?” slide requires user input to proceed and was occasionally preceded by a comprehension question.

wine, p’s > 0.05; see Fig. 2B). In sum, we observed a significant medial posterior P200 effect for the implausible critical word following the demonstrative but not the definite determiner. This effect appears to index the double violation, and critically, this difference is observed in the contrast (**that #roof* vs **that wine*) where the baseline is held constant.

3.3.2. N400 effect: 300–500 ms

Fig. 3A shows the grand average ERPs, time-locked to the onset of the critical word (*#roof* vs. *wine*) at medial and lateral electrode sites, with a post-stimulus baseline of 100–300 ms as opposed to the pre-stimulus baseline of 0–200 ms used in Fig. 2A. Visually, it is evident that the N400 effect is typical with respect to both latency and topography (see Fig. 3B). The highest order interaction, of Object, Determiner, and Anteriority (and Hemisphere for lateral) models significantly improved fit beyond all other reduced models, medial: χ (8) $= 17.862$, $p = 0.022$, lateral: $\chi(12) = 28.248$, $p = 0.005$. Pairwise contrasts here confirmed an N400 effect robustly across both Determiner types, with a Central-Posterior distribution with slight right lateralization. N400 amplitudes were slightly attenuated for the demonstrative condition (medial: $\Delta =$ 1.08–1.80 μV, lateral: $\Delta =$ 0.175–0.975 μV) and spatially restricted (not significant in the left anterior ROI) as compared to the definite condition (medial: $\Delta =$ 1.67–1.97 μV, lateral: $\Delta =$ 0.837–1.356 μV, significant in all ROIs).

3.4. Critical word minus one position, determiner, the vs *that

Next, grand average ERPs, time-locked at the position of the Determiner, or critical word minus one position (*the* vs. *that*), are shown at medial sites in Fig. 4. Visual inspection of the waveform reveals a difference in voltage starting at 300 ms which persists until 500 ms. Although it is maximal at right, centroparietal sites, there is no peak as characteristic of a typical N400 component. To investigate these differences, as above, we evaluated linear mixed effect regression models separately for medial and lateral ROIs, omitting the factor of Object type and all associated interactions: medial = mean voltage ~ determiner * anteriority + (1 + determiner | electrode) + (1 + determiner | pptid) + (1 + determiner | itemid); lateral = mean voltage ~ determiner * anteriority * hemisphere + (1 + determiner | electrode) + (1 + determiner | pptid) + (1 + determiner | itemid). The highest order interaction, of Determiner and Anteriority (and Hemisphere for lateral) models significantly improved significantly observed fit for medial: $\chi(4) =$ 10.009, $p = 0.040$, and lateral: $\chi(6) = 18.361$, $p = 0.005$ ROIs. Pairwise comparisons revealed significantly more negative mean voltages for **that* as compared to “*the*” for central and posterior medial ROIs and the right posterior ROI (p’s < 0.05). These results confirm a right lateralized centroparietal negativity for the demonstrative determiner as compared to the definite determiner 300–500 ms following stimulus onset.

While the effect observed here resembles the N400 in timing and topography, the shape of the waveform does not correspond to this component. This negativity likely indexes the differences in word-frequency differences between “*the*” and “*that*” (van Petten & Kutas, 1990; van Petten, 1995).

4. Discussion

In the present study, we were interested in neural responses to words that exhibited dual meaning violations: first, in terms of real-world plausibility, and second, in terms of referential meaning. That is, “*The connoisseur tasted *that #roof…*” (vs. “*… *that wine*”; also vs. control condition “*… the wine*”) is incongruent both in terms of contextual plausibility and in terms rules of reference. Given our previous findings regarding semantic anomaly associated with reference, we predicted an N400-P600 complex. We predicted “*roof*”, an implausible direct object in its immediate sentence context, would elicit an N400 component, and the use of the demonstrative determiner “*that*”, would result in an independent (semantic) P600 effect, since “*that*” violated discourse structure algorithms regarding semantic reference, (as in Dwivedi et al., 2006, 2010).

Our predictions were partially borne out. We did, in fact, see independent neural responses to the combined violation condition, **that #roof*. However, instead of an N400 followed by a P600 component at the critical word “*roof*”, we observed a P200-N400 complex. Given the clear implausibility of tasting a roof vs. wine, the N400 was an expected neural response at this condition. However, the P200 was not. Below, we discuss the cognitive significance of the P200-N400 complex and then conclude with why the P600 was not observed.[14]

4.1. P200, attention and the algorithm of meaning

The P200 component has been associated with allocation of attention, where stimuli that are attended to yield larger P200 components vs. unattended stimuli (Hillyard & Münte, 1984; Luck & Hillyard, 1994). With respect to language, studies with P200 effects often discuss this ERP component in terms of attention and salience of the relevant linguistic cue.

In a recent ERP language experiment, Vergis et al. (2020) showed that when participants listened to sentences that were spoken with either rude or polite voices, P200 effects were found at sentence-final words in the rude prosody conditions. The researchers hypothesized

[14] This section was revised thanks in part to reviewers’ suggestion to better explicate the concept of presupposition at the critical words.

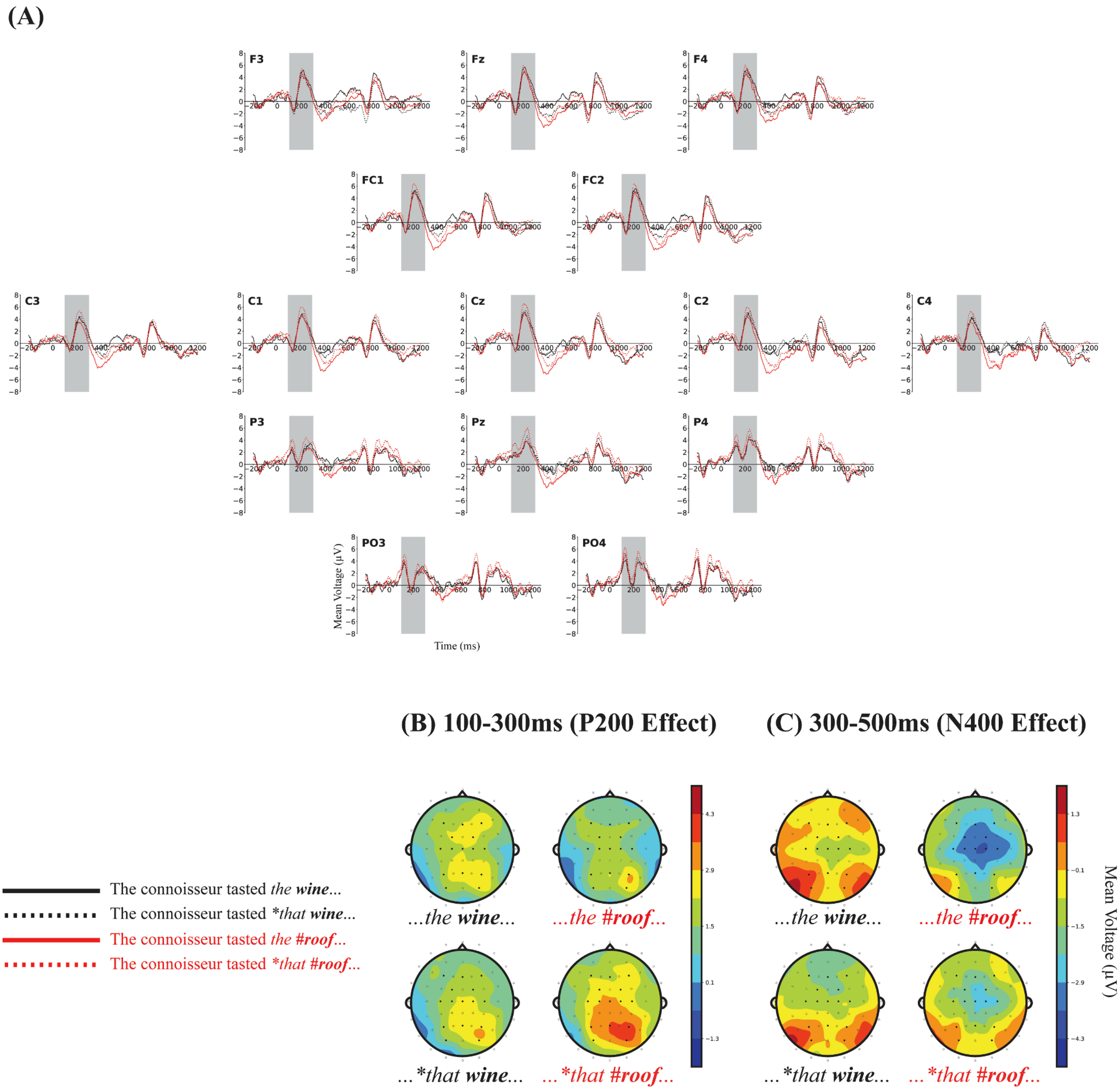

Fig. 2. Grand average ERPs at the critical word (*wine/#roof*) with a 200 ms pre-stimulus baseline. **(A)** ERP waveforms at medial electrode sites for all four conditions *the wine* (solid black), **that wine* (dotted black), *the #roof* (solid red), **that #roof* (dotted red). Topographic plots of mean amplitudes (μV) during time windows for P200 (100–300 ms) in **(B)** and N400 (300–500 ms) in **(C)** after stimulus onset at the critical word. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

that the P200 effect found for rude-sounding intonation reflected greater attention by listeners since that cue was salient to the task at hand; during the experiment, participants had to rate how likely it was that someone might comply with rude vs. polite requests by the speaker. Thus, the rude vocal cue was more noticeable since it was germane to the task of deciding whether someone might comply or co-operate with the speaker. In another experiment, Zhao et al. (2021) examined scalar implicature sentences in Mandarin and showed that focus conditions elicited larger P200 components. They reasoned that focus conditions would require more attentional resources than non-focused conditions for interpretation.

On a view where the P200 indexes attention, the results for the current experiment become straightforward. As mentioned previously, meaning is compositional such that the interpretation of a sentence varies when the noun is preceded by a different determiner (e.g., "*the apples…*" vs. "*all apples…*"). When "*roof*" was perceived in the sentence containing "*connoisseur… taste…*", it was clearly not expected or associated with the local sentence context (Dwivedi, Goertz, Selvanayagam, 2018), in contrast to "*wine*", which fit perfectly. The extra effort required in retrieving the meaning of "*roof*" vs. "*wine* (Aurnhammer et al., 2021; Federmeier & Kutas, 1999) would necessarily result in extra attentional resources for interpreting "*that roof*", indexed here by a P200-N400 complex. No evidence of extra attention or salience is there for "*wine*" due to its 'good enough' fit with the local context (see more

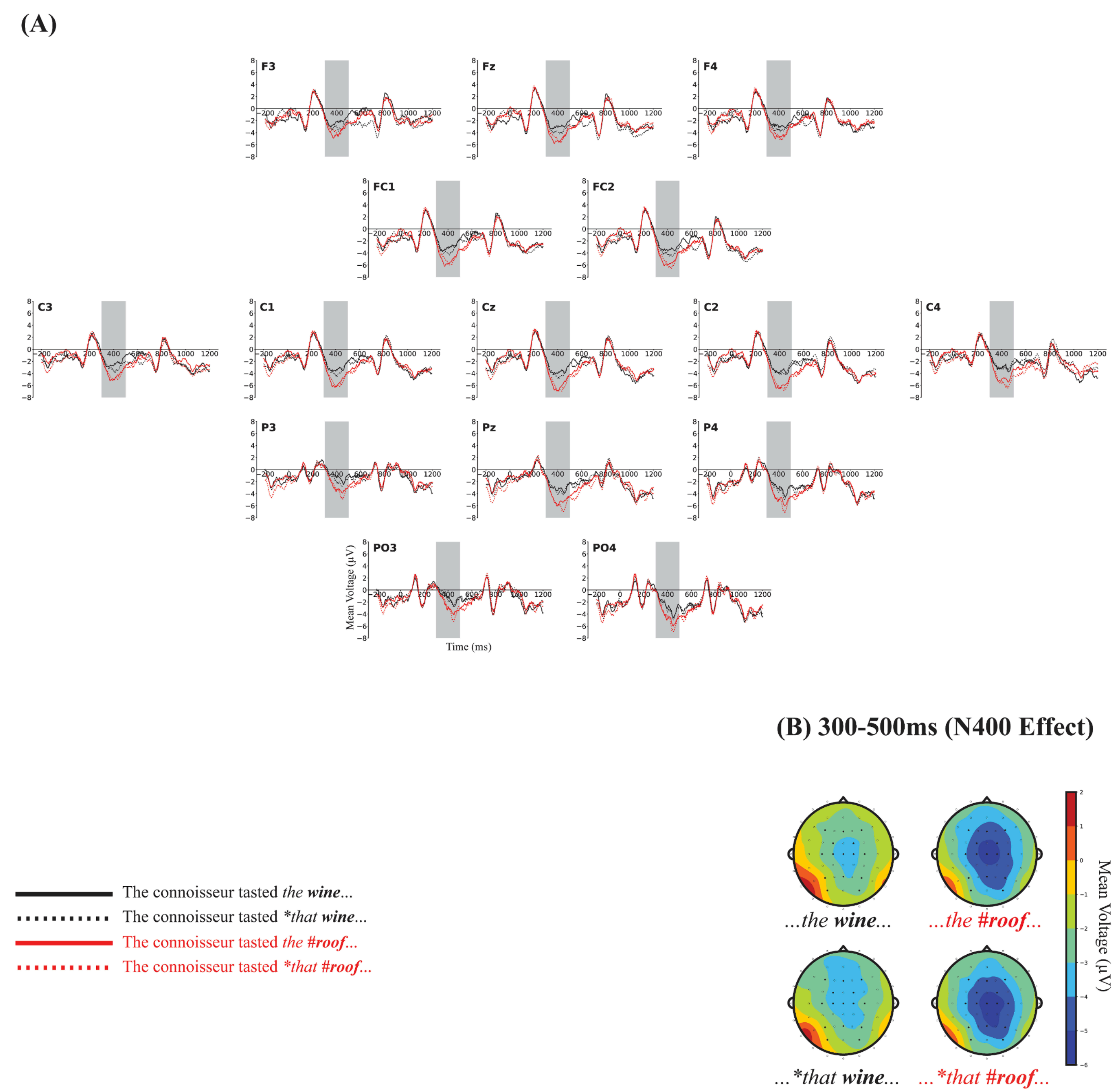

Fig. 3. Grand average ERPs at the critical word (*wine/#roof*) with a 100–300 ms post-stimulus baseline to compensate for differences due to the preceding P200 effect. **(A)** ERP waveforms at medial electrode sites for all four conditions *the wine* (solid black), **that wine* (dotted black), *the #roof* (solid red), **that #roof* (dotted red). **(B)** Topographic plot of mean amplitude (μV) during 300–500 ms time window for renormalized N400. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

discussion below).

We note that a P200-N400 was also observed in another ERP language experiment (Carreiras et al., 2005) and that a clear (though speculative) link can be made between those previous findings and ours; especially on account that takes difficulty/ease of lexical retrieval into account.

Carreiras et al. (2005) were interested in whether and how sublexical rules of syllabification applied to single words during reading. In two experiments, syllable boundaries were marked by colour boundaries for (both high and low frequency) words and pseudo-words in a lexical decision task. These colour boundaries either matched (e.g., “***casa***”), or did not match (e.g., “***casa***”) syllable boundaries. When a mismatch between syllable and colour boundaries occurred, a P200-N400 complex emerged for low frequency and pseudo-words— but not high frequency words. Presumably, interpreting low frequency and pseudo-words required more cognitive effort resulting in more attention—the same idea as proposed above. This increase in salience would have elicited the P200 effect for low-frequency words, vs. high frequency words. Regarding the lack of a P200 effect for incongruently marked high frequency words, they indicated that “[s]yllabic parsing may routinely occur for high-frequency words but may be quickly overshadowed by the fast lexical access to the word itself,” (Ibid., p. 1811). Without extra attentional resources, the effort required for syllabic parsing would not occur. Similarly, in the present experiment,

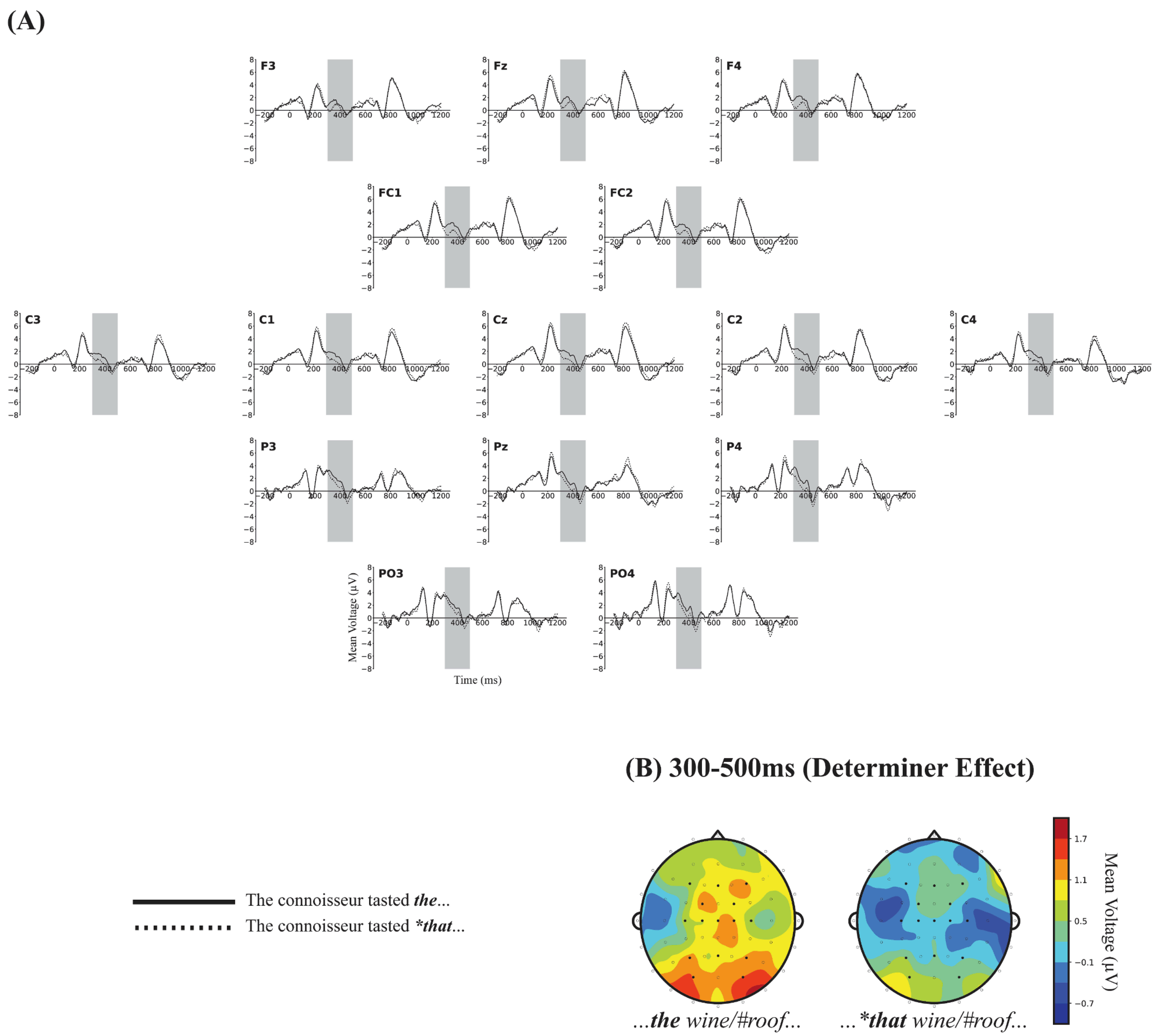

Fig. 4. Grand average ERPs at the determiner (critical word minus one; *the/*that*) with a 200 ms pre-stimulus baseline. **(A)** ERP waveforms at medial electrode sites for both conditions *the* (solid black) and **that* (dotted black). **(B)** Topographic plots of mean amplitude (μV) 300–500 ms after stimulus onset at the determiner to investigate the distribution of the observed effect.

interpreting *"wine"* in the sentence did not require extra attentional resources due to its 'good enough' (Townsend & Bever, 2001; Ferreira, 2003; Dwivedi, 2013) fit with the immediate context—resulting in no P200 effect.

4.2. Addendum: P200 and presupposition

We note here that P200 effects were also found in a series of ERP language studies by Regel and colleagues (2010, 2011, 2014) that examined comprehension of ironic vs. literal sentences. That is, sentences, when interpreted on their ironic vs literal interpretation, elicited P200 effects (exhibiting similar parietal topography as reported herein) at sentence-final words when preceded by appropriate context. For example, P200 effects were shown in sentences such as "*You should take a break*" only on the ironic interpretation (where the context sets up the addressee as someone who has barely worked at all) vs. the literal interpretation (the context is about an addressee who has worked for several hours). Ironic sentences have a presuppositional meaning, in that they require context for interpretation, or (i.e., the sentence expresses the opposite of its literal meaning, which can only be derived by context, see Bollobás, 1981; Schlöder, 2017). Thus, another related way of interpreting the P200 effect found in the current experiment would be that, provided attentional resources are available, the P200 marks presupposition. If so, then the finding of this early ERP response emerging before the N400 would suggest that the interpretation of presupposition occurs at the earliest stages of nominal processing. In fact, one could further speculate that the P200 component is a neural signature of discourse linking (Pesetsky, 1987), and consider that recent P300 findings by Jouravlev et al. (2016) examining presuppositional failure consist of the same component, or family of components (relatedly, see also Leckey and Federmeier, 2020). We leave these questions for further research.

Furthermore, we note that P200 effects found for sentences in visual field studies conducted by Federmeier and colleagues (2002, 2005) are likely not related to the current findings. First, larger P200 effects were found for expected vs. unexpected words, which is the opposite of our findings. Moreover, the aforementioned studies manipulated highly constraining vs. weakly constraining sentences, which was not an aspect of the present design.

4.3. The lack of a P600

We did not observe the P600 effect at **that roof*, as predicted, given our earlier studies (Dwivedi et al., 2006; 2010) examining referential anomaly. Perhaps differences between our preceding work and the present experiment could explain why. First, our previous work examined anaphoric dependencies between two sentences vs. the current single sentence study. Second, it was suggested in the previous studies that the observed semantic P600 might reflect the cost of cognitive procedure of revision. That is, for co-reference to occur between the pronoun *"it"* and the (hypothetical) antecedent "a *novel*", interpretation of the antecedent in the first sentence would need to be adjusted. In the current experiment, the only context available is the single sentence, and it cannot be revised in any way to help with interpretation. That is, when "*roof*" is perceived after **that,* there is no probable or possible adjustment to be made. This could explain the difference in the ERP components—different cognitive procedures are at play.

A difference in cognitive procedures would also explain the lack of an empirical effect at **that wine*… That is, there is no previous context to adjust to accommodate the presuppositional meaning associated with *that wine*; and/or if any adjustments are made, the cost of updating the common ground is minimal (see von Fintel, 2008). This would be because "*wine*" is a 'good enough' fit with the internal sentence context and few resources (if at all) would be required for accommodation (Ferreira, 2003; Chwilla & Kolk, 2005; Dwivedi, 2013).

Interestingly, our off-line findings did show empirical differences between *the/that wine* sentences. This is likely due to the differences in methodology of the norming study vs. ERP methods. That is, the off-line norming study displayed the entire sentence all at once, and participants were tasked with rating sentences for naturalness, under zero-time pressure. This contrasts with the ERP experiment, where no task was associated with critical sentences (Kaan & Swaab, 2003a; Kolk et al., 2003; Schacht et al., 2014). In addition, the presentation was timed, using standard RSVP methods. As a result, participants would not have the opportunity to look back and review the sentence for interpretation, resulting in potentially different interpretive processes.

4.4. Conclusion

In sum, our findings support the notion that meaning can be derived from separate sources of information; both contextual heuristics and grammar, as indexed by independent ERP components. We note that although we did not find a semantic P600 effect, we did find a P200-N400 complex for the combined violation condition, supporting the independence of meaning derived by context vs. grammar (in contrast to Hagoort et al., 2004). We construed the P200 effect as a marker of increased attention, perhaps due to the increased effort associated with interpreting an implausible noun. We further speculated that the P200 effect could be a marker of presupposition, as argued for the P300 component found in Jouravlev et al. (2016). Finally, we note that our findings are consistent with our sentence processing model of Heuristic first, algorithmic second (Dwivedi, 2013), as well as the Retrieval-Integration account of language processing (Aurnhammer et al., 2021).

CRediT authorship contribution statement

Veena D. Dwivedi: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Janahan Selvanayagam:** Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All materials (stimuli, data, and scripts) associated with this experiment are available at https://gitlab.com/dwivedilab/erp_reference.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bandl.2024.105438.

References

Anderson, J. E., & Holcomb, P. J. (2005). An electrophysiological investigation of the effects of coreference on word repetition and synonymy. *Brain and Language, 94*(2), 200–216. https://doi.org/10.1016/j.bandl.2005.01.001

Aurnhammer, C., Delogu, F., Schulz, M., Brouwer, H., & Crocker, M. W. (2021). Retrieval (N400) and integration (P600) in expectation-based comprehension. *PLOS ONE, 16* (9). https://doi.org/10.1371/journal.pone.0257430

Bollobás, E. (1981). Who's afraid of Irony? An analysis of uncooperative behavior in Edward Albee's 'Who's afraid of Virginia Woolf'. *Journal of Pragmatics, 5*(4), 323–334. https://doi.org/10.1016/0378-2166(81)90035-7

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "Semantic P600" effects in language comprehension. *Brain Research Reviews, 59*(1), 55–73. https://doi.org/10.1016/j.brainresrev.2008.05.003

Bornkessel-Schlesewsky, I., Sharrad, I., Howlett, C. A., Alday, P. M., Corcoran, A. W., Bellan, V., Wilkinson, E., Kliegl, R., Lewis, R. L., Small, S. L., & Schlesewsky, M. (2022). Rapid adaptation of predictive models during language comprehension: Aperiodic EEG slope, individual alpha frequency and idea density modulate individual differences in real-time model updating. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.817516

Briggs, G. G., & Nebes, R. D. (1975). Patterns of hand preference in a student population. *Cortex, 11*(3), 230–238. https://doi.org/10.1016/S0010-9452(75)80005-0

Carreiras, M., Vergara, M., & Barber, H. (2005). Early event-related potential effects of syllabic processing during visual word recognition. *Journal of Cognitive Neuroscience, 17*(11), 1803–1817. https://doi.org/10.1162/089892905774589217

Chierchia, G., & McConnell-Ginet, S. (1990). Meaning and grammar: An introduction to semantics.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

Chwilla, D. J., & Kolk, H. H. J. (2005). Accessing world knowledge: Evidence from N400 and reaction time priming. *Cognitive Brain Research, 25*(3), 589–606. https://doi.org/10.1016/j.cogbrainres.2005.08.011

Cortech Solutions. (2013). *EMSE Suite* (Version 5.5.1) [Computer software]. Cortech Solutions. https://cortechsolutions.com/.

Diessel, H., Coventry, K., Gudde, H., Capirci, O., eds. (2021). Demonstratives, deictic pointing and the conceptualization of space. *Lausanne: Frontiers Media SA*. https://doi.org/10.3389/978-2-88966-782-6.

Dwivedi, V. D. (1996). Modality and Discourse Processing. *McGill Working Papers in Linguistics, 12*, 17–52.

Dwivedi, V. D. (2013). Interpreting quantifier scope ambiguity: Evidence of heuristic first, algorithmic second processing. *PLOS ONE, 8*(11), 1–20. https://doi.org/10.1371/journal.pone.0081461

Dwivedi, V. D., Drury, J. E., Molnar, M., Phillips, N. A., Baum, S., & Steinhauer, K. (2010). ERPs reveal sensitivity to hypothetical contexts in spoken discourse. *NeuroReport, 21*(11), 791–795. https://doi.org/10.1097/wnr.0b013e32833cae0d

Dwivedi, V. D., & Gibson, R. M. (2017). An ERP investigation of quantifier scope ambiguous sentences: Evidence for number in events. *Journal of Neurolinguistics, 42*, 63–82. https://doi.org/10.1016/J.JNEUROLING.2016.11.006

Dwivedi, V. D., Goertz, K. E., & Selvanayagam, J. (2018). Heuristics in language comprehension. *Journal of Behavioural and Brain Sciences, 08*(07), 430–446. https://doi.org/10.4236/jbbs.2018.87027

Dwivedi, V. D., Phillips, N. A., Laguë-Beauvais, M., & Baum, S. R. (2006). An electrophysiological study of mood, modal context, and anaphora. *Brain Research, 1117*(1), 135–153. https://doi.org/10.1016/j.brainres.2006.07.048

Dwivedi, V.D, Selvanayagam, J. (2021). Effects of dispositional affect on the N400: language processing and socially situated context. *Frontiers in Psychology: Language Sciences*. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.566894/full.

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language, 41*(4), 469–495. https://doi.org/10.1006/jmla.1999.2660

Federmeier, K. D., & Kutas, M. (2002). Picture the difference: Electrophysiological investigations of picture processing in the two cerebral hemispheres. *Neuropsychologia, 40*(7), 730–747. https://doi.org/10.1016/s0028-3932(01)00193-2

Federmeier, K. D., Mai, H., & Kutas, M. (2005). Both sides get the point: Hemispheric sensitivities to sentential constraint. *Memory & Cognition, 33*(5), 871–886. https://doi.org/10.3758/BF03193082

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*(2), 164–203. https://doi.org/10.1016/S0010-0285(03)00005-7

Filik, R., Sanford, A. J., & Leuthold, H. (2008). Processing pronouns without antecedents: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience, 20* (7), 1315–1326. https://doi.org/10.1162/jocn.2008.20090

Gisladottir, R. S., Chwilla, D. J., & Levinson, S. C. (2015). Conversation electrified: ERP correlates of speech act recognition in underspecified utterances. *PLOS ONE, 10*(3). https://doi.org/10.1371/journal.pone.0120068

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience, 15*(6), 883–899. https://doi.org/10.1162/089892903322370807

Hagoort, P., Brown, C. M., & Groothusen, J. (1993). The syntactic positive shift as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8*(4), 439–483. https://doi.org/10.1080/01690969308407585

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. https://doi.org/10.1126/science.1095455

Hammer, A., Jansma, B. M., Lamers, M., & Münte, T. F. (2005). Pronominal reference in sentences about persons or things: An electrophysiological approach. *Journal of Cognitive Neuroscience, 17*(2), 227–239. https://doi.org/10.1162/0898929053124947

Heim, I. (1982). *The semantics of definite and indefinite noun phrases*. [Doctoral dissertation, University of Massachusetts-Amherst]. ScholarWorks@UMass. https://scholarworks.umass.edu/dissertations/AAI8229562/.

Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Oxford: Blackwell Publishers.

Hillyard, S. A., & Münte, T. F. (1984). Selective attention to color and location: An analysis with event-related brain potentials. *Perception and Psychophysics, 36*(2), 185–198. https://doi.org/10.3758/BF03202679

Jouravlev, O., Stearns, L., Bergen, L., Eddy, M., Gibson, E., & Fedorenko, E. (2016). Processing temporal presuppositions: An event-related potential study. *Language, Cognition and Neuroscience, 31*(10), 1245–1256. https://doi.org/10.1080/23273798.2016.1209531

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes, 15*(2), 159–201. https://doi.org/10.1080/016909600386084

Kaan, E., & Swaab, T. Y. (2003a). Electrophysiological evidence for serial sentence processing: A comparison between non-preferred and ungrammatical continuations. *Cognitive Brain Research, 17*(3), 621–635. https://doi.org/10.1016/S0926-6410(03)00175-7

Kaan, E., & Swaab, T. Y. (2003b). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience, 15*(1), 98–110. https://doi.org/10.1162/089892903321107855

Kaplan, D. (1978). Dthat. In P. Cole (Ed.), *Syntax and Semantics:* Vol. 9. (pp. 221–243). New York, NY: Academic Press.

Karttunen, L. (1976). Discourse Referents. In J. D. McCawley (Ed.), *Syntax and Semantics* (Vol. 7, pp. 363–386). Academic Press. https://philpapers.org/rec/KARDR-2.

Katz, J. J., & Fodor, J. A. (1963). The Structure of a Semantic Theory. *Language, 39*(2), 210. https://doi.org/10.2307/411200

Kiel, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E. S., Luck, S. J., Luu, P., Miller, G. A., & Yee, C. M. (2014). Committee report: Publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology, 51*(1), 1–21. https://doi.org/10.1111/psyp.12147

King, J. C. (2006). Singular terms, reference and methodology in semantics. *Philosophical Issues, 16*(1), 141–161. https://doi.org/10.1111/j.1533-6077.2006.00107.x

Kolk, H. H., Chwilla, D. J., Van Herten, M., & Oor, P. J. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and language, 85*(1), 1–36. https://doi.org/10.1016/s0093-934x(02)00548-5

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, 1146*, 23–49. https://doi.org/10.1016/j.brainres.2006.12.063

Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2020). A tale of two positivities and the N400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience, 32*(1), 12–35. https://doi.org/10.1162/jocn_a_01465

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*(1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203–205. https://doi.org/10.1126/science.7350657

Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition, 11*(5), 539–550. https://doi.org/10.3758/BF03196991

Leckey, M., & Federmeier, K. D. (2020). The P3b and P600(s): Positive contributions to language comprehension. *Psychophysiology., 57*, e13351. https://doi.org/10.1111/psyp.13351

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology, 31*(3), 291–308. https://doi.org/10.1111/j.1469- 8986.1994.tb02218.x

Magidor, O. (2019). Category Mistakes. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Fall, 2019)*. Stanford University. https://plato.stanford.edu/archives/fall2019/entries/category-mistakes/.

Murphy, G. L. (1984). Establishing and accessing referents in discourse. *Memory & Cognition, 12*(5), 489–497. https://doi.org/10.3758/BF03198311

Nieuwland, M. S., & Van Berkum, J. J. A. (2008). The neurocognition of referential ambiguity in language comprehension. *Language and Linguistics Compass, 2*(4), 603–630. https://doi.org/10.1111/j.1749-818x.2008.00070.x

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*(6), 785–806. https://doi.org/10.1016/0749-596X(92)90039-Z

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 34*(6), 739–773. https://doi.org/10.1006/JMLA.1995.1033

Pesetsky, D. (1987). Wh-in-situ: Movement and unselective binding. In E. Reuland & A. ter Meulen (Eds.), *The Representation of (in) Definiteness*, (pp. 98–129). Cambridge, Mass.: MIT Press.

Pflieger, M. E. (2001). Theory of a spatial filter for removing ocular artifacts with preservation of EEG. EMSE Workshop, September 2001, 7–8. https://doi.org/10.13140/RG.2.1.4810.0086.

Qualtrics XM- Experience Management Software. Qualtrics (2020). Qualtrics. https://www.qualtrics.com.

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. https://doi.org/10.1016/j.brainres.2009.10.077

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. https://doi.org/10.1162/jocn.2010.21411

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). https://doi.org/10.1371/journal.pone.0096840

Roberts, C. (1989). Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy, 12*(6), 683–721. https://doi.org/10.1007/bf00632602

Roberts, C. (1996). Anaphora in intensional contexts. In S. Lappin (Ed.), *The Handbook of Contempory Semantic Theory.* (pp. 215–246). Hoboken: Blackwell Publishers.

Roberts, C. (2002). Demonstratives as definites. In Stevenson, R., Deemter, K. V., & Kibble, R. (Eds.) *Information sharing: Reference and presupposition in language generation and interpretation* (pp. 89-136). CSLI Publications.

Roberts, C. (2003). Uniqueness in definite noun phrases. *Linguistics and Philosophy, 26*(3), 287–350. https://doi.org/10.1023/A:1024157132393

Rullmann, H. (2003). Additive particles and polarity. *Journal of Semantics, 20*(4), 329–401. https://doi.org/10.1093/jos/20.4.329

Schacht, A., Sommer, W., Shmuilovich, O., Martíenz, P. C., & Martín-Loeches, M. (2014). Differential task effects on N400 and P600 elicited by semantic and syntactic violations. *PLOS ONE, 9*(3). https://doi.org/10.1371/journal.pone.0091226

Schwarz, F. (2009). *Two types of definites in natural language*. [Doctoral dissertation, University of Massachusetts Amherst]. ScholarWorks@UMass. https://scholarworks.umass.edu/open_access_dissertations/122.

Selvanayagam, J., Witte, V., Schmidt, L. A., & Dwivedi, V. D. (2019). A preliminary investigation of dispositional affect, the P300, and sentence processing. *Brain Research, 1721*, Article 146309. https://doi.org/10.1016/j.brainres.2019.146309

Stalnaker, R. (1978). Assertion. In P. Cole (Ed.), *Syntax and Semantics* (pp. 315–332). New York, NY: Academic Press.

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules*. MIT Press. https://doi.org/10.7551/mitpress/6184.001.0001

Van Berkum, J. J. (2009). The neuropragmatics of 'simple' utterance comprehension: An ERP review. In *Semantics and pragmatics: From experiment to theory* (pp. 276–316). Palgrave Macmillan.

van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods, 38*(4), 584–589. https://doi.org/10.3758/BF03193889

van Herten, M., Chwilla, D. J., & Kolk, H. H. J. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience, 18*(7), 1181–1197. https://doi.org/10.1162/jocn.2006.18.7.1181

van Petten, C. (1995). Words and sentences: Event-related potential measures. *Psychophysiology, 32*(6), 511–525. https://doi.org/10.1111/j.1469-8986.1995.tb01228.x

van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition, 18*(4), 380–393. https://doi.org/10.3758/bf03197127

Vergis, N., Jiang, X., & Pell, M. D. (2020). Neural responses to interpersonal requests: Effects of imposition and vocally expressed stance. *Brain Research, 1740*, 1–16. https://doi.org/10.1016/j.brainres.2020.146855 von Fintel, K. (2008). What is presupposition accommodation, again? *Philosophical Perspectives, 22*(1), 137–170. https://doi.org/10.1111/j.1520-8583.2008.00144.x

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology, 54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Zhao, M., Liu, X., Dai, X., Dong, S., & Han, Z. (2021). Scalar implicature is not a default process: An ERP study of the scalar implicature processing under the effect of focus factor. *Brain Research, 1765*, Article 147499. https://doi.org/10.1016/j.brainres.2021.147499

Further reading

IBM Corp. (2018). SPSS (Version 27.0) (No. 27). IBM Corp.