JOURNAL OF NEUROLINGUISTICS

An International Journal of Experimental, Clinical and Theoretical Research on Language and the Brain

---

Aims and Scope

The *Journal of Neurolinguistics* is an international forum for the integration of the neurosciences and language sciences. JNL provides for rapid publication of novel, peer-reviewed research into the interaction between language, communication and brain processes. The focus is on rigorous studies of an empirical or theoretical nature and which make an original contribution to our knowledge about the involvement of the nervous system in communication and its breakdowns. Contributions from neurology, communication disorders, linguistics, neuropsychology and cognitive science in general are welcome. Published articles will typically address issues relating some aspect of language or speech function to its neurological substrates with clear theoretical import. Interdisciplinary work on any aspect of the biological foundations of language and its disorders resulting from brain damage is encouraged. Studies of normal subjects, with clear reference to brain functions, are appropriate. Group-studies on well defined samples and case studies with well documented lesion or nervous system dysfunction are acceptable. The journal is open to empirical reports and review articles. Special issues on aspects of the relation between language and the structure and function of the nervous system are also welcome.

Author inquiries: You can track your submitted article at http://www.elsevier.com/track-submission. You can track your accepted article at http://www.elsevier.com/trackarticle. You are also welcome to contact Customer Support via http://support.elsevier.com.

Publication information: *Journal of Neurolinguistics* (ISSN 0911-6044). For 2017, volumes 42-45 are scheduled for publication. Subscription prices are available upon request from the Publisher or from the Elsevier Customer Service Department nearest you or from this journal's website (http://elsevier.com/locate/jneuroling). Further information is available on this journal and other Elsevier products through Elsevier's website: (http://www.elsevier.com). Subscriptions are accepted on a prepaid basis only and are entered on a calendar year basis. Issues are sent by standard mail (surface within Europe, air delivery outside Europe). Priority rates are available upon request. Claims for missing issues should be made within six months of the date of dispatch.

Journal of Neurolinguistics 45 (2018) 13–35

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

N400 and P600 modulation in presupposition accommodation: The effect of different trigger types

Filippo Domaneschi [a,*], Paolo Canal [b,1], Viviana Masia [c,1], Edoardo Lombardi Vallauri [c], Valentina Bambini [d]

[a] EXPRESS – Laboratory of Psychology of Language, Department of Educational Sciences, Psychology Unit, University of Genoa, Corso Andrea Podestà 2, 16128, Genoa, Italy
[b] Laboratorio di Linguistica "G. Nencioni", Scuola Normale Superiore, Piazza dei Cavalieri 7, 56100 Pisa, Italy
[c] Dipartimento di Lingue, Letterature e Culture Straniere, Università Roma Tre, Via Ostiense 236, 00146 Rome, Italy
[d] Center for Neurocognition, Epistemology and theoretical Syntax (NEtS), School of Advanced Study (IUSS), Piazza della Vittoria 15, 27100 Pavia, Italy

ARTICLE INFO

Article history:
Received 8 March 2017
Received in revised form 24 July 2017
Accepted 22 August 2017

Keywords:
Presupposition
N400
P600
P200
Linking and updating
Experimental pragmatics
Neuropragmatics

ABSTRACT

This study investigates the neurophysiological correlates of presupposition processing in conditions of satisfaction and accommodation, comparing two types of triggers: definite descriptions and change-of-state verbs. Results showed that, for both types, the accommodation of presuppositions is associated with a biphasic N400-P600 pattern at the processing point. With definite descriptions, we observed a more clear involvement of the N400, while for change-of-state verbs the costs of accommodation were associated with a more pronounced P600. Moreover, when conveyed by change of state predicates, presuppositions seem to elicit also a P200 visible already at the trigger verb. The data nicely fit into the Linking-Updating model and support two main conclusions. First, presupposition accommodation is a sequential process unfolding through a biphasic ERP pattern presumably related to search for antecedent and discourse update. Second, the kind of presupposition trigger seems to affect the cognitive cost of presupposition accommodation at different processing times, with definite description capitalizing more on the earlier search for antecedent and change-of-state verbs capitalizing more on the later updating of the discourse mental model with the presupposed information. Overall, our findings suggest that the brain understands information taken for granted by going through a process whose time course involves several phases, differently modulated based on specific linguistic expressions.

* Corresponding author. University of Genoa, DISFOR - Department of Educational Sciences, Psychology Unit, Corso Podestà, 2, Room 3C3, Laboratory of Psychology of Language, 16128 Genova, Italy.

E-mail addresses: filippo.domaneschi@unige.it (F. Domaneschi), paolo.canal@iusspavia.it (P. Canal), viviana.masia@uniroma3.it (V. Masia), edoardo.lombardivallauri@uniroma3.it (E. Lombardi Vallauri), valentina.bambini@iusspavia.it (V. Bambini).

[1] Equally contributed.

http://dx.doi.org/10.1016/j.jneuroling.2017.08.002

1. Introduction

1.1. Presupposition satisfaction and accommodation

Traditionally, linguists and philosophers of language characterize presuppositions as background information which is communicated as taken for granted. For example, the utterance in (1) introduces the presupposition in (1a):

(1) John stopped smoking
(1a) John used to smoke

Presuppositions are usually carried by ‘presupposition triggers’, namely, lexical items and syntactic constructions that, when used in an utterance, activate presuppositions – e.g., definite descriptions, focus sensitive particles, or, as in (1), change-of-state verbs like *give up* etc. (Karttunen, 1974; Kiparsky & Kiparsky, 1971; Levinson, 1983; Stalnaker, 1974).

Presuppositions are generally considered a condition for the appropriateness of an utterance: according to Stalnaker (2002), a sentence *p* presupposes *q* if the use of *p* would be inappropriate when *q* does not belong to the *common ground* in a conversation. If *q* is entailed by the context, then the presupposition *q* is said to be *satisfied*. Conversely, if *q* does not belong to the common ground this leads to *presupposition failure*. If failure occurs, speakers are supposed to *accommodate* the presupposition (Heim, 1982; Lewis, 1979) in order to make sense of the utterance. Accommodation is the process whereby the content of a presupposition that is not satisfied is introduced into the discourse in order to make it possible for the context to be updated with the assertive component of the utterance.

1.2. Processing presupposition accommodation

Compared with other topics of pragmatic research such as figurative language and scalar implicatures (Bambini & Resta, 2012; Noveck & Reboul, 2008; Sauerland & Schumacher, 2016), in the field of Experimental Pragmatics (i.e., the study of pragmatics via empirical methods) an experimental research line on presuppositions is still underdeveloped. In recent years, however, some empirical studies investigated presuppositions using mainly behavioral methods and focusing on processing times. Overall, the contemporary experimental literature offers some preliminary results in support of the idea that presuppositions are processed rapidly in online language comprehension (Schwarz, 2015). Moreover, several studies showed that whether or not a presupposition is processed easily and fast depends first of all on the kind of information available in the context. In particular, an interesting finding suggested by the current psycholinguistic literature is that presupposition accommodation, i.e., the case of presupposition not entailed by the context, as opposed to presupposition satisfaction, seems to involve longer processing times associated with higher cognitive costs, reflecting a process of context repair. Behavioral works on definite descriptions showed, on the one hand, that the processing of context non-supported definite descriptions takes longer than that of contextually supported definite descriptions (Arnold, Wasow, & Losongco, 2000; Haviland & Clark, 1974; Yekovich & Walker, 1978), and, on the other, that falsified definite descriptions are costlier to process than asserted information (Schwarz, 2015). The idea that accommodation involves higher processing costs was also confirmed by evidence on other categories of presupposition triggers. For instance, presupposition accommodation in sentences containing the German additive particle *auch* (“too”) generally entails longer reading times compared with presupposition satisfaction in intra-sentential contexts (Schwarz, 2007). Tiemann et al. (2011) showed that not only presupposing sentences in general require longer reading times than non-presupposing sentences, but, independently of the trigger in use, a presupposition falsified by the context requires longer processing times than a verified presupposition, and that presupposition accommodation takes longer than the processing of falsified presuppositions. More recently, Tiemann (2014) used a word-by-word self-paced reading task to look at *wieder* (“again”) in conditions of presupposition satisfaction versus accommodation, finding longer reading times on the critical word in the latter case.

Presupposition accommodation constitutes therefore a condition that involves higher processing costs. Two factors, moreover, may affect presupposition accommodation. First, the type of presupposition trigger: Domaneschi, Carrea, Penco, and Greco (2014) showed that the accommodation of the presuppositions activated by triggers like definite descriptions and factive verbs is more mandatory than the accommodation of more optional presuppositions like those triggered by focus sensitive particles (e.g., *too*) and iterative expressions (e.g., *again*). Second, the role of plausibility: Frazier (2006) found an effect of plausibility on the reading times of plausible and implausible definite descriptions, while Singh, Fedorenko, Mahowald, and Gibson (2016) showed that accommodation results inappropriate only in implausible contexts.

1.3. Neurolinguistic evidence on presupposition accommodation

Event-Related brain Potentials (ERPs) are voltage changes of the electrical activity of the brain recorded from the scalp (EEG) that are induced by an external stimulation or an internal cognitive event (e.g., Rugg & Coles, 1995). The study of ERPs has provided fundamental evidence on which kinds of cognitive costs a reader may incur during language comprehension, focusing on the functional role of ERP components (e.g., Luck & Kappenman, 2012). The available literature about the cognitive processes underlying ERP components can help in clarifying the nature and the time-course of the cognitive costs of presupposition accommodation. As it will be detailed below, the handful of studies investigating the processing of presuppositions reported effects on the two most studied ERP components in language processing, the N400 and the P600.

The N400 (Kutas & Hillyard, 1980) is a negative deflection of the ERPs peaking approximately 400 ms after stimulus onset and maximally distributed in Centro-Parietal electrodes. Perhaps the most established finding on sentence comprehension mechanisms is that the less expected a word is, the larger the N400. Almost four decades of research allowed to frame N400 related processes within the realms of semantic memory access/retrieval (Kutas & Federmeier, 2000; Lau, Phillips, & Poeppel, 2008) and semantic integration/unification (Hagoort & Van Berkum, 2007) mechanisms. Set aside the evidence on the processing of words in isolation or figurative language (more closely investigating the access to or the retrieval of information from semantic memory) we want to stress that contextual expectations are crucial for both accounts, and that a fundamental source of expectations is the incremental representation of the ongoing discourse. A large number of studies showed that words providing information that is inconsistent with the discourse representation (e.g., Nieuwland & Van Berkum, 2006), individual world knowledge (e.g., Hagoort, Hald, Bastiaansen, & Petersson, 2004), individual beliefs (e.g., Van Berkum, Holleman, Nieuwland, Otten, & Murre, 2009) or the hearer's information about the speaker (e.g., Van Berkum, Van den Brink, Tesink, Kos, & Hagoort, 2008) affect the amplitude of the N400.

The P600 component (Kim & Osterhout, 2005) is a positive deflection of the ERPs that does not show a clear peak and affects the brainwaves during a longer time interval ranging from 500 to 1000 ms after word onset (e.g., Friederici, 2011), with a scalp distribution that may vary (typically Posterior but sometime more Fronto-Central, (cf. Kaan & Swaab, 2003). Due to its sensitivity to syntactic errors or structural/syntactic complexity, the P600 was initially taken to reflect purely syntactic mechanisms of reanalysis or revision (e.g., Kaan & Swaab, 2003). More recently, after the discovery of P600 effects for semantic reversal anomalies (e.g., Kim & Osterhout, 2005; Van Herten, Kolk, & Chwilla, 2005), or for its involvement in the interpretation of pragmatic phenomena such as ironical utterances (e.g., Regel, Gunter, & Friederici, 2011; Spotorno, Cheylus, Van Der Henst, & Noveck, 2013) and figurative meanings (Bambini, Bertini, Schaeken, Stella, & Di Russo, 2016; Canal, Pesciarelli, Vespignani, Molinaro, & Cacciari, 2017), the interpretation of the functional role of this component has changed and is now taken to reflect a) difficulties in integrating semantic and syntactic information (Friederici, 2011; Kuperberg, 2007), b) the integration and update in discourse and conversation via inferential processes (Burkhardt, 2007; Hoeks & Brouwer, 2014) or c) attention reorientation processes (Sassenhagen & Bornkessel-Schlesewsky, 2015).

With respect to presupposition, evidence is still limited to very few studies, mainly on definite descriptions. In Burkhardt (2006), the processing of definite phrases was probed relative to three types of context-dependency of a definite phrase (*given*, *bridged* and *new*) differing for the degree of availability of the antecedent of the definite description. Burkhardt found that new definite phrases elicited N400 and P600 differences when compared with already given definite phrases. Also Kirsten et al. (2014) found a N400/P600 pattern in response to definite descriptions read in conditions where more referents were available in the context. In a following study, Burkhardt (2007) compared the processing of definite phrases following three types of context, varying in terms of inferential demands needed to form a relationship between the definite phrase (e.g., *the pistol*) and the information provided by the previous sentence, which could either be necessary (e.g., *a student was shot*), probable (e.g., *a student was killed*) or inducible (e.g., *a student was found dead*). The increased efforts for drawing more demanding inferences were associated with larger P600 components, while no N400 effects were observed. On the basis of these and other findings, Schumacher and Hung (2012) and Wang and Schumacher (2013) proposed a model in which the N400 indexes Discourse Linking mechanisms, i.e., the attempts of locating an entity or a referent in the ongoing discourse representation, and late positive ERP effects, such as those involving the P600 component, reflect Discourse Updating mechanisms, i.e., the correction, modification or enrichment of the current representation of the discourse. As pointed out by Schumacher and Hung (2012: 298), both linking and updating processes massively contribute to the construction of the discourse representation. Notably, the processor first links a referent to prior discourse and then updates the mental representation of the discourse with that referent. Linking mechanisms are also deeply influenced by the salience and likelihood of the referent and are generally discourse-dependent because they come about in cohesion building. On the contrary, updating mechanisms are less dependent on prior discourse and mainly reflect strategies of maintenance and updating of the discourse representation structure.

In a recent study, Masia, Canal, Ricci, Lombardi Vallauri, and Bambini (2017) investigated how presupposed content is accommodated with the discourse model using definite descriptions (*the migration* in a context in which the noun is contextually appropriate but no explicit mention to a migration is provided) compared with assertive controls (*there was a migration* following the same context) and reported a larger N400 for the former case. The larger N400 and the absence of a P600 effect for presuppositions compared with assertions was interpreted as related to both conditions providing information that was equally new and similarly compatible with the preceding context, thus requesting the same amount of context updating. The larger N400 reflected a more elaborate search of the antecedent (i.e., discourse linking) that was elicited by the fact that presuppositions convey information communicated as already shared by the interlocutors. This expectation forced the reader to perform a more extended search for the antecedent in the discourse model.

Neurolinguistic research on other kinds of presupposition triggers is still on the way to be conducted. A recent study by Jouravlev et al. (2016) tested the time course of presuppositions projected by the temporal trigger *again* in conditions of satisfaction and violation. Positive effects with an early onset around 300 ms were observed in response to presupposition violation. Although useful for a more in-depth understanding of the electrophysiological responses to different presupposition triggers, this finding is less informative about the accommodation of presuppositions triggered by this adverb (because no condition of context-non supported presuppositions was tested), and so it weakly contributes to the elaboration of working hypotheses for the present study.

The literature reviewed allows for a preliminary characterization of the electrophysiological components associated with the accommodation of presuppositions triggered by definite descriptions. The N400 is more tightly related to difficulties in *linking* the presupposed content to the preceding context set. Differently, the P600 reflects the resolution of a prior incongruence and, in particular, the *updating* of the presupposed information into the discourse mental model.

To the best of our knowledge, what is still lacking in the neurolinguistic research on presuppositions is a more direct comparison between presupposition accommodation and satisfaction where different trigger types are taken under exam, to explore potential differences in cognitive mechanisms and in time-course.

1.4. Research questions and predictions

This study is aimed at investigating, for the first time, the brain signature of presupposition accommodation vs. satisfaction by comparing two different types of triggers: definite descriptions and change of state verbs. The comparison is theoretically relevant since it might reveal possible analogies or differences in the mechanisms underlying the processing of different presupposition triggers in two different conditions of contextual dependency.

The reason for focusing on definite descriptions and change-of-state verbs as triggers of interest is threefold. First, definite descriptions and temporal triggers are the two trigger types that have been mainly considered by the neurolinguistic literature. Second, it was shown that temporal triggers like iterative expressions and change-of-state verbs involve higher processing costs in presupposition accommodation compared with definite descriptions (Domaneschi et al., 2014). In terms of cognitive demands, such a difference allows for predicting variations in terms of neurophysiological responses. Third, although the only ERP study on temporal triggers on the market concerns the iterative expression *again* (Jouravlev et al., 2016), *again* does not allow for a direct comparison with definite descriptions since, as it was broadly argued in the semantic literature, while the presuppositions of *again* and *too* are not necessarily accommodated, the presuppositions of definite descriptions, factive verbs and change-of-state verbs like *stop* or *give up* are mandatory, as they provide a meaningful contribution to the assertive component of an utterance (Beaver & Zeevat, 2007; Glanzberg, 2003; Kripke, 2009). Therefore, from a semantic point of view, change-of-state verbs are the temporal trigger that best fits a neurolinguistic comparison with definite descriptions.

Based on the above premises, three research questions will be addressed:

(RQ1). What are the neurophysiological correlates involved in presupposition accommodation with definite descriptions?

The first issue to be addressed concerns a further clarification of the ERP components associated with definite descriptions in presupposition accommodation, comparing a condition of presupposition satisfaction with a condition of accommodation, with no intermediate state of context-dependency of the critical referent, as this latter condition would be less relevant to the experimental rationale of our study. The crucial point at stake, in response to the available neurolinguistic literature, is to determine whether the accommodation of definite phrases compared with the presupposition satisfaction elicits an N400, a P600 or an N400/P600 pattern. The most probable scenario is that accommodation, compared with satisfaction, capitalizes on a biphasic N400/P600 pattern as observed when comparing new vs given referents (Burkhardt, 2006), since the accommodation vs satisfaction contrast involves the modulation of the givenness of the noun phrase. Yet we cannot exclude to replicate the N400 only pattern observed for presupposition accommodation compared with assertion, as in Masia et al. (2017), or a P600 only, as observed for referents with different contextual fit (Burkhardt, 2007).

(RQ2). What are the ERP components associated with the accommodation of change of state verbs?

Tiemann et al. (2011: 583) noted that, in studying presupposition triggers, there are some crucial regions of interest in a presupposing sentence where presupposition processing is expected to occur. Two of these regions may be defined as the *triggering point* and the *computation point* (or *processing point*). Take for instance the definite description in (2):

(2) *The waiter* is nice
(2a) There is a (unique, identifiable) waiter

and the change-of-state verb in (3):

(3) Mark *stopped* smoking
(3a) Mark used to smoke

In the case of the definite description in (2), the triggering point is the definite article *the* that activates a presupposition of existence, uniqueness and identifiability (Hawkins, 1978):

(2b) $[[\text{the}]]w = \lambda f\langle e,t\rangle$: there is exactly one x s.t. $f(x) = 1$. the unique y s.t. $f(y) = 1$.

The head noun *waiter*, conversely, represents the computation point of the presupposition where the hearer computes the content of (2a).[2] Similarly, in the phrase *stopped smoking, stop* is the triggering point where the presupposition of existence of a previous state of affair is activated:

(3b) [[give up]]w,t = λP<i,et>. λx. ∃t': t' < t & P(t')(x). ~P(t)(x)

The critical region *smoking* is the computation point where, again, the content of the presupposition is computed.[3]

Since the ERP literature on the accommodation of definite descriptions highlighted the effect of electrophysiological correlates on the head noun of the definite phrases, i.e., on the computation point, with change-of-state verbs a reasonable expectation is to find an effect of the manipulation on the final word of the verb phrase, e.g., *smoking* in (3), which we assume to be the critical computation point for this trigger type. Keeping on with building the predictions based on the literature on definite phrases, we can predict that, also with change-of-state verbs, accommodation, compared with satisfaction, would capitalize on a biphasic N400/P600 pattern. Yet specific aspects of the change-of-state verbs might elicit different modulation of the biphasic response, as investigated in (RQ3).

(RQ3). Is there any neurophysiological difference between the accommodation of definite descriptions and that of change-of-state verbs?

On the basis of the assumptions presented in the previous section, if the N400 can be considered an index of discourse linking mechanisms (Masia et al., 2017; Wang & Schumacher, 2013), then a possible scenario can be predicted. According to Zeevat (1992), definite descriptions are considered an example of resolution triggers like *again*, *even* or *too*, which involve the anaphoric retrieval of an entity or an eventuality from the common ground; conversely, change-of-state verbs (e.g., *stop*) are a lexical trigger which directly encode in their conventional meaning a precondition for their asserted content. We predict therefore that the accommodation with definite descriptions could elicit a more demanding linking process than the accommodation with change-of-state verbs, reflecting a more extended search for the antecedent of the anaphoric expression in the context set.

Instead, if the P600 can be considered an index of discourse updating mechanisms (Wang & Schumacher, 2013), then a larger P600 effect can be predicted with change-of-state verbs. It has been shown that temporal triggers like iterative expressions and change-of-state verbs involve higher processing costs compared with definite descriptions and factive verbs, possibly due to the fact that they entail the mental representation of temporally displaced events (Domaneschi et al., 2014; Tiemann, 2014). Therefore, it is reasonable to predict that the process of recalling a temporally displaced event might require a surplus of cognitive resources, which makes discourse model updating more demanding.

2. Material and methods

2.1. Participants

Twenty-four university students (12 men–12 female, mean age = 23.6, SD = 3) took part in the experiment and were paid for their participation. All subjects were right-handed (as attested by the Edinburgh Handedness questionnaire: mean laterality = 0.90; cf. Oldfield, 1971), native speakers of Italian, with normal or corrected to normal vision and no history of neurological or psychiatric disorders. Informed consent was obtained. The experiment was approved by the local Ethics Committee (Comitato Etico Area Vasta Nord Ovest, Azienda Ospedaliero-Universitaria Pisana).

2.2. Stimuli

Experimental materials were composed of a set of 120 pairs of short stories in Italian. Each story presented fictional ordinary circumstances and was composed of a *context sentence 1*, a *context sentence 2* and a *target sentence.* Each target sentence contained a presupposition trigger: 60 pairs with definite descriptions (DDs) and 60 pairs with change-of-state verbs (CSs). Each pair included 2 conditions, *satisfaction* (SAT) and *accommodation* (ACC), obtained by manipulating the information provided by *context sentence 1*, which either made explicit the content presupposed by the trigger (i.e., SAT) or did not satisfy the presupposition of the target sentence (i.e., ACC). *Context sentence 2* and the *target sentence* were kept constant across conditions (see Table 1 for an example pair; see Appendix 1 for a representative sample of DD pairs and Appendix 2 for a representative sample of CS pairs).

[2] Formally: [[the waiter]]w = [[the]]w([[waiter]]w) = [λf<e,t>: there is exactly one x s.t. f(x) = 1. the unique y s.t. f(y) = 1] ([λz. z is a waiter at w]) = there is exactly one x s.t. [λz. z is a waiter at w](x) = 1. the unique y s.t. [λz. z is a waiter at w](y) = 1 = there is exactly one x s.t. x is waiter at w = 1. the unique y s.t. y is a waiter at w = 1.

[3] As a result, we have the presupposition that someone was smoking at a previous time: [[give up smoking]]w,t = [[give up]]w,t ([[smoking]] w) = [λP<i,et>. λx. ∃t′: t′ < t & P(t′)(x). ~P(t)(x)] ([λt″.λy. y smokes at t″]) = λx. ∃t′: t′ < t & [λt″.λy. y smokes at t″](t′)(x). ~[λt″.λy. y smokes at t″](t)(x) = λx. ∃t′: t′ < t & x smokes at t′. ~ x smokes at t.

Table 1
Example of stimulus pairs for both trigger types (DD: Definite Descriptions; CS: Change of State) in the satisfaction (SAT) and accommodation (ACC) conditions (original Italian and English translation; target sentence with English glossa). Antecedent in the context sentence 1 for the SAT condition in italics; target word in the target sentence in underlined bold.

DD set	**SAT**		**ACC**
Context sentence 1	Nell'ufficio di Paolo lavorava un *grafico* davvero scontroso. [Eng. In Paolo's office, there used to be a very bad-tempered graphic *designer*.]		Nell'ufficio di Paolo lavorano diversi impiegati [Eng. In Paolo's office there are many employees.]
Context sentence 2	La sede necessita di consulenti per diversi rami dell'attività. [Eng. The office needs consultants for several branches of the activity.]		
Target sentence	Per esubero del personale circa un mese fa il **<u>grafico</u>** è stato messo in mobilità Due.to excess of.ART staff about a month ago the graphic **<u>designer</u>** put.PASS.PST.3SG PREP.redundant 'Due to overstaffing problems, about a month ago the graphic designer was made redundant'		
Verification questions	**Question 1 (target):** *C'era un grafico nell'ufficio di Paolo?* [Eng. Was there a graphic designer in Paolo's office?]	**Question 2 (filler):** *Nell'ufficio di Paolo, l'attività è suddivisa in diversi rami?* [Eng. Is Paolo's office divided in several branches?]	**Question 3 (filler):** *Nell'ufficio di Paolo il personale è in esubero?* [Eng. Is the personnel redundant in Paolo's office?]
CS set	**SAT**		**ACC**
Context sentence 1	Durante gli anni di liceo Simone *ha consegnato pizze* per il ristorante di suo zio. [Eng. During his high school years, Simone *used to deliver pizza* for his uncle's restaurant.]		Durante gli anni di liceo, Simone ha lavorato nel ristorante di suo zio. [Eng. During his high school years, Simone used to work in his uncle's restaurant.]
Context sentence 2	Spesso finiva i suoi turni oltre la mezzanotte. [Eng. He often finished working after midnight.]		
Target sentence	Poi si è diplomato così ha smesso di consegnare **<u>pizze</u>** ed ha rilevato una trattoria Then 3SG.REFL.graduate.PST so 3SG.have stop.PST PREP delivering **<u>pizzas.</u>**PL and 3SG.have take over.PST a small.restaurant 'Then he graduated, and so he stopped delivering pizza and took over a small restaurant'		
Verification questions	**Question 1 (target):** *Simone consegnava pizze?* [Eng. Did Simone deliver pizza?]	**Question 2 (distractor):** *Simone ha lavorato nel ristorante di suo zio?* [Eng. Did Simone work in his uncle's restaurant?	**Question 3 (distractor):** *Simone si è diplomato?* [Eng. Did Simone get the high school diploma?]

A number of aspects were taken into account while building the items. Target sentences were controlled to avoid the presence of logical negation operators as well as conversational and conventional implicatures on or close to the presupposition triggers. Moreover, the overlap of different presupposition triggers in the same target sentence was avoided. The length of the target sentences was fixed at 15 words for all conditions.

Concerning definite descriptions, only DDs with a referential interpretation were used in the experiment (Donnellan, 1966). In order to avoid positional biases, all definite descriptions had their head noun – i.e., their processing point – in 10th position in the target sentence. For the items containing change-of-state verbs, in order to reduce potential biases associated with different semantic properties of the verbs, only one token (i.e., *smettere*, Engl. "to stop") was selected as a trigger. Although verbs such as *stop*, *continue*, *leave*, *arrive*, etc. all denote a change of state, the kind of semantic representation entailed by each of them is somewhat different. Therefore, comparing different types of change-of-state verbs could have resulted in extremely indirect measures on their processing behavior in both SAT and ACC conditions. We also tried to avoid using verbs with a too specific meaning prior to the region of interest represented by the last word of the verb phrase – e.g., instead of saying: "Jane stopped cooking *dinner*" we opted for "Jane stopped making *dinner*", so as to make the occurrence of "dinner" less predictable. We created target sentences with change-of-state verbs falling always in 7th position, whereas the last word of the phrase (the processing point), was always in 10th position, as for DD items. In this way, the regions of interest of DDs and CSs were in the same sentence position, and this was a noun for both trigger types.[4]

Each story was followed by three verification questions: a *target question* verifying the content of the presupposition activated by the target sentence and two *distractor questions* – see Table 1. The target question was always true, while the

[4] Although we did our best to control for confounding variables, we aknowledge two potentially critical aspects in out materials. First, there is a difference between DDs and CSs in terms of internal structure, with the computation point preceded by a determiner in the case of DDs and by an infinite verb in CSs. Future ERP works might overcome this intrinsic limitation by contrasting more syntactically homogeneous categories of triggers like for instance CSs and factive verbs. The second difference concerns morphological and semantic features of the nouns used as computation points: with DDs we used singular nouns; conversely, with CSs we mainly used plural nouns, mostly countable and in few cases uncountable.

number of the true and false responses was counterbalanced across distractors. In order to distract the subjects from the purpose of the target questions, all verification questions were presented in a randomized order within each trial. Due to the purpose of the study, these verification questions allowed to ascertain that critical presuppositions were actually accommodated by the subjects.

Two lists of 120 stories each were created and stimuli were distributed according to a Latin Square design, so that participants who read a story in the ACC condition never read the same story in SAT. Forty-five fillers were added to each list, consisting of 3-sentence stories and containing other categories of presupposition trigger, for a total number of 165 stories in each list. The order of presentation of the stimuli was pseudo-randomized twice for each list.

2.3. Measures and pre-tests

All experimental items were controlled for a number of potentially confound variables. For all stories, we calculated length (number of words) as well as readability based on the Gulpease index (Piemontese, 1996). Similar to the Flesch test elaborated for English, the Gulpease index is calibrated to the Italian language and is used to measure text complexity in psycholinguistic studies employing Italian materials (e.g., Bambini, Resta, & Grimaldi, 2014; Masia et al., 2017). The Gulpease index calculates the readability of a text based on the number of letters, the number of words, and the number of sentences. The calculation returns values indicating the ease of reading for different populations with different degrees of formal education. Average Gulpease level for our stimuli was in the range of 40–60, meaning easy independent reading for individuals with high school level education – see Table 2 for mean values of Gulpease and length. Two-way ANOVAs with Trigger (DD and CS) and Condition (SAT and ACC) as factors were run to exclude potential asymmetries in length and readability across conditions. Results showed no interaction between the two independent variables in terms of length ($F(1,116) = 1.95$, $p = 0.16$) nor in terms of readability ($F < 1$).

In a series of pre-tests run on-line, materials were judged on 5-point scales according to parameters of *naturalness*, *plausibility*, and *predictability*. All participants were matched in age and education with those of the electrophysiological session. The naturalness questionnaire was administered to 30 participants, the predictability rating to 14 participants and the plausibility rating to 15 participants. The naturalness parameter was measured on all critical stimuli in both the SAT and ACC. For the test, participants were asked to judge the naturalness of the target sentences with respect to the preceding context, to assess whether target sentences represented a good continuation in terms of discourse felicity. For the plausibility parameter (cf. Singh et al., 2016), participants were asked to rate how plausible was the existence of the referent or event mentioned in the target sentence. They were thus presented with questions like *How plausible is it to find a graphic designer in an office?* for the DD set, or *How plausible is it that a person used to deliver pizza?* for the CS set. Finally, to make sure that the degree of expectedness of the presupposition in the ACC condition was low enough to prevent the subjects from deriving the critical presupposition from prior contextual information, a predictability rating was run only on critical stimuli in the ACC condition. For this test, participants were presented with the first two context sentences of each experimental text and were then asked to judge how predictable the referent or event presupposed in the target sentence was. For example, for the DD set in ACC, they were presented contexts likes *In Paolo's office there are many employees. The office needs consultants for several branches of the activity* followed by a target sentence like *How predictable is it that there is a graphic designer in Paolo's office?*. For the CS set in ACC, they read contexts like *Matteo has never looked after his health. Over the past few months, he has not been very well* followed by a question, e.g. *How predictable is it that Matteo used to buy cigars?*. All materials were rated as highly natural and plausible (range of acceptance 3–5), and as weakly predictable (range of acceptance 1–3) (see means in Table 2). Two-way ANOVAs were conducted to assess the crossed effect of Trigger (DD and CS) and Condition (SAT and ACC) for the Naturalness measure ($F(1,116) = 1.6$, $p = 0.18$). Results showed no significant interaction between the two independent variables. Since only the Trigger factor was considered for the plausibility and predictability parameter, two one-way ANOVAs were run on these measures to test the interaction between the DD and CS trigger types. Results again showed no significant effect of trigger type, neither for plausibility ($F(1,118) = 2.84$, $p = 0.09$) nor for predictability ($F(1,118) = 1.52$, $p = 0.22$). The absence of statistically significant results suggests that materials were overall natural and easily readable for participants, and that the presuppositions were on the whole plausible and weakly predictable (when accommodated).

Table 2
Means and standard deviations on pre-testing measurements.

	DD		CS		Filler
	SAT	ACC	SAT	ACC	
Gulpease	53.5 (17.3)	43.4 (4.1)	65 (4.7)	47.2 (16)	52 (6.3)
Length	35.8 (3.2)	33.4 (3.1)	35.9 (2.5)	33.8 (3.2)	31.3 (3)
Naturalness	4.2 (0.2)	4.1 (0.2)	4.1 (0.2)	4.2 (0.2)	–
Predictability	–	2.2 (0.2)	–	2.1 (0.2)	–
Plausibility	4.2 (0.2)		4.3 (0.2)		

2.4. Procedure

Participants sat approximately 60 cm from the display. The room had normal lighting. Only a keyboard (no mouse) was available for performing the experiment. During the experiment, the two context sentences were presented as a whole in the centre of the computer screen, while the target sentences were presented word-by-word for 400 ms with a 200 ms inter-stimulus interval (Fig. 1A). Participants moved from the first to the second sentence, and from the second to the target sentence in a self-paced manner using the space bar. ERPs were time-locked to the presentation of the target word (head noun for definite descriptions and last word of the verb phrase for change-of-state verbs) in the target sentence (Fig. 1B). After having read the stories, participants answered the three verification questions by pressing TRUE/FALSE buttons (green/red) on a keyboard.

2.5. EEG recordings and ERP pre-processing

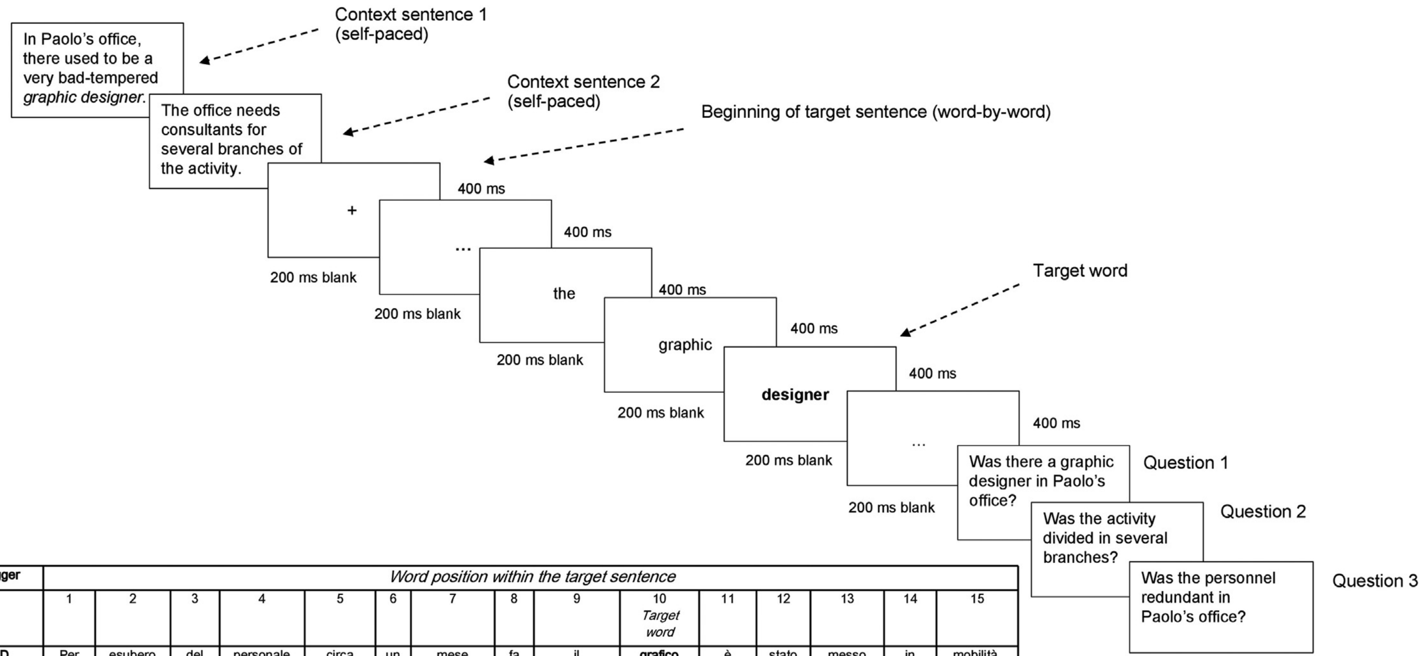
EEG was recorded continuously from 59 electrodes placed on the scalp according to the 10–20 International System (Fp1, Fpz, Fp2, AF7, AF3, AF4, AF8, F7, F5, F3, F1, Fz, F2, F4, F6, F8, FT7, FC5, FC3, FC1, FC2, FC4, FC6, FT8, T7, C5, C3, C1, Cz, C2, C4, C6, T8, CP5, CP3, CP1, CPz, CP2, CP4, CP6, P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO7, PO3, POz, PO4, PO8, O1, Oz, O2, M1, M2). The signal was sampled at 512 Hz using a Brain Amp®System (Brain Products GmbH, Gilching, DE), which amplified, recorded and stored the EEG signal on the acquisition computer. Four electrodes were placed around the eyes (below and above the right eye and at the outer canthi of both eyes) to monitor eye-movements. The EEG signal was online referenced to an electrode close to the Vertex and offline re-referenced to the average activity recorded from the two mastoids. Two opensource toolboxes (EEGLAB, Delorme & Makeig, 2004; FieldTrip, Oostenveld, Fries, Maris, & Schoffelen, 2010) were used for EEG preprocessing in Matlab® (The MathWorks, Inc, Natick, US) environment. The EEG signal was time-locked to the presentation of the critical word and ERP epochs (from −350 to 1200 ms) were extracted, and further filtered using a conservative band pass filter from 0.1 to 40 Hz (Tanner, Morgan-Short, & Luck, 2015). Eye-related activity was corrected using ICA decomposition (e.g., Mennes, Wouters, Vanrumste, Lagae, & Stiers, 2010). We used a semi-automatic rejection procedure, by which epochs with voltage values analyzed according to a fixed threshold ($\pm 150\ \mu V$) were rejected and remaining artifacts were identified manually, with visual inspection. One participant was excluded because less than 60% of the epochs were retained. The overall rejection rate for the remaining participants was 16.7%, and for each single subject ERP a similar number of epochs per condition was retained (CS-SAT = 25.2; CS-ACC = 25.6; DD-SAT = 25.2; DD-ACC = 25.1). A 200 ms pre-stimulus interval was used for baseline correction. To further explore the processing costs of presupposition triggered by CS verbs, for CS passages only, we performed an additional analysis extracting longer EEG epochs, spanning from −700 to 2800 ms around the presentation of the verb, up to the computation point, to capture possible earlier effects. In this longer epoch we used the same pre-processing routine described above, which resulted in a slightly higher rejection rate (19.13%): we retained on average 23.69 epochs for CS-SAT and 24.17 for CS-ACC.

2.6. Statistical analysis

Statistical analyses were carried out in R (R Core Team, 2015). Behavioral data were analyzed within the mixed models framework (e.g., Baayen, Davidson, & Bates, 2008) using *lme4* package (Bates, Maechler, Bolker, & Walker, 2015): categorical True/False questions were analyzed with generalized linear mixed models for binomial responses (e.g., Jaeger, 2008), and reaction times were analyzed with linear mixed models (e.g., Baayen & Milin, 2010). To assess the reliability of the interaction between Trigger Type and Condition we computed likelihood ratio tests between two nested models fitted using Maximum Likelihood estimation (Zuur, Ieno, & Smith, 2007): one model included the main effects associated with the two experimental factors and the other included the interaction between the two. We reported chi-square (χ^2) statistics to describe how the inclusion of the interaction term brought to a change in the goodness of fit. The random structure allowed for intercept adjustments for each participant and item, and by-participant slope adjustments for Trigger Type. The random structure was not maximal (Barr, Levy, Scheepers, & Tily, 2013) as the inclusion of additional terms induced convergence failures, so it was determined on grounds of feasibility (Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2015).

For ERP data analysis we used Analysis of Variance. The dependent variable was the single-participant average voltage amplitude in two time-windows of interest: effects on N400 and P600 components were evaluated during canonical latency intervals (300–500 ms and 600–900 ms, respectively). To account for the scalp distribution of the topographic differences in the ERP response, two types of analyses (Lateral and Midline) were carried out on a subset of 43 electrodes. In the Lateral analysis 36 electrodes were organized into two topographical dimensions (Fig. 2): Hemisphere (two levels, 18 electrodes each: Left, Right) and Longitude (three levels, 12 electrodes each: Frontal, Central and Parietal). In the Midline analysis all 7 electrodes along the midline were organized along one Longitudinal dimension. In the Lateral analysis we carried out two four-way repeated-measures ANOVAs with Trigger Type (Definite Description, Change-of-State), Condition (ACC, SAT), Longitude (Anterior, Central, Posterior) and Hemisphere (Left, Right) as within-subjects factors. In the Midline analysis two three-way ANOVAs were carried out with Trigger Type, Condition (ACC, SAT) and Longitude (FPz, Fz, Cz, CPz, Pz, Poz, Oz). We adopted Greenhouse-Geisser correction against violations of sphericity (corrected *p* value and uncorrected degrees of freedom are reported).

(A)



(B)

Trigger	*Word position within the target sentence*														
	1	2	3	4	5	6	7	8	9	10 *Target word*	11	12	13	14	15
DD	Per	esubero	del	personale	circa	un	mese	fa	il	**grafico**	è	stato	messo	in	mobilità.
	[Eng. Due to overstaffing problems, about a month ago the graphic designer was made redundant.]														
CS	Poi	si	è	diplomato	così	ha	smesso	di	consegnare	**pizze**	e	ha	rilevato	una	trattoria.
	[Eng. Then he graduated, so he stopped delivering pizza and took over a restaurant.]														

Fig. 1. Stimulus presentation procedure. Panel A shows presentation timing for each story. Panel B focuses on the target sentence, displaying the position of the target word for both types of triggers: Definite Descriptions (DD) and Change of State verbs (CS).

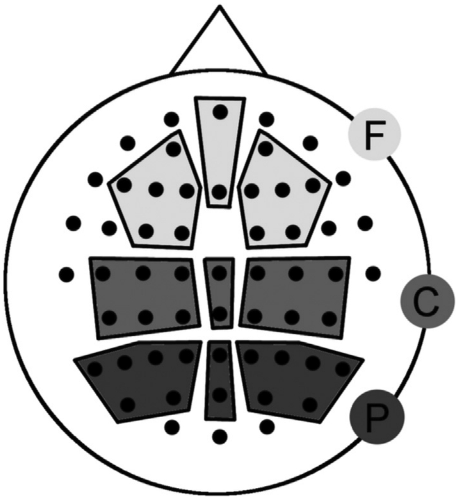


Fig. 2. Topographic sub-regions selected for ERP data analysis (Frontal: F; Central: C; Parietal: P).

For the additional analysis of the longer epoch for CS passages, we used a non-parametric permutation test (Maris & Oostenveld, 2007) implemented in fieldtrip (Oostenveld et al., 2010). Non-parametric tests have "considerable and growing appeal" in ERP research (Keil et al., 2014, p. 9; see also; Groppe, Urbach, & Kutas, 2011; Luck & Gaspelin, 2017). The use of this test is particularly suited to highlight voltage differences between conditions when no prior hypotheses on the spatio-temporal differences between two conditions are available. In our data, the input matrix consisted of 1400 time points (from 0 to 2.8 s from the verb presentation) and 57 channels: for each (sensor-time)-pair the two conditions are compared, and to control for Type-1 error due to the large number of comparisons (79800) the permutation test is carried out on clusters of data, in the following manner. A dependent *t*-test between the voltage of the two conditions is performed in each sensor-time sample. Data samples for which the *t*-test exceeds a threshold (here we used alpha $= 0.05$) and are adjacent in time and space (two channels were neighbors if their distance was less than 6 cm) are grouped together in a set of positive and negative clusters. The observed cluster statistic corresponds to the sum of all t values composing the data cluster (maxsum). The permutation test is then performed by collecting data for both conditions in a single set and then randomly partitioning (for 1000 iterations) this set into two equally sized subsets and calculating the test statistic on each random partition, for each cluster of data. A distribution of the test statistic is obtained from the permutation and corresponds to the distribution of the null hypothesis, which is rejected when the observed cluster-based test statistic falls within the highest or lowest 2.5 percentile of the distribution (Maris & Oostenveld, 2007). For each cluster, we reported maxsum and associated p value.

3. Results

3.1. Behavioral results

The models estimates for participants' accuracy to the true-false questions at the end of each passage were 96.65% for CS in SAT, 95.53% for CS in ACC, 94.26% for DD in SAT and 94.75% for DD in ACC. Questions on CS passages were accurately judged slightly more often than questions on DD [+1.58%, $z = 1.90$, $p = 0.056$]. Moreover, the likelihood ratio test revealed a marginally reliable interaction between Trigger Type and Condition [χ^2 (1df) $= 3.61$, $p = 0.057$]. This marginally significant interaction was driven by the fact that for CS accuracy in SAT was higher than in ACC (+1.12%) while for DD accuracy in SAT was lower than in ACC (−0.48%): although numerically small, this difference across levels of Trigger Type resulted to be robust [$\Delta acc = +1.61\%$, $z = +1.96$, $p < 0.05$]. The model's estimates for response times were 1861 ms for CS in SAT, 1907 ms for CS in ACC, 2030 ms for DD in SAT, and 2032 ms for DD in ACC (see Fig. 3). The effect of Trigger Type was significant due to faster response times [-147 ms, $t = -4.15$, $p < 0.001$] to CS (1884 ms) compared with DD (2031 ms). No significant Trigger Type by Condition interaction emerged [χ^2 (1df) $= 1.17$, ns].

3.2. ERP results

Grand averages ERPs for the four conditions are shown in Fig. 4.

3.2.1. N400

Lateral and Midline ANOVAs carried out during the N400 time window attest that ERP differences associated with Condition are reliable and can be interpreted as affecting the N400 component, with ACC eliciting more negative voltages

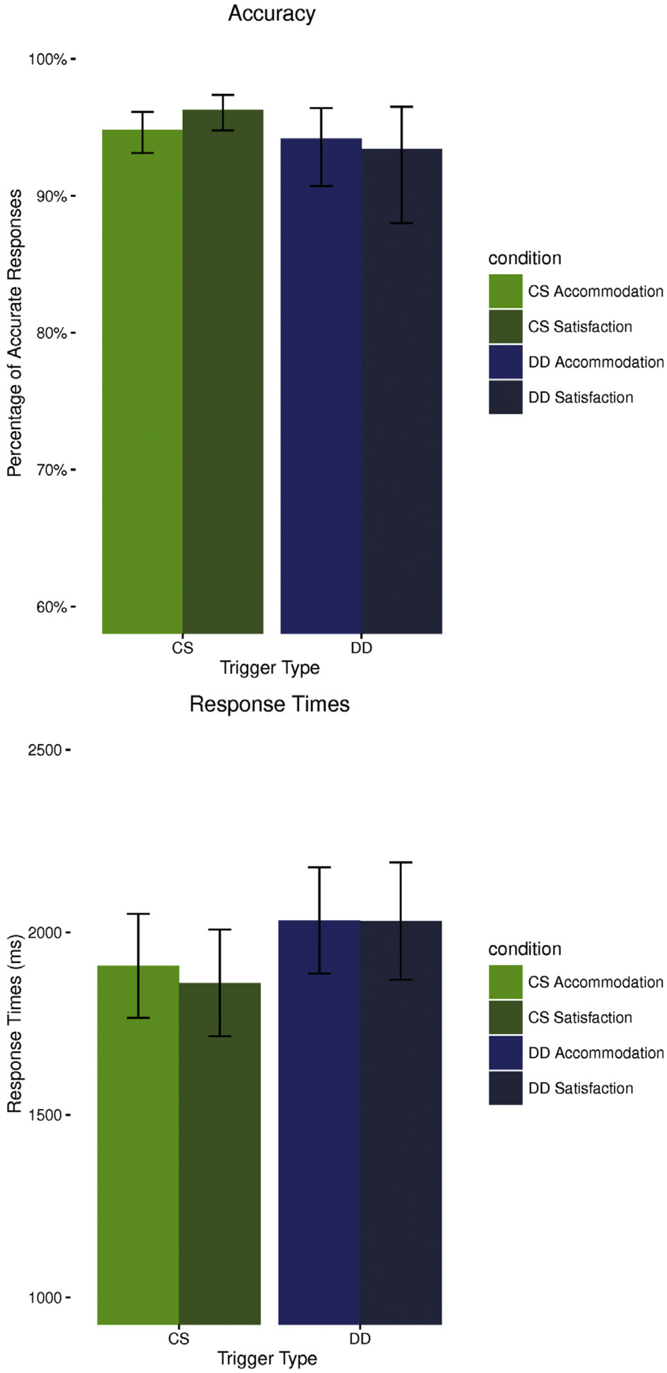


Fig. 3. Reaction times and accuracy results in the behavioral task.

compared with SAT (Table 3). The main effect of Condition is robust [Midline: $F(1,22) = 13.77$, $p < 0.01$; Lateral: $F(1,22) = 12.28$, $p < 0.01$] and the significant Condition by Longitude interaction [Midline: $F(6,132) = 6.62$, $p < 0.01$; Lateral: $F(2,44) = 5.16$, $p < 0.05$] shows that the effect is more prominent over Parietal locations. Lateral ANOVA also revealed a significant Hemisphere by Condition interaction [$F(1,22) = 10.51$, $p < 0.01$], showing that, although the effect of Condition (independently of Trigger Type) is widely distributed, the amplitude differences between ACC and SAT are more pronounced over the Left [$F(1,22) = 12.28$, $p = 0.002$] rather than over the Right [$F(1,22) = 5.69$, $p = 0.026$] hemisphere.

Most importantly, ANOVAs revealed a significant three-way interaction between Trigger Type, Condition, and Longitude [Midline: $F(6,132) = 4.56$, $p < 0.01$; Lateral: $F(2,44) = 6.59$, $p < 0.01$]. To better describe this interaction, two two-way

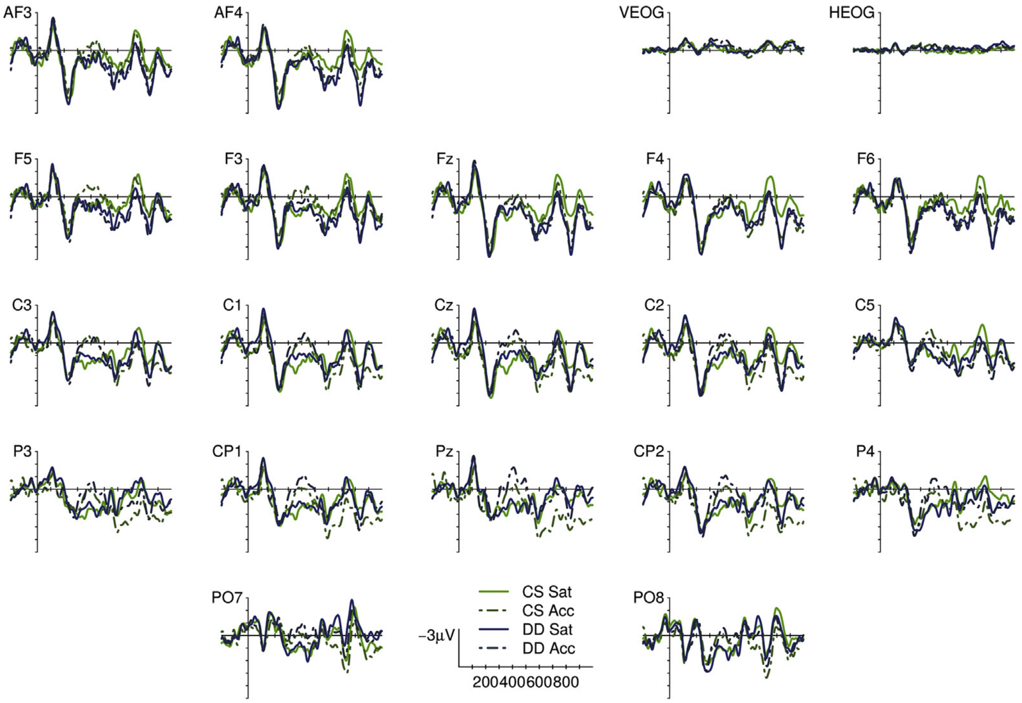


Fig. 4. ERP Grand averages on the critical regions for the Condition and Trigger factors.

Table 3
Lateral and Midline ANOVAs on the N400 amplitude (300–500 ms).

	df	F	p	ε	Eta
Lateral					
Trigger Type	1, 22	0.24	ns		0.000
Condition	1, 22	12.28 **	0.002		0.057
Trigger Type X Condition	1, 22	0.02	ns		0.000
Trigger Type X Longitude	2, 44	3.19^	0.07	0.65	0.003
Trigger Type X Hemisphere	1, 22	0.11	ns		0.000
Trigger Type X Longitude X Hemisphere	2, 44	1.28	0.288	0.66	0.000
Condition X Longitude	2, 44	5.16 *	0.025	0.61	0.004
Condition X Hemisphere	1, 22	10.51 *	0.003		0.001
Condition X Longitude X Hemisphere	2, 44	2.29	0.126	0.80	0.000
Trigger Type X Condition X Longitude	2, 44	6.59 **	0.008	0.73	0.004
Trigger Type X Condition X Hemisphere	1, 22	1.45	0.241		0.000
Trigger Type X Condition X Longitude X Hemisphere	2, 44	1.52	0.234	0.90	0.000
Midline					
Trigger Type	1, 22	1.27	0.272		0.006
Condition	1, 22	13.77 **	0.001		0.057
Trigger Type X Condition	1, 22	0.56	ns		
Trigger Type X Longitude	6, 132	3.37 *	0.037	0.377	0.007
Condition X Longitude	6, 132	6.62 **	0.002	0.355	0.010
Trigger Type X Condition X Longitude	6, 132	4.56 **	0.009	0.411	0.006

(Condition by Longitude) ANOVAs were carried out separately for each Trigger Type (using both Lateral and Midline electrodes). For CS passages, the effect of Condition did not show any difference across levels of Longitude, as attested by the significant main effect [-0.89 μV, $F(1,22) = 6.29$, $p = 0.020$] and by the lack of interaction with Longitude [$F < 1$]. Instead, for DD passages, the effect of Condition was strongly modulated by Longitude, as attested by the significant interaction [$F(2,44) = 16.31$, $p < 0.001$]. In particular, the N400 effect in DD passages was null over Frontal electrodes [-0.31 μV, $F < 1$] but robust over Central [-1.17 μV, $F(1,22) = 7.64$, $p = 0.014$] and Parietal [-1.51 μV, $F(1,22) = 15.56$, $p < 0.001$] electrodes. We then compared the size of the effect of Accommodation (ACC vs SAT) across trigger types in Central and Parietal electrodes, with simple t tests (Bonferroni corrected for four comparisons). Planned contrasts revealed a robust N400 effect in the DD set, and a marginally significant effect for CS (Table 4; Fig. 5).

Table 4
Planned contrasts on the N400 amplitude on central and parietal electrodes. Significant level of alpha = 0.0125.

		ACC vs SAT in DD set	ACC vs SAT in CS set
Central Electrodes	Effect size in μV	−1.17	−1.00
	t(22) value	−2.76 *	−2.41 ^
	95% confidence intervals	−2.04: −0.29	−1.87: −0.14
Parietal Electrodes	Effect size in μV	−1.51	−0.83
	t(22) value	−3.94 **	−2.59 ^
	95% confidence intervals	−2.31: −0.71	−1.49: −0.17

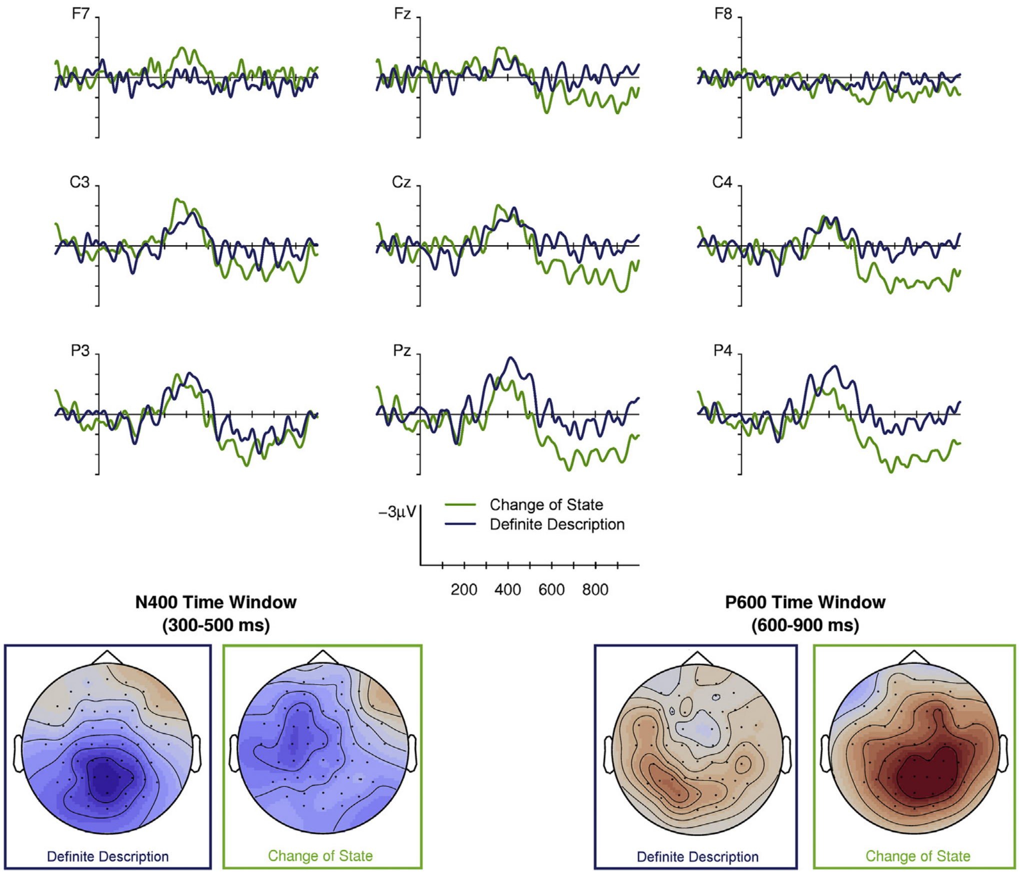


Fig. 5. Difference waves (ACC minus SAT) for each Trigger Type, and scalp topography of these differences during the N400 and P600 time windows.

The analysis of the ERPs during this time-window confirms that Condition induced reliable differences in the size of the N400 component and that N400 differences affect both levels of Trigger Type, with some differences in terms of scalp topography and reliability of the effect: CS sentences were associated with an N400 effect that was widely distributed over the scalp, whereas the N400 for DD passages was slightly more robust and showed the typical Central-Parietal distribution.

3.2.2. P600

Lateral and Midline ANOVAs carried out during the P600 time window confirmed that presupposition ACC elicited more positive ERPs than presupposition SAT (Table 5). This was attested by the significant main effect of Condition [Midline: $F(1,22) = 16.72$, $p < 0.001$; Lateral $F(1,22) = 20.54$, $p < 0.001$] and by the significant two-way interactions between Condition and Longitude [Midline: $F(6,132) = 6.69$, $p < 0.01$; Lateral: $F(2,44) = 14.20$, $p < 0.001$]. The differences between ACC and SAT were positive and broadly distributed, but with a maximum over Parietal [+1.18 μV, $F(1,22) = 33.20$, $p < 0.001$] and Central [+0.95 μV, $F(1,22) = 23.21$, $p < 0.001$] rather than Frontal [+0.43 μV, $F(1,22) = 3.77$, $p = 0.065$] electrodes. The effect of Condition was marginally different across types of Trigger. The Trigger Type by Condition interactions were marginally significant in both analyses [Midline: $F(1,22) = 4.06$, $p = 0.056$; Lateral: $F(1,22) = 3.36$, $p = 0.08$]. Considering all electrodes from both Midline and Lateral arrangements, the effect of Condition is prominent for CS [+1.34 μV, $F(1,22) = 15.47$, $p < 0.001$]

Table 5
Lateral and Midline ANOVAs on the P600 amplitude (600–900).

	df	F	p	ε	Eta
Lateral					
Trigger Type	1, 22	0.01	ns		
Condition	1, 22	20.54 ***	<0.001		0.067
Trigger Type X Condition	1, 22	3.36 ^	0.080		0.021
Trigger Type X Longitude	2, 44	14.09 ***	<0.001	0.65	0.029
Trigger Type X Hemisphere	1, 22	<1	ns		
Trigger Type X Longitude X Hemisphere	2, 44	1.40	0.25	0.75	0.000
Condition X Longitude	2, 44	14.20 ***	0.001	0.64	0.008
Condition X Hemisphere	1, 22	1.30	0.226		0.000
Condition X Longitude X Hemisphere	2, 44	2.66 ^	0.09	0.82	0.000
Trigger Type X Condition X Longitude	2, 44	<1	ns		
Trigger Type X Condition X Hemisphere	1, 22	8.08 **	0.009		0.003
Trigger Type X Condition X Longitude X Hemisphere	2, 44	<1	ns		
Midline					
Trigger Type	1, 22	1.06	0.31		0.006
Condition	1, 22	16.72	<0.001		0.050
Trigger Type X Condition	1, 22	4.06	0.056		0.021
Trigger Type X Longitude	6, 132	11.18	<0.001	0.36	0.050
Condition X Longitude	6, 132	6.97	0.001	0.39	0.012
Trigger Type X Condition X Longitude	6, 132	1.92	0.159	0.32	0.005

and non reliable for DD [+0.84 μV, $F(1,22) = 1.58$, ns], although the direction of the effect (a positive difference for ACC compared with SAT) is the same for both Trigger Types. Moreover, strong differences due to Trigger Type emerged in this time window, as attested by the Trigger Type by Longitude interactions [Midline: $F(6,132) = 11.18$, $p < 0.01$; Lateral: $F(2,44) = 14.09$, $p < 0.001$]. Over Frontal electrodes, CS elicited more negative ERPs than DD [-0.80 μV, $F(22) = 6.91$, $p = 0.015$], almost null differences in Central electrodes [+0.12 μV, $F < 1$], and more positive ERPs in Parietal [+0.66 μV, $F(22) = 4.85$, $p = 0.038$].

Critically, the three-way interaction between Trigger Type, Condition, and Hemisphere [$F(1,22) = 8.08$, $p < 0.01$] was followed up by carrying out two more two-way (Condition X Trigger Type) ANOVAs separately in Left and Right lateralized electrodes. This analysis further clarified that, only a main effect of Condition (and no interaction with Trigger Type) emerged over the Left Hemisphere [$F(1,22) = 13.56$, $p = 0.001$]. This suggests that differences between trigger types were less strong [CS: +1.06 μV; DD: +0.48 μV] in the Left hemisphere compared with the pattern over the Right hemisphere [CS: +1.59 μV; DD: +0.29 μV], where Condition [$F(1,22) = 23.12$, $p < 0.001$] was further modulated by the interaction with Trigger Type [$F(1,22) = 6.97$, $p = 0.014$]. We then assessed the size of the P600 for Condition (ACC vs SAT) considering only parietal electrodes in Left and Right Hemisphere. For CS, the effect was robust in both Left and Right electrodes, while for DD the effect could be detected on the Left Hemisphere only (Table 6; Fig. 5).

The analysis of the P600 component revealed that presupposition ACC is associated with larger P600 effects compared with SAT, and that differences between Trigger Types are relevant (Figs. 4 and 5): CS passages elicited a clear P600 effect due to Condition that was distributed in Central and Parietal electrodes, whereas in DD passages the P600 effect was detectable but tiny and Left lateralized (Fig. 5).

3.2.3. Additional analysis on the longer epoch for CS trigger type

The cluster permutation test on the longer ERP epoch (from the trigger verb to the processing point) revealed reliable differences between ACC and SAT[5] (ACC minus SAT). Two clusters of data represented the most consistent differences between the conditions and occurred in two partially overlapping time windows (Fig. 6). The earlier difference was associated with a negative cluster of data ($maxsum = -7374$, $p = 0.042$) in the latency ranging from 2084 ms to 2288 ms, in Parietal and Occipital electrodes. The other difference was described by a positive cluster of data ($maxsum = 15183$, $p = 0.004$) ranging from 2225 ms to 2800 ms, in Parietal electrodes. Interestingly, although the test was not constrained by any a priori hypothesis on the timing or scalp distribution of the effects, the latency range of the ERP differences overlapped to the latency of the time windows that were selected to investigate the N400 and P600 components on the last word of the phrase (284–488 ms vs 300–500 ms for the N400 and 452–1000 vs 600–900 ms for the P600). The non-parametric test substantially confirmed the N400-P600 pattern associated with the ACC condition for the CS Trigger Type on the processing point, and did not reveal any earlier additional effect.

However, taking a much more liberal approach that can be useful as exploratory analysis, the visual inspection of the longer ERP epoch (Fig. 6) suggests that an additional difference between conditions may actually exist: upon presentation of the Verb a larger P200 was associated with ACC compared with SAT. ANOVAs carried out on a time window of 80 ms centered around the peak of the P200 (from 200 to 280 ms) suggests that the effect is significant. A marginally significant effect of Condition [$F(1,22) = 2.98$, $p < 0.1$] and a significant Condition by Hemisphere interaction [$F(1,22) = 9.48$, $p < 0.01$] emerged

[5] We thank an anonymous reviewer for having suggested this analysis.

Table 6
Planned contrasts on the P600 amplitude on parietal electrodes. Significant level of alpha = 0.0125.

		ACC vs SAT in DD set	ACC vs SAT in CS set
Left Parietal Electrodes	Effect size in μV	0.85	1.41
	t(22) value	+2.58 ^	+2.97 *
	95% confidence intervals	0.16: 1.53	0.43: 2.39
Right Parietal Electrodes	Effect size in μV	0.50	1.86
	t(22) value	+1.59	+4.75 ***
	95% confidence intervals	−0.15: 1.15	1.05: 2.67

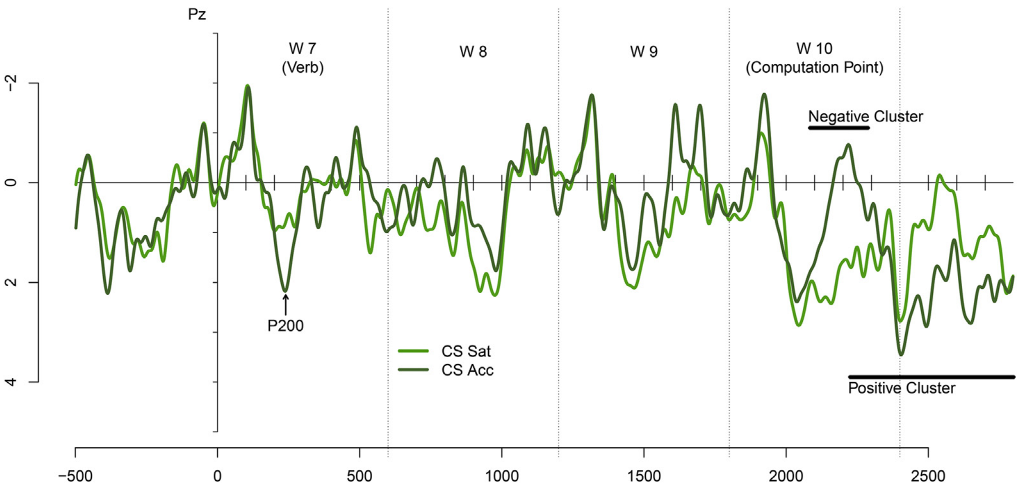


Fig. 6. ERP Grand averages of the longer epoch for CS Trigger Type from one representative electrode (Pz). Waves are shown from the presentation of word 7 (W 7, the CS verb) to the presentation of word 10 (W 10, the computation point). The two black horizontal lines represent the time windows of the Negative and Positive clusters that captured significant differences between ACC and SAT in the permutation test.

over lateral electrodes. This interaction shows that the effect is more prominent over the right [$F(1,22) = 6.50$, $p < 0.05$] rather than the left hemisphere [$F < 1$]. Furthermore, the main effect of Condition was significant for the channels distributed along the midline [$F(1,22) = 4.40$, $p < 0.05$]. Overall, the ANOVAs suggest that the P200 is larger for the ACC condition, and that the effect is distributed over midline and right lateralized electrodes.

4. Discussion

Results clearly revealed that the ERP correlates of accommodation observed on the “computation point” (i.e., the noun) of the two presupposition triggers consist of a biphasic ERP pattern, with a more negative N400 and a more positive P600 compared with presupposition SAT. Differences between types of triggers were also manifest in the three-way interactions observed in both time windows. With definite descriptions, we observed a more clear involvement of the N400 component, whereas change-of-state verbs elicited a more pronounced P600. The analysis of the responses to the behavioral task further showed that CS sentences were judged slightly more accurately and faster compared with DD sentences. The effect of Condition was not significant for judgment times, but the interaction between Trigger Type and Condition for accuracy suggests that the responses to satisfied CSs were more accurate than responses to satisfied DDs.

Concerning our initial expectations, the first research question at stake (RQ1) regarded the neurophysiological correlates of presupposition ACC, as compared with SAT, with definite descriptions. We observed that processing definite descriptions elicited a centro-parietal N400 followed by a P600 effect that could only be detected in left-lateralized posterior electrodes. The results are in line with previous experimental studies reporting a centro-parietal N400 response for “discourse-new” or “discourse-bridged” DDs (cf. Burkhardt, 2006; Masia et al., 2017). In these works, the N400 effect was described as reflecting additional costs related to a linking process, i.e., the attempt of anchoring a new discourse referent in a mental model where it was not previously introduced. In Burkhardt (2006) the larger N400 was found with respect to a given referent (repetition of the antecedent). In a similar vein, here the larger N400 for new referents DDs was also compared with a baseline condition in which the referent was already given, and the N400 effect was fully replicated. Burkhardt (2006) found also more positive ERPs for new referents compared with given referents, accounted for in terms of discourse updating. The positivity observed for DDs in the present study seems less robust than the one observed in Burkhardt (2006), although sharing the same, left-lateralized scalp distribution (similar distributions are not new in ERP research on this topic: see also Burkhardt, 2007; and Burkhardt & Roehm, 2007). The effect that we observed may be weaker (i.e., left-localized) because the DDs to be accommodated were completely plausible and related to the preceding context (as attested in the rating studies), whereas in Burkhardt (2006) the new referent was rather unrelated to the preceding context and this may have required more efforts in context updating mechanisms.

One may argue that the more negative N400 for new referents compared with given referents may be due to the well-known repetition effects (Schumacher, Bambini, & Weiland, 2012) attenuating the N400 in the SAT condition. Indeed, the presupposition satisfaction condition was carried out by repeating the antecedent and therefore it was impossible to completely avoid repetition priming with this manipulation. Nonetheless, and although we cannot quantify how much of the N400 difference was due to priming mechanisms, three observations suggest that the role of these mechanisms was not crucial: a) the N400 increases as function of the distance between repeated words; b) when discourse referents are repeated, other factors (e.g., prominence) may impact on the behavior of the N400; c) when repetition is avoided, genuine N400 effects have been reported for linking incomplete/non fully given content to the previous discourse. Concerning (a), experiments on repetition priming in word lists (Rugg, 1990) and texts (Van Petten, Kutas, Kluender, Mitchiner, & McIsaac, 1991) showed that with increasing lag between repeated words, the priming effect is reduced compared with shorter repetition lags. Van Petten et al. (1991) presented their participants with new and repeated words in natural texts. They subdivided the class of repeated words according to the number of intervening words since last occurrences (more or less than 20) and found that the N400 amplitude was larger at long lags compared with short lags. Given the long distance between repeated words in our experiment (between 20 and 25 words), repetition priming is likely to have exerted only limited influence on the modulation of the N400.

Concerning (b), research on the so-called repeated name penalty (e.g., Camblin, Ledoux, Boudewyn, Gordon, & Swaab, 2007; Swaab, Camblin, & Gordon, 2004) showed that other discourse factors, namely prominence, affect the N400 amplitude even when the target words are repetitions of the antecedent. Using written (Swaab et al., 2004) or auditory materials (Experiment 1 in Camblin et al., 2007) the authors had participants presented with short discourses. Prominence was manipulated by introducing either one or two antecedents, creating a prominent condition in which only one antecedent was available and thus highly prominent (*At the office Daniel moved the cabinet* …) and a non-prominent condition in which two antecedents were available (*At the office Daniel and Amanda moved the cabinet* …). In the final part of the passages, a name coreferential with the antecedent appeared (the repeated name condition: *because Daniel needed room for the desk*). Although the names were repeated within a very short distance, the N400 component was sensitive to how available the antecedent was, being it larger when referents were more prominent. This is a case in which the N400 is modulated by discourse factors that go beyond simple repetition priming. Similar evidence is reported also for common nouns (Cowles, Kluender, Kutas, & Polinsky, 2007). In our study too, thus, it seems plausible to think that the N400 in SAT was not exhausted by the repetition of the antecedent.

Finally, concerning (c), some experiments on information structure processing (Masia et al., 2017) offer further evidence to support the hypothesis that the N400 is sensitive to subtle differences in how information is packaged, also in the absence of repetition. For instance, Masia et al. (2017) showed that when new information is presented as presupposed (e.g, *the migration*) the N400 is larger than when it is conveyed as asserted (*there was a migration*). In that study, contexts were unaltered for the two conditions and so they rendered all presuppositions and assertions equally plausible and predictable, which suggests that the observed modulation in the N400 signature was a reflection of variations at the sole presupposition-assertion packaging level.

Taken together, these observations support the hypothesis that a significant part of the N400 difference observed in the present investigation is a genuine correlate of accommodation processes, reflecting increasing difficulty in linking new information to prior discourse. Note, moreover, that as concerns differences between the trigger types (see RQ3 below) repetition priming cannot be responsible as it equally affects both trigger types.

There is a further argument – outside the ERP literature - against the idea that our results were mainly due to repetition effects. Different behavioral experiments compared SAT vs ACC using self-paced reading time paradigms based on a method similar to the one adopted in our experiment, where the same target sentence was presented in two conditions, in which the information presupposed was either previously introduced (SAT) or not (ACC), e.g., Haviland and Clark (1974), Tiemann et al. (2011, 2014). In particular, Haviland and Clark (1974), in a first seminal experiment, compared SAT vs ACC with different trigger types and measured the reading times of the target sentences containing the triggers on the whole sentence region. Data collected revealed longer reading times in ACC. Haviland and Clark interpreted these data within the Given-New strategy theory, and attributed the longer reading times in ACC to the extra cost needed for the process of integrating the new information with the given information. In order to exclude that such an effect could be the result of a mere repetition of the information (i.e., priming), Haviland & Clark run a follow up study. In this second experiment, they observed that their indirect antecedent condition (i.e., our ACC) still elicited longer reading times than the direct antecedent condition (i.e., our SAT), even if in this case the information (potentially primed) was equally repeated in both conditions. For example, an item in the Direct Antecedent Condition was *We got some beer out of the trunk* (context sentence). *The beer was warm* (target sentence), while an item in the Indirect Antecedent Condition was *Andrew was especially fond of beer* (context sentence). *The beer was warm* (target sentence). Although we cannot exclude a potential role of repetition, we think that Haviland & Clark's results reliably suggest that the effect of the condition observed with the present manipulation reflects more the cognitive costs involved in presupposition processing than a priming effect.

The second question (RQ2) under exam concerned the ERP components associated with the accommodation of change-of-state verbs. For CSs, data showed an N400 effect followed by a strong modulation of the P600 component. The N400 difference associated with CSs accommodation compared with SAT was statistically reliable when considering all scalp locations, but when looking at planned comparisons in Central and Parietal electrodes, where the N400 is usually distributed, the observed differences resulted less reliable. The N400 effect may be too weak (i.e., non-canonically distributed) to unambiguously reflect the involvement of discourse linking processes because, following the same argument on the N400 effect observed with DDs, it is difficult to exclude that part of the N400 response for CSs is due to repetition priming. By contrast, the later positivity seems considerably robust. Consistently with the framework provided in Burkhardt (2007), the strong modulation of the P600 component for the accommodation of CSs may reflect higher cognitive costs in terms of discourse updating mechanisms.

These results provide new evidence for P600 effects in sentences that are syntactically well-formed, as for semantic reversal anomalies of the type *the hearty meal was devouring* (e.g., Kim & Osterhout, 2005). The novel contribution of the present work is in fact that CSs were perfectly well-formed also from a semantic point of view, as attested by the rating studies conducted on the experimental materials. Indeed, our data support the view of the P600 as a component responsible for integration processes, which operates at the discourse level (Hoeks & Brouwer, 2014) and, more generally, at the contextual level, via inferential mechanisms (Bambini et al., 2016; Spotorno et al., 2013).

Finally, concerning (RQ3), we observed interesting differences between the two types of triggers, attested by the several significant interactions between Trigger Type and Condition. In the N400 time window presupposition accommodation was associated with larger N400 components for both triggers, but the planned comparisons carried out in posterior electrodes, where the N400 is usually distributed, revealed a reliable effect for DDs and a marginally significant effect for CSs. This result is consistent with our prediction that a resolution trigger like DDs could elicit a more demanding linking process than a lexical trigger like CSs, because DDs involve a process of anaphoric retrieval of the antecedent presupposition, while CSs directly encode in their conventional meaning a precondition for their asserted content. Other differences emerged during the P600 time window, where presupposition accommodation was associated with more positive ERPs for both triggers. However, the P600 effect was prominent in CSs while in DDs it could only be detected over left-lateralized electrodes. This result seems to support our expectation that, in condition of accommodation, CSs constitute a more demanding category of presupposition triggers as compared with DDs in terms of updating mechanisms, by which the discourse model is integrated or restructured with the content of the new presupposition. As predicted by the behavioral literature, CSs instantiate a category of presupposition trigger involving more complex and demanding processes of presupposition accommodation as compared with DDs, because they require the more cognitively demanding process of construing temporally displaced events (Domaneschi et al., 2014; Tiemann, 2014).

A limitation of this analysis is related to the structural differences of the materials preceding the computation point in the two conditions (the noun in the 10th position). With CSs, accommodation could start before the noun and precisely at the verb following the trigger – i.e., "stopped SMOKING cigarettes …". This issue was explored in the additional analysis on the longer epoch spanning from the verb to the noun. Results confirmed the biphasic N400-P600 pattern at the computation point, which was the most robust ERP response visible in whole time course. However, visual inspection revealed a larger P200 for ACC at the point where the verb was presented, which proved significant in an ad-hoc analysis. The latency of the effect is consistent with the previous literature assessing the P200 component in language processing (e.g., 180–260 ms in Carreiras, Vergara, & Barber, 2005), although the scalp topography is less typical, being right lateralized. Because of the ad-hoc analysis and the right lateralization, this result should be taken with caution, but it offers nice preliminary evidence suggesting that the time course of accommodation in CSs might indeed be already affected upon reading the verb. In language studies, ERP effects on the P200 has been reported for a variety of factors, including word level processing, e.g., phonological and orthographic recognition (Barnea & Breznitz, 1998; Carreiras et al., 2005), as well as at a higher level, e.g., for prosodic integration (Liu, Wang, & Jin, 2010) and processing of irony (Regel et al., 2011). Across studies, authors seem to converge on the link between the P200 and the recruitment of attentional resources (Lee, Liu, & Tsai, 2012). Future studies are needed to replicate the P200 result and consider its role within the previous literature.

The above differences allow us to sketch considerations on the time-course of accommodation across trigger types. While accommodating a presupposition compared with the simple satisfaction of it seems to unfold through a biphasic pattern with both triggers, the efforts might be differently distributed over time. DDs seems to capitalize on earlier mechanisms (of discourse linking) while CSs seem to be engage more cognitive resources in the later stage of updating the mental model of discourse, in the integration of temporally displaced events. Moreover, the topographic distribution of the two effects for the two trigger types may suggest qualitatively different linking and updating mechanisms in the accommodation of DDs and CSs. In terms of topography and effect size, the N400 was more prominent in DDs and the P600 was more prominent in CSs. Although the ERP technique cannot be used to localize the neural sources of the cognitive processes under scrutiny, the different scalp distribution of the N400 and P600 effects across trigger types is potentially indicative of non-overlapping neural generators and consequently different processing mechanisms. For both N400 effects we could think of a summation of linking and repetition priming effects. Being the contribution of these two factors different across trigger types (more linking effort in DD), the distribution on the scalp resulted different as well. Concerning the P600, the left-lateralized P600 for ACC in DDs replicates previous findings and may be related to update mechanisms required by the inclusion of a new discourse referent (as interpreted in Burkhardt, 2007), whereas the more widely distributed effect for CS may be related to more demanding update mechanisms not only due to the inclusion of a new referent but also to the representation of temporally displaced events. Clarifying further the extent to which DDs and CSs differ in terms of linking and updating mechanisms is difficult with one single study available, and the scant literature on presupposition does not support any other speculation. For example, it might not be excluded that the P200 observed on the triggering point of CSs might signal that the linking process starts on the critical region of the verb. However, we take the result of this early effect as an explorativestep towards a more fine-grained comparison of the ERP components involved in presupposition accommodation with different trigger types.

A closer inspection to the behavioral data is also of some interest. In particular, we observed a trigger effect with shorter response times in CS, and a Condition × Trigger interaction with higher accuracy for CSs overall and, in particular, in SAT. Interestingly, results in response times replicate the findings of Domaneschi and Di Paola (2017) where faster response was obtained for CSs compared with definite descriptions, focus particles and factive verbs. In this experiment, we have also observed that the presuppositions of CSs were accommodated accurately more frequently and, once introduced in the discourse mental model, were more easily retrieved than DDs (i.e. higher accuracy with CSs than with DDs). These data are only apparently in contrast with the ERP results and with the interpretation of the CSs as more demanding in the updating stage. It is worth recalling that, according to Zeevat (1992), with resolution triggers like DDs, the triggering process consists in the retrieval of an entity from the preceding context. Failure in such retrieving process results in a discourse failure rather than in a *logical* failure. Conversely, with lexical triggers like CSs, presupposition failure yields a logical failure since the presupposition constitutes a logical precondition for the (temporal) implication conveyed by the trigger. The fact that CSs were judged more accurately than DDs is indicative of a substantial difference between lexical triggers and resolution triggers in the speakers' availability to accept and accommodate the backgrounded content. As shown in other behavioral studies, the presuppositions of lexical triggers like CSs are less likely to be dismissed compared with the presuppositions of resolution triggers (cf. Amaral & Cummins, 2015; Cummins, Amaral, & Katsos, 2012). Furthermore, the fact that the presuppositions of CSs were recovered more easily and rapidly might indicate that they are derived via a direct logical implication. On the contrary, the presuppositions of DDs are inferred via a discourse-based anaphoric inference, which makes the recovering process more indirect and demanding. Roughly speaking, it is plausible to think that if someone remembers that "Sue has given up smoking", it is easy and logical to remember that "She used to smoke". Conversely, if someone remembers that "The black cat crossed the street" the information that a unique black cat exists has to be recalled from prior discourse. Although the experimental paradigm used in the present work does not leave too much space for speculations and post-hoc explanations about the behavioral measures, we believe that the more direct/indirect retrieval of the presupposition associated with anaphoric and lexical triggers offers a preliminary explanation to the higher accuracy observed with CSs as compared to DDs.

5. Conclusion

Using the event-related brain potential technique, this study aimed at taking a first step towards a neurolinguistic investigation of brain response to the accommodation of presuppositions as compared with presupposition satisfaction with two types of presupposition triggers: definite descriptions and change of state verbs. Data showed that presupposition accommodation is associated with a biphasic time-course that shows up with different modulations for the two trigger types. In the N400 component, the accommodation of DDs seems to elicit slightly stronger (i.e., more canonically distributed) response than CSs. An even more remarkable difference between the two triggers is reflected in the P600 response, which is more prominent (i.e., more extended and canonically distributed) for CSs than for DDs in ACC condition.

This work supports two main conclusions. First, independently of the trigger in use, presupposition accommodation involves a N400 + P600 response associated with linking and updating mechanisms reflecting i) a process of research of a previous antecedent in the prior discourse and ii) a subsequent process of context repair where the content of the presupposition is integrated in the discourse model. Second, variance in the cognitive load of different presupposition triggers can be observed along the time course of the ERP response.

The present study adds some evidence in support of a better understanding of the cognitive processes involved in presupposition processing in on-line language comprehension. Understanding the information taken for granted seems to involve (at least) two different mechanisms of antecedent retrieval and of discourse mental model update. Such a result, moreover, contributes to a more clear characterization of two key ERP components involved in discourse level processing: N400 and P600. P600 activity, in particular, seems to be associated with mechanisms involved in the discourse mental model construction.

Authors' contribution

Study concept and design: FD, VB, PC, VM, ELV. Data collection: VM, PC. Data analysis: PC. Manuscript writing: FD, VM, PC, ELV, VB.

Acknowledgments

The authors would like to thank Gianmarco Giordano, Irene Ricci, and Chiara Bertini for their support in data collection. This research has been funded by the Italian Ministry of Education, University and Research within the three-year project SIR_2014 - *EXPRESS – Experimenting on Presuppositions* directed by FD, project code RBSI147WM0. VB was supported by the

MIUR PRIN (Progetti di Ricerca di Rilevante Interesse Nazionale) 2015, project "The Interpretative Brain: Understanding and Promoting Pragmatic Abilities across Lifespan and in Mental Illness," project code 201577HA9M.

Appendixes

Appendix 1

Representative sample of stimulus pairs for definite description (DD) trigger in the satisfaction (SAT) and accommodation (ACC) conditions (original Italian; English translation in brackets). Antecedent in the context sentence 1 for the SAT condition in italics; target word in the target sentence in underlined bold.

Trigger Type	Condition	Context sentence 1	Context sentence 2	Target sentence
DD-1	SAT	Nell'ufficio di Paolo lavorava un *grafico* davvero scontroso. [Engl. In Paolo's office, there used to be a very bad-tempered *designer*.]	L'ufficio necessita di consulenti per diversi rami dell'attività. [Eng. The office needs consultants for several branches of the activity.]	Per esubero del personale, circa un mese fa il **grafico** è stato messo in mobilità. [Eng. Due to overstaffing problems, about a month ago, the graphic **designer** was made redundant.]
	ACC	Nell'ufficio di Paolo lavorano diversi impiegati. [Eng. In Paolo's office there are many employees.]		
DD-2	SAT	Vicino a casa di Manuela c'è un *cinema* aperto tutto il giorno. [Eng. Close to Manuela's house, there is a *cinema* open 24-h a day.]	La zona è molto organizzata e tutto sommato vivibile. [Eng. The area is well organized and liveable.]	Con forti reazioni da parte degli abitanti, ieri il **cinema** è stato chiuso per ristrutturazione. [Eng. With strong reactions on the part of the locales, yesterday the **cinema** was closed.]
	ACC	Il quartiere di Manuela offre molti servizi. [Eng. The quarter where Manuela lives offers many services.]		
DD-3	SAT	Per la sua nuova camera, Marco ha comprato un *baule*. [Eng. Marco has bought a *trunk* for his new bedroom.]	I suoi genitori lo hanno aiutato nelle spese di arredamento. [Eng. His parents helped him paying the furniture.]	I nuovi mobili sono grandi e capienti, e il **baule** è in puro legno massello. [Eng. The new furniture is big and capacious, and the **trunk** is made of solid wood.]
	ACC	Marco ha da poco affittato una camera nuova. [Eng. Marco has recently rented a new bedroom.]		
DD-4	SAT	Nel liceo in cui insegna Chiara c'è una *piscina*. [Eng. In the school where Chiara teaches, there is a *pool*.]	Gli studenti della scuola praticano alcuni sport. [Eng. The students play several sports.]	Per via di numerose gare e campionati regionali, la **piscina** è spesso affollata e rumorosa. [Eng. Due to the many competitions and regional championships, the **pool** is often crowded and noisy.]
	ACC	Chiara insegna inglese in un liceo scientifico. [Eng. Chiara teaches English in a scientific high school.]		
DD-5	SAT	Nello zoo di Trieste c'è un *ghepardo* dal manto bellissimo. [Eng. In Trieste zoo, there is a *cheetah* with a beautiful fur.]	Molti animali sono nati in cattività. [Eng. Many animals were born in captivity.]	Dopo un lungo periodo di attesa, l'altro ieri il **ghepardo** ha partorito tre bellissimi cuccioli. [Eng. After a long period, the day before yesterday, the **cheetah** gave birth to three beautiful cubs.]
	ACC	A Trieste c'è uno zoo aperto tutta la settimana. [Eng. In Trieste, there a zoo open all week.]		
DD-6	SAT	Matteo vive in un appartamento con una *coinquilina* ungherese. [Eng. Matteo lives in a flat with a Hungarian *roommate*.]	La casa gli è stata regalata da sua nonna due anni fa. [Eng. He got the house from his granny two years ago.]	Di solito lui studia in camera sua, invece la **coinquilina** lavora molto spesso in cucina. [Eng. He usually studies in his room, while the **roommate** often works in the kitchen.]
	ACC	Matteo si è trasferito in un appartamento vicino alla sua università. [Eng. Matteo moved in a flat near the campus.]		
DD-7	SAT	Giulio ha una *figlia* con cui va spesso al cinema. [Eng. Giulio has a *daughter* and he often goes to the cinema with her.]	Ieri sono andati a vedere un film dell'orrore. [Eng. Yesterday, they saw a horror movie.]	Lui ha impiegato venti minuti per parcheggiare mentre la **figlia** aspettava impaziente davanti alla biglietteria. [Eng. It took him twenty minutes to park the car, while his **daughter** was waiting in front the ticket office.]
	ACC	Giulio va spesso al cinema con i suoi familiari. [Eng. Giulio often goes to the cinema with his family.]		

(*continued on next page*)

(*continued*)

Trigger Type	Condition	Context sentence 1	Context sentence 2	Target sentence
DD-8	SAT	Nel palazzo di Marco c'è un *solaio*. [Eng. In Marco's building there is an *attic*.]	L'edificio ha sei piani ed è piuttosto vecchio. [Eng. The building has six floors and is rather old.]	Lo scorso anno, durante la ristrutturazione del palazzo, il **<u>solaio</u>** è stato tinteggiato di grigio. [Eng. Last year, during renovation works, the **<u>attic</u>** was painted grey.]
	ACC	Marco vive in un palazzo antico con molti appartamenti. [Eng. Marco vive in an old building with many flats.]		
DD-9	SAT	Ad Altamira c'è un sito archeologico in cui fino a qualche tempo fa lavorava un *custode*. [Eng. In Altamira, there is an archeological site in which a *guardian* used to work.]	Purtroppo, da tempo non si scoprono più reperti interessanti. [Eng. Unfortunately, it's been a long time since the last finds were discovered.]	Alcuni settori del sito sono stati chiusi, così il **<u>custode</u>** è stato mandato a casa. [Eng. Some sections of the site have been closed, so the **<u>guardian</u>** has been dismissed.]
	ACC	Ad Altamira c'è un sito archeologico che riceve numerosi visitatori ogni anno. [Eng. In Altamira, there is an archeological site visited by many tourists every year.]		
DD-10	SAT	Nel nuovo quartiere di Carla c'è un *tram* che passa ogni 15 minuti. [Eng. In Carla's quarter, there is a *tram* running every 15 min]	Lei non ha un'automobile e si sposta solo con i mezzi pubblici. [Eng. She does not have a car and so she always uses public transport.]	La zona è in generale ben collegata e il **<u>tram</u>** ferma proprio sotto casa sua. [Eng. On the whole, the area is well-connected and the **<u>tram</u>** stops right in front of her house.]
	ACC	Carla si è trasferita in un piccolo quartiere di Caserta. [Eng. Carla moved in a little quarter in Caserta.]		

Appendix 2

Representative sample of stimulus pairs for change-of-state verbs (CS) trigger in the satisfaction (SAT) and accommodation (ACC) conditions (original Italian; English translation in brackets). Antecedent in the context sentence 1 for the SAT condition in italics; target word in the target sentence in underlined bold.

Trigger Type	Condition	Context sentence 1	Context sentence 2	Target sentence
CS-1	SAT	Matteo comprava spesso *sigari* costosi e dal sapore forte. [Eng. Matteo often bought expensive and strong-flavoured *cigars*.]	Negli ultimi mesi non è stato molto bene. [Eng. Over the past few months, he has not been very well.]	Così, da circa due settimane ha smesso di comprare **<u>sigari</u>** e fuma solo sigarette elettroniche. [Eng. So, two weeks ago, he stopped buying **<u>cigars</u>** and now only smokes electronic cigarettes.]
	ACC	Matteo non è mai stato attento alla sua salute. [Eng. Matteo has never looked after his health.]		
CS-2	SAT	Sara ha sempre fatto *giardinaggio* durante il fine settimana. [Eng. Sara has always practised *gardening* during the weekends.]	Da quando ha un nuovo impiego, però, ha poco tempo libero. [Eng. Now she has a new job and so she does not have much free time.]	Da quasi tre mesi ormai ha smesso di fare **<u>giardinaggio</u>** poiché lavora spesso il sabato. [Eng. Three months ago, she stopped practising **<u>gardening</u>** as she often works on Saturdays.]
	ACC	Da giovane, Sara ha sempre avuto molti hobby interessanti. [Eng. As a young woman, Sara used to have many interesting hobbies.]		
CS-3	SAT	Il parroco della diocesi di Grosseto organizzava *concerti* per appassionati di musica gospel. [Eng. The priest of Grosseto's church used to organize *concerts* for classical music lovers.]	Purtroppo, però, di recente i fondi a disposizione sono stati ridotti drasticamente. [Eng. Unfortunately, though, the funds available have been drastically reduced over the last years.]	Per questo motivo il parroco ha smesso di organizzare **<u>concerti</u>** e ora promuove alcune conferenze. [Eng. For this reason, the priest has stopped organizing **<u>concerts</u>** and now promotes conferences.]
	ACC	Il parroco della diocesi di Grosseto è molto attivo nella comunità. [Eng. The priest of Grosseto's church is extremely dynamic in the community.]		

(*continued*)

Trigger Type	Condition	Context sentence 1	Context sentence 2	Target sentence
CS-4	SAT	Da giovane, Carlo vendeva *quadri* nel negozio di suo zio. [Eng. As a young man, Carlo used to sell *paintings* in his uncle's shop.]	Lavorava molte ore al giorno e a volte tornava a casa alle dieci di sera. [Eng. He used to work many hours a day and sometimes went back home at 10 p.m.]	Su richiesta di sua moglie, ha smesso di vendere **quadri** e ha avviato un'attività commerciale. [Eng. To make his wife happy, he stopped selling **paintings** and started a business of his own.]
	ACC	Da giovane, Carlo lavorava nel negozio di suo zio. [Eng. As a young man, Carlo used to work in his uncle's shop.]		
CS-5	SAT	Da circa sei anni, Lia vendeva *scarpe* in un negozio di Perugia. [Eng. Until six years ago, Lia had been selling *shoes* in a shop in Perugia.]	Da tempo desiderava un lavoro più creativo e stimolante. [Eng. For a while already, she desired a more creating and challenging job.]	Dopo un periodo di riflessione, ha smesso di vendere **scarpe** e ha aperto una pasticceria. [Eng. After a period of reflection, she stopped selling **shoes** in a shop and opened a patisserie.]
	ACC	Lia ha sempre lavorato con grande diligenza e dedizione. [Eng. Lia has always worked with remarkable zeal and dedication.]		
CS-6	SAT	Alle medie, Rebecca prendeva *ripetizioni* di matematica. [Eng. At the junior high school, Rebecca used to take private math *lessons*.]	Non le piace lo studio, e a scuola consegue voti negativi. [Eng. She does not like to study and she gets low marks.]	Poi, all'inizio del nuovo quadrimestre ha smesso di prendere **ripetizioni** perché preferisce studiare da sola. [Eng. Then, at the beginning of the new term, she stopped taking **lessons** because she prefers to study by herself.]
	ACC	Quasi tutti i pomeriggi, Rebecca usciva con le sue amiche. [Eng. Rebecca used to go out with her friends almost all afternoons.]		
CS-7	SAT	I professori dell'Università di Trieste hanno sempre organizzato *tirocini* per gli studenti di Ingegneria. [Eng. Professors at the University of Trieste have always organized *training courses* for Engineering students.]	Purtroppo, il numero di iscritti è calato di molto in questi anni. [Eng. Unfortunately, the number of enrolled students has remarkably shrunk in the recent years.]	Per questo motivo, i professori hanno smesso di organizzare **tirocini** e tengono solo corsi specialistici. [Eng. For this reason, professors have stopped organizing **training courses** and only hold endorsement courses.]
	ACC	All'università di Trieste l'offerta didattica è molto ampia per gli studenti di Ingegneria. [Eng. At the University of Trieste, education programs are very rich for Engineering students.]		
CS-8	SAT	Il nonno di Nicola faceva spesso *artigianato* nella sua casa di campagna. [Eng. Nicola's grandfather used to practise *craft* in his house in the countryside.]	Sia in estate che in inverno si dedicava a molti passatempi. [Eng. In summer and in winter he spent much time on his hobbies.]	Ultimamente soffre di artrosi, così ha smesso di fare **artigianato** poiché ha frequenti dolori articolari. [Eng. Recently he has suffered from arthrosis, so he stopped practising **craft** because his joints hurt.]
	ACC	Il nonno di Nicola è sempre stato una persona attiva. [Eng. Nicola's grandfather has always been a very dynamic person.]		
CS-9	SAT	Fino allo scorso anno, Marco e Chiara compravano spesso *verdure* scadenti al supermercato. [Eng. Until last year, Marco and Chiara often bought low-quality *vegetables* at the supermarket.]	Ora hanno iniziato a curare un orto. [Eng. Now, they have begun cultivating their own products.]	Da un paio di settimane hanno smesso di comprare **verdure** e consumano i loro prodotti. [Eng. A couple of weeks ago, they stopped buying **vegetables** and now eat their own ones.]
	ACC	Marco e Chiara si sono trasferiti in un agriturismo. [Eng. Marco and Chiara have moved into a farmhouse.]		
CS-10	SAT	Al liceo, i nipoti di Franca passavano tutto il pomeriggio ad ascoltare *musica*. [Eng. At the high school, Franca's grandchildren used to spend all afternoons listening to *music*.]	Quando non avevano compiti si dedicavano al loro passatempo preferito. [Eng. When they did not have homework, they spent their time on their favourite pastime.]	Ora hanno iniziato l'università, così hanno smesso di ascoltare **musica** perché hanno molto da studiare. [Eng. Now they go to university, so they stopped listening to **music** because they have a lot to study.]
	ACC	Al liceo, i nipoti di Franca erano sempre in giro a divertirsi. [Eng. At the high school, Franca's grandchildren were always around having fun.]		

References

Amaral, P., & Cummins, C. (2015). A cross-linguistic study on information backgrounding and presupposition projection. In F. Schwarz (Ed.), *Experimental perspectives on presuppositions* (pp. 157–172). Dordrecht: Springer. http://dx.doi.org/10.1007/978-3-319-07980-6_7.

Arnold, J. E., Wasow, T., & Losongco, A. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language, 76*, 28–55. http://dx.doi.org/10.1353/lan.2000.0045.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. http://dx.doi.org/10.1016/j.jml.2007.12.005.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3*(2), 12–28. http://dx.doi.org/10.21500/20112084.807.

Bambini, V., Bertini, C., Schaeken, W., Stella, A., & Di Russo, F. (2016). Disentangling metaphor from context: An ERP study. *Frontiers in Psychology, 7*, 559. http://dx.doi.org/10.3389/fpsyg.2016.00559.

Bambini, V., & Resta, D. (2012). Metaphor and experimental pragmatics: When theory meets empirical investigation. *Humana. Mente Journal of Philosophical Studies, 23*, 37–60.

Bambini, V., Resta, D., & Grimaldi, M. (2014). A dataset of metaphors from the Italian literature: Exploring psycholinguistic variables and the role of context. *PLoS One, 9*. https://doi.org/10.1371/journal.pone.0105634.

Barnea, A., & Breznitz, Z. (1998). Phonological and orthographic processing of Hebrew words: Electrophysiological aspects. *Journal of Genetic Psychology, 159*, 492–504.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. http://dx.doi.org/10.1016/j.jml.2012.11.001.

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

Beaver, D., & Zeevat, H. (2007). Accommodation. In G. Ramchand, & C. Reiss (Eds.), *The oxford handbook of linguistic interfaces* (pp. 503–538). Oxford: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780199247455.0010001.

Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language, 98*(2), 159–168. http://dx.doi.org/10.1016/j.bandl.2006.04.005.

Burkhardt, P. (2007). The P600 reflects cost of new information in discourse memory. *Neuroreport, 18*(17), 1851–1854. http://dx.doi.org/10.1097/WNR.0b013e3282f1a999.

Burkhardt, P., & Roehm, D. (2007). Differential effects of saliency: An event-related brain potential study. *Neuroscience Letters, 413*, 115–120. http://dx.doi.org/10.1016/j.neulet.2006.11.038.

Camblin, C. C., Ledoux, K., Boudewyn, M., Gordon, P. C., & Swaab, T. Y. (2007). Processing new and repeated names: Effects of coreference on repetition priming with speech and fast RSVP. *Brain Research, 1146*, 172–184. http://dx.doi.org/10.1016/j.brainres.2006.07.033.

Canal, P., Pesciarelli, F., Vespignani, F., Molinaro, N., & Cacciari, C. (2017). Basic composition and enriched integration in idiom processing: An EEG study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(6), 928–943. http://dx.doi.org/10.1037/xlm0000351.

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

Cowles, H. W., Kluender, R., Kutas, M., & Polinsky, M. (2007). Violations of information structure: An electrophysiological study of answers to wh-questions. *Brain and Language, 102*, 228–242. http://dx.doi.org/10.1016/j.bandl.2007.04.004.

Cummins, C., Amaral, P., & Katsos, N. (2012). Experimental investigations of the typology of presupposition triggers. *Humana Mente, 23*, 1–16.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*, 9–21. http://dx.doi.org/10.1016/j.jneumeth.2003.10.009.

Domaneschi, F., Carrea, E., Penco, C., & Greco, A. (2014). The cognitive load of presupposition triggers: Mandatory and optional repairs in presupposition failure. *Language, Cognition and Neuroscience, 29*(1), 136–146. http://dx.doi.org/10.1080/01690965.2013.830185.

Domaneschi, F., & Di Paola, S.., (2017). The Cost of Context Repair: Presupposition Accommodation, CEUR-Proceedings of the Workshop on Contexts in Philosophy – Paris, June, 20, 2017, Vol. 1845, urn:nbn:de:0074-1845-7, 36-47.

Donnellan, K. S. (1966). Reference and definite descriptions. *Philosophical Review, 75*(3), 281–304.

Frazier, L. (2006). *The big fish in a small pond: Accommodation and the processing of novel definites*. Amherst: University of Massachusetts. Unpublished manuscript.

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Psychological Reviews, 91*, 1357–1392. http://dx.doi.org/10.1152/physrev.00006.2011.

Glanzberg, M. (2003). *Felicity and presupposition triggers, University of Michigan workshop in philosophy and linguistics, Michigan, USA*.

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48*(12), 1711–1725. http://dx.doi.org/10.1111/j.1469-8986.2011.01273.x.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304*, 438–441. http://dx.doi.org/10.1126/science.1095455.

Hagoort, P., & Van Berkum, J. J. A. (2007). Beyond the sentence given. Philosophical transactions of the royal society. Series B. *Biological Sciences, 362*, 801–811. http://dx.doi.org/10.1098/rstb.2007.2089.

Haviland, S. E., & Clark, H. H. (1974). What's New? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior, 13*, 512–521.

Hawkins, J. A. (1978). *Definiteness and indefiniteness. A study in reference and grammaticality prediction*. London: Croom Helm.

Heim, I. R. (1982). *The semantics of definite and indefinite noun phrases*. Amherst: University of Massachussetts (Doctoral dissertation).

Hoeks, J. C. J., & Brouwer, H. (2014). Electrophysiological research on conversation and discourse processing. In T. Holtgraves (Ed.), *Oxford handbook of language and social psychology* (pp. 365–386). Oxford: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780199838639.013.024.

Jaeger, T. F. (2008). Categorical data Analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434–446. http://dx.doi.org/10.1016/j.jml.2007.11.007.

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. http://dx.doi.org/10.1080/23273798.2016.1209531.

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

Karttunen, L. (1974). Presupposition and linguistic context. *Theoretical Linguistics, 1*, 181–194. http://dx.doi.org/10.1515/thli.1974.1.1-3.181.

Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E. S., Luck, S. J., et al. (2014). Committee report: Publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology, 51*(1), 1–21. http://dx.doi.org/10.1111/psyp.12147.

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52*, 205–225. http://dx.doi.org/10.1016/j.jml.2004.10.002.

Kiparsky, C., & Kiparsky, P. (1971). Fact. In D. D. Steinberg, & L. A. Jakobovitz (Eds.), *Semantics: An interdisciplinary reading* (pp. 345–369). Cambridge: Cambridge University Press.

Kirsten, M., Tiemann, S., Seibold, V. C., Hertrich, I., Beck, S., & Rolke, B. (2014). When the polar bear encounters many polar bears: Event-related potential context effects evoked by uniqueness failure. *Language, Cognition and Neuroscience, 29*, 1147–1162. http://dx.doi.org/10.1080/23273798.2014.899378.

Kripke, S. (2009). Presupposition and Anaphora: Remarks on the formulation of the projection problem. *Linguistic Inquiry, 40*(3), 367–386. http://dx.doi.org/10.1162/ling.2009.40.3.367.

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

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

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

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience, 9*, 920–933. http://dx.doi.org/10.1038/nrn2532.

Lee, C.-Y., Liu, Y.-N., & Tsai, J.-L. (2012). The time course of contextual effects on visual word recognition. *Frontiers in Psychology, 3*(285), 285. http://dx.doi.org/10.3389/fpsyg.2012.00285.

Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic, 8*, 339–359. http://dx.doi.org/10.1007/BF00258436.

Liu, B., Wang, Z., & Jin, Z. (2010). The effects of punctuations in Chinese sentence comprehension: An ERP study. *Journal of Neurolinguistics, 23*, 66–80. https://doi.org/10.1016/j.jneuroling.2009.08.004.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology, 54*(1), 146–157. http://dx.doi.org/10.1111/psyp.12639.

Luck, S. J., & Kappenman, E. S. (2012). *The Oxford handbook of event-related potential components*. Oxford: Oxford University Press.

Maris, E., & Oostenveld, R. (2007). Non-parametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190. http://dx.doi.org/10.1016/j.jneumeth.2007.03.024.

Masia, V., Canal, P., Ricci, I., Lombardi Vallauri, E., & Bambini, V. (2017). Presupposition of new information as a pragmatic garden path: Evidence from event-related brain potentials. *Journal of Neurolinguistics, 42*, 31–48. http://dx.doi.org/10.1016/j.jneuroling.2016.11.005.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2015). *Balancing type I error and power in linear mixed models*. preprint available at: arXiv:1511.01864.

Mennes, M., Wouters, H., Vanrumste, B., Lagae, L., & Stiers, P. (2010). Validation of ICA as a tool to remove eye movement artifacts from EEG/ERP. *Psychophysiology, 47*, 1142–1150. http://dx.doi.org/10.1111/j.1469-8986.2010.01015.x.

Nieuwland, M. S., & Van Berkum, J. J. A. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience, 18*(7), 1098–1111. http://dx.doi.org/10.1162/jocn.2006.18.7.1098.

Noveck, I., & Reboul, A. (2008). Experimental pragmatics: A gricean turn in the study of language. *Trends in Cognitive Sciences, 12*, 425–431. http://dx.doi.org/10.1016/j.tics.2008.07.009.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*, 97–113. http://dx.doi.org/10.1016/0028-3932(71)90067-4.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2010). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Article ID 156869 *Computational Intelligence and Neuroscience*, 9. http://dx.doi.org/10.1155/2011/156869.

Piemontese, E. (1996). *Capire e farsi capire. Teorie e tecniche della scrittura controllata*. Napoli: Tecnodid.

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. http://dx.doi.org/10.1162/jocn.2010.21411.

Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory & Cognition, 18*, 367–379.

Rugg, M. D., & Coles, M. G. H. (1995). *Electrophysiology of mind: Event-related brain potentials and cognition*. Oxford: Oxford University Press.

Sassenhagen, J., & Bornkessel-Schlesewsky, I. (2015). The P600 as a correlate of ventral attention network reorientation. *Cortex, 66*, 3–20. http://dx.doi.org/10.1080/23273798.2016.1209531.

Sauerland, U., & Schumacher, P. B. (2016). Pragmatics: Theory and experiment growing together. *Linguistische Berichte, 245*, 3–24.

Schumacher, P., Bambini, V., & Weiland, H. (2012). Event-related brain potentials of masked repetition and semantic priming while listening to sentences. *Neuroscience Letters, 530*, 138–143. https://doi.org/10.1016/j.neulet.2012.09.057.

Schumacher, P. B., & Hung, Y.-C. (2012). Positional influences on information packaging: Insights from topological fields in German. *Journal of Memory and Language, 67*(2), 295–310. http://dx.doi.org/10.1016/j.jml.2012.05.006.

Schwarz, F. (2007). Processing presupposed content. *Journal of Semantics, 24*(4), 373–416. http://dx.doi.org/10.1093/jos/ffm011.

Schwarz, F. (2015). Presupposition vs. asserted content in online processing. In F. Schwarz (Ed.), *Experimental perspectives on presupposition, studies in theoretical psycholinguistics* (pp. 89–108). Dordrecht: Springer.

Singh, R., Fedorenko, E., Mahowald, K., & Gibson, E. (2016). Accommodating presuppositions is inappropriate in implausible contexts. *Cognitive Science, 40*(3), 607–634. http://dx.doi.org/10.1111/cogs.12260.

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*, e66839. http://dx.doi.org/10.1371/journal.pone.0066839.

Stalnaker, R. (1974). Pragmatic presuppositions. In M. Munitz, & P. Under (Eds.), *Semantics and philosophy* (pp. 197–213). New York: New York University Press.

Stalnaker, R. (2002). Common ground. *Linguistics and Philosophy, 25*, 701–721. http://dx.doi.org/10.1023/A:1020867916902.

Swaab, T. Y., Camblin, C. C., & Gordon, P. C. (2004). Electrophysiological evidence for reversed lexical repetition effects in language processing. *Journal of Cognitive Neuroscience, 16*(5), 715–726. http://dx.doi.org/10.1162/089892904970744.

Tanner, D., Morgan-Short, K., & Luck, S. J. (2015). How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition. *Psychophysiology, 52*(8), 997–1009. http://dx.doi.org/10.1111/psyp.12437.

Tiemann, S. (2014). *The processing of "wieder" (again) and other presupposition triggers*. University of Tübingen. Doctoral dissertation.

Tiemann, S., Schmid, M., Bade, N., Rolke, B., Hertrich, I., Ackermann, H., et al. (2011). Psycholinguistic evidence for presuppositions: On-line and off-line data. In I. Reich, et al. (Eds.), *Proceedings of Sinn & Bedeutung* (Vol. 15, pp. 581–595). Saarbrücken: Saarland University Press.

Van Berkum, J. J. A., Holleman, B., Nieuwland, M. S., Otten, M., & Murre, J. (2009). Right or wrong? The brain's fast response to morally objectionable statements. *Psychological Science, 20*(9), 1092–1099. http://dx.doi.org/10.1111/j.1467-9280.2009.02411.x.

Van Berkum, J. J. A., Van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience, 20*(4), 580–591. http://dx.doi.org/10.1162/jocn.2008.20054.

Van Herten, M., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research, 22*, 241–255. http://dx.doi.org/10.1016/j.cogbrainres.2004.09.002.

Van Petten, C., Kutas, M., Kluender, R., Mitchiner, M., & McIsaac, R. (1991). Fractionating the word repetition effect with event-related potentials. *Journal of Cognitive Neuroscience, 3*(2), 131–150. http://dx.doi.org/10.1162/jocn.1991.3.2.131.

Wang, L., & Schumacher, P. B. (2013). New is not always costly: Evidence from online processing of topic and contrast in Japanese. *Frontiers in Psychology, 4*, 363. http://dx.doi.org/10.3389/fpsyg.2013.00363.

Yekovich, F. R., & Walker, C. H. (1978). Identifying and using referents in sentence comprehension. *Journal of Verbal Learning and Verbal Behavior, 17*, 265–278. http://dx.doi.org/10.1016/S0022-5371(78)90174-3.

Zeevat, H. (1992). Presupposition and accommodation in update semantics. *Journal of Semantics, 9*(4), 379–412. https://doi.org/10.1093/jos/9.4.379.

Zuur, A., Ieno, E. N., & Smith, G. M. (2007). *Analyzing ecological data*. Springer Science & Business Media.