Distinguishing between different syntactic roles of identical words in normal reading: an ERP study

Lakretz Yair (yair.lakretz@post.tau.ac.il)
Tel-Aviv University
69978, Tel-Aviv, Israel

Marjanovic Katarina (kmarjano@sissa.it)
SISSA, Via Bonomea 265
34136 Trieste, Italy

Gu YuQiao (yuqiaogu@sissa.it)
SISSA, Via Bonomea 265
34136 Trieste, Italy

Treves Alessandro (ale@sissa.it)
SISSA, Via Bonomea 265
34136 Trieste, Italy

Abstract

Separating between semantic and syntactic aspects of language processing in the brain is a difficult task. In an attempt to distinguish between the two, many studies so far have measured responses to semantic or syntactic violations in reading comprehension tasks. However, this methodology may be inaccurate in describing semantic and syntactic processing during normal reading. In this study, we use a novel task, measuring responses to identical target words as they assume different syntactic roles. All sentences presented in the task are syntactically correct sentences without lexical-semantic anomalies. We present results from a behavioral experiment, testing the validity of the experimental design, and results from a pilot ERP study, measuring brain responses to the difference in the syntactic role of the target words. We conclude that the proposed design is valid and may be used to shed light on semantic and syntactic processing during language comprehension, in normal reading.

Keywords: Language comprehension; syntactic violation; ERPs; normal reading task; noun-plus-noun constructions.

Introduction

Semantic and syntactic aspects of language processing are associated with characteristic electrophysiological responses to language stimuli. For example, many studies have shown that the N400 component systematically correlates with lexical-semantic aspects of language processing, and is concluded to reflect lexical-integration processing, e.g. (Kutas & Hillyard, 1980, 1984). Other studies have shown that the syntactic aspect of language processing correlates with the P600 component, e.g. (Friederici, Pfeifer, & Hahne, 1993) - See *related literature* section.

Many of these studies use reading comprehension tasks, involving stimuli of sentences which contain either a semantic or syntactic violation. The type of violation is used to reveal the syntactic or semantic process in question. However, this methodology offers no insight into how these types of processing are distinguished in normal reading, that is, without violating the syntactic rules or semantic expectations.

In this study, we present a novel task for event-related potentials (ERP) studies which enables distinguishing between semantic and syntactic aspects of language processing in normal reading. In the experimental design, participants are presented with target words which differ by their syntactic role but have similar semantic content, and are otherwise (e.g. orthographically) identical. For this, we make use of noun-plus-noun constructions in English, in which the first noun preserves its meaning while changing its position on the syntactic tree, moving from the position of the specifier of a head in a noun phrase (NP), to the head of the NP in a simple sentence without such construction. For example, compare between the word 'family' in 'It's a *family* discount' to the same word in 'It's a *family* from Sweden'.

We conducted two experiments, a behavioral and an ERP experiment. The behavioral experiment was designed to test the validity of the novel task described below, and it is also used to select pertinent stimuli for the task. We then conducted an ERP pilot study using the task and the stimuli which were selected according to the behavioral experiment. We present here qualitative results from this pilot study.

Related literature

ERP signatures of semantic and syntactic aspects of language processing

In order to study the first language (L1) syntactic aspect of language processing, many ERP studies adopted violation paradigms where non-grammatical sentences are compared with correct sentences, which are otherwise similar to the violation stimuli. These studies assume that when all other linguistic variables are held constant, the brain response to the target stimulus, compared to the control stimulus, reflects processes which are related to the grammatical rule in question. The major ERP signatures reported in L1 sentence processing are (for a recent review see (Caffarra, Molinaro, Davidson, & Carreiras, 2015)):

The early left anterior negativity (ELAN) The ELAN component peaks at around 200 ms, with left-anterior distribution, in response to violations of an obligatory phrase structure. It is ascribed to automatic early syntactic parsing processes, during which an initial phrase structure is built - see, e.g. (Hahne & Friederici, 1999; Friederici, 2002; Friederici & Weissenborn, 2007; Steinhauer & Drury, 2012).

The left anterior negativity (LAN) The LAN component peaks at around 400 ms, with left-anterior distribution, in response to morphosyntactic violations such as grammatical agreement violations, tense-marking violations and case-marking violations. It is ascribed to difficulties in integrating morphosyntactic information within a sentence structure with the final goal of thematic role assignment, or mismatch detection during linking processes of agreement computation - see, e.g. (Molinaro, Vespignani, Zamparelli, & Job, 2011; Molinaro, Barber, & Carreiras, 2011; Friederici, 2002; Barber & Carreiras, 2005).

N400 The N400 component peaks at around 400ms, with centro-posterior distribution, in response to lexical-semantic anomalies. It is ascribed to difficulties in processing lexical-semantic information - see, e.g. (Kutas & Federmeier, 2011; Federmeier, 2007; Hagoort, 2003; Traxler & Gernsbacher, 2011; Kutas & Federmeier, 2000).

P600 The P600 component peaks at around 600 ms, with posterior distribution, in response to various violations of syntactic and morphosyntactic features, thematic-rule structure violations, temporary ambiguities, semantic anomalies, and long-distance dependencies - see, e.g. (Friederici et al., 1993; Molinaro, Vespignani, et al., 2011; Molinaro, Barber, & Carreiras, 2011; Carreiras, Salillas, & Barber, 2004). It is ascribed to processes of syntactic reanalysis and repair, and to late integration processes which are not syntactic specific - see, e.g. (Friederici, 2002; Molinaro, Vespignani, et al., 2011; Molinaro, Barber, & Carreiras, 2011; van de Meerendonk, Kolk, Vissers, & Chwilla, 2010; Brouwer, Fitz, & Hoeks, 2012).

First language (L1) - second language (L2) similarity

In this study, we tested native Italian speakers, with high proficiency in English, on a task in English. The brain responses of these participants were recorded in an ERP design. Previous studies have found that participants with high proficiency in L2 have similar brain responses compared to L1 speakers. Several ERP experiments were conducted by Rossi, Gugler, Friederici, & Hahne (2006) on L2 speakers presenting sentences with morphosyntactic and phrase structure violations. Results show that low-proficiency L2 speakers did not show a LAN effect for morphosyntactic violations with a delayed P600 for both types of violations (Ojima, Nakata, & Kakigi, 2005), as compared to natives. However, participants with high proficiency showed similar response to that observed with L1 controls. The authors concluded that, at high-enough L2 proficiency levels, an L1-like brain response can be observed, reflecting early automatic parsing processes followed by late processes of re-analysis and repair.

Noun-plus-noun constructions in English

Noun-plus-noun (NNs) constructions are composite nominals in which both the head and the attributive dependendat(s) are nouns (Garnier, 2011), e.g. *family discount*, *bus driver*. These types of NNs are a common type of constructions in the English language, however their role in its grammar is nevertheless still an ongoing debate among linguists, as some classify them as a phrase, originating in the syntax (Garnier, 2011), while the others claim them as compounds, originating in the lexicon (Giegerich, 2004). Another group of studies claims they can belong to both categories (Payne & Huddleston, 2002). Furthermore, (Giegerich, 2004) also distinguishes between the fore-stressed and end-stressed NNs, assigning them into two different origins. Avoiding this debate, and in order to keep a homogeneous set of stimuli in the experiment, we therefore chose NNs with a fore-stress only.

Methods

Participants

The participants of this study are native Italian speakers, students in Scuola Internazionale Superiore di Studi Avanzati (SISSA, Italy), with high proficiency of English. All the participants are between 18 and 40 years of age and right-handed. In the behavioral experiment, 9 participants have taken part. In the ERP experiment, 5 participants have participated who did not take part in the behavioral experiment.

Stimuli

The stimuli consist of 52 syntactically and semantically correct sets of English sentences. Each of the 52 sets consists of a short quiz question, a beginning of an answer (including the target word), and 4 possible completions. There are 52 different quiz questions, 26 different beginnings of answers (each of these repeats twice, in each of the two conditions), and 26 * 4 different completion options.

Design and procedure

We present a novel task for ERP studies which enables to distinguish between semantic and syntactic aspects of language processing in normal reading. To do so, participants are presented with target words which differ by their syntactic role but have similar semantic content, and are otherwise (e.g. orthographically) identical. We make use of noun-plus-noun constructions in English to contrast between, e.g. the word 'family' in

(S1) It's a *family* discount. (*noun-plus-noun condition*)
(S2) It's a *family* from Sweden. (*single noun condition)*

However, such a comparison is only possible if the syntactic expectation of the participant in (S1) is the desired one while reading the target-word. For example, such a comparison would fail if the participant comprehends ’family’ in (S1) as a pre-head only after having completed and reanalysed the sentence. We therefore manipulate the syntactic expectation of the participant by preceding the sentence with a quiz question the answer to which requires the desired syntactic role only. Continuing with the above example, we precede (S1) with the following quiz question:

(Q): It will get you a cheaper entrance to the pool. What is it?

The quiz is then followed by a beginning of an answer:

(A) ’It’s a family _____’.

Note that whether the participant knows the correct answer to the above quiz question is unimportant. Even without knowing the correct answer to (Q), we hypothesised that one would expect the answer (A) to the type of question in (Q) to end with a noun, thus reading ’family’ as an adjective. We assume that this kind of expectations are also enhanced after the practice block. Note also, that data analysis focuses on the time during which the participant reads the target word (e.g. ’family’), *before* she is asked to complete the sentence. We therefore regard it as normal reading, and are not concerned with other processes that may follow.

Importantly, while creating syntactic expectation, the preceding quiz question must not have created semantic expectation to the target word. Therefore, all words in the quiz questions were made sure not to be semantically related to the target word in their answers (as can be assessed with Latent Semantic Analysis). For example, no word in (Q) semantically primes the target-word in (A). An additional benefit to this manipulation is that it enhances the engagement of the participants in the task, by challenging them with quiz questions.

In order to test our hypothesis that the quiz questions induce the correct syntactic expectation, we ran a behavioral experiment that assesses the syntactic expectation of the participant when reading the target word. This experiment is described below. According to the results of the behavioral experiment, we chose the quizzes which best manipulated the syntactic expectations of the participants. These quizzes were then used in the ERP experiment, described in the following section.

Behavioral experiment In order to test the syntactic expectation of the quiz question, we ran a behavioral experiment in which participants are asked to complete target sentences after reading the quiz. The design of this task is shown in Figure 1. Participants are presented with 5 subsequent screens: (1) A quiz question, (2) ”It’s a” (beginning of the answer), (3) fixation cross, (4) target word (e.g. ’family’), (5) Textbox. participants are asked to read the quiz question and then to complete the sentence in a textbox, using the keyboard, after having read the beginning of the sentence. Responses and reaction times are recorded during the experiment.

The list of quiz questions contained 51 quizzes from the noun-plus-noun condition, and 46 quizzes from the simple-sentence condition. All quizzes were presented to the participants in a random order.

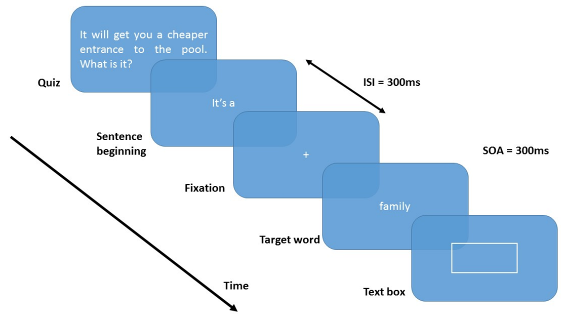

Figure 1: The design of the behavioral experiment.

ERP experiment Brain responses of another group of participants, none of whom participated in the behavioral experiment, were recorded with 128-channel EEG. Stimuli were presented to the participants on a computer screen, in a similar manner to that in the behavioral experiment. Before starting the experiment, participants execute a practice section containing 10 quiz questions which are different from the ones later presented in the experiment. The experiment consists of 5 blocks, each containing 52 quiz questions, to which four possible answers are given (see section *Stimuli*). The list of sentences within each block is presented to participants in a random manner, which is different for each block.

The quiz question is presented on the screen until the participants decide to continue by pressing a key. After having pressed a key, the following three screens are presented: (1) ”It’s a”, (2) target word (e.g. ’family’), (3) dashed line (_____) (ISI=300ms). The last screen of the dashed line is then followed by an option screen, containing four possible answer completions, randomly ordered in each trial and block (Figure 2). The five blocks are separated by breaks, during which the participants remain in their position in front of the screen. The participants determine the length of the breaks by themselves.

EEG recording

The EEG was continuously recorded using the ActiveTwo BioSemi EEG system (BioSemi V.O.F., Amsterdam, Netherlands) with 128 channels covering the entire scalp. EEG signals were sampled at 512 Hz with band-pass filters set at 0.1 – 100 Hz.

Data analysis

Acquired data is analysed using EEGLAB, open source MATLAB (The Matworks, Natick, MA) toolbox for EEG processing. Data is first high-pass filtered at 1 Hz and low-pass filtered at 40Hz, and re-referenced. Next, follows an extrac-

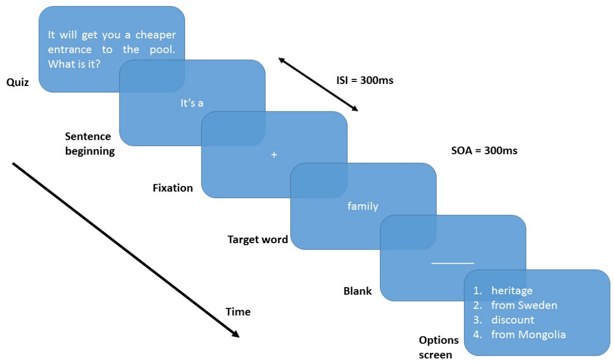

Figure 2: The design of the ERP experiment.

tion of the epochs of the two conditions. In both epoch sets, the answer of the participants are divided into three groups: the correct answer (e.g. 'discount' in (Q)); the semi-correct answer (e.g. 'heritage' in (Q)), which is syntactically correct, suggesting that the participant comprehended the target word in the desired syntactic role; and wrong answers (e.g. 'from Sweden' and 'from Mongolia' in (Q)). The answers are counterbalanced over conditions such that, for example, 'from Sweden' is the correct answer, 'from Mogolia' the semi-correct, and 'heritage' and 'discount' are the wrong answers in the second condition. All the wrong answers are omitted and only the correct and semi-correct answers are included in the analysis. The data is pre-processed and artifacts are omitted using independent component analysis (ICA).

Results

Behavioral experiment

We tested whether quiz questions can induce the desired syntactic expectation when participants read the target sentence. Figure 3 presents the results for this experiment, showing for each quiz the mean accuracy, calculated over all participants.

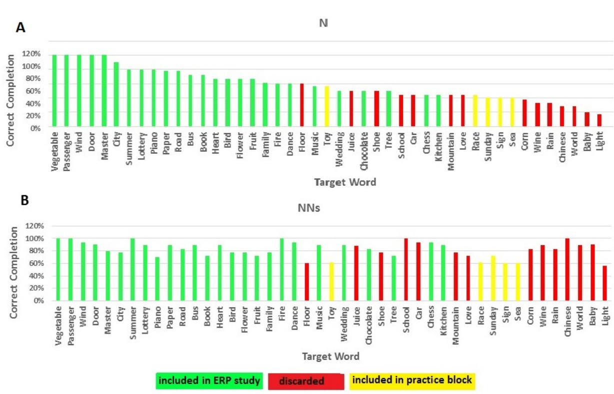

Figure 3: Results of the behavioral experiment (A) Correct completion rates for each quiz question, for the condition of noun (N) (B) Correct completion rates for each quiz, for the condition of noun-plus-noun construction (NNs).

We found that the manipulations for the noun-plus-noun construction achieve higher average accuracy (average accuracy = 0.80 ± 0.15), in comparison to the single nouns (average accuracy = 0.54 ± 0.25).

For the ERP experiment, we then chose for each of the two conditions 26 quiz questions with the highest accuracy score (NNs - 0.86 ± 0.09, single nouns - 0.72 ± 0.17). These quiz questions are then used in the ERP experiment.

ERP experiment

Since the experiment is still ongoing, we present ERP results from our pilot study (Figure 4).

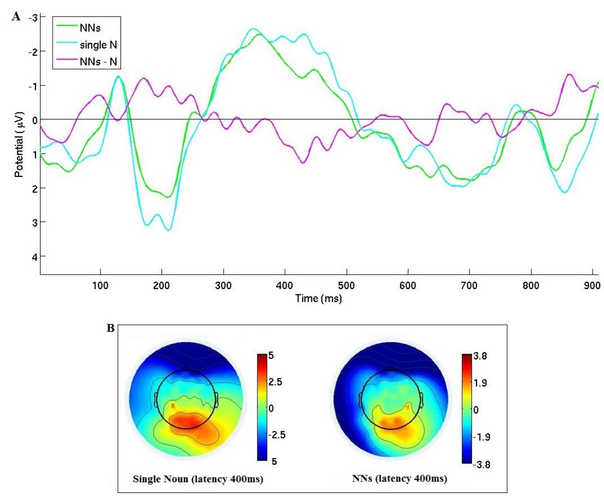

Figure 4: Results of the ERP experiment for an arbitrarily chosen participant: (A) The ERP signatures for syntactic and semantic processing, revealing different responses between the two conditions (B) brain response during processing of single nouns (left) and noun-plus-noun constructions (right).

Summary and discussion

In this study, we present a novel paradigm to disentangle the syntactic aspect of language processing from its semantic aspect. The paradigm avoids the use of syntactic violations, and focuses on normal reading of correct sentences. We make use of noun-plus-noun constructions in English to contrast between two conditions which differ by their syntactic aspect only. In order to manipulate the syntactic expectation of the participant before reading the target word, we use quiz questions, such that the answer to the questions allows only one possible syntactic role to the target word.

We tested the method of manipulating the syntactic expectations of the participants in a behavioral experiment. Results support the validity of the proposed experimental design, and are also used to choose quiz questions with highest score of manipulation. Results suggest that the syntactic manipulations are more effective for the noun-plus-noun conditions, but are reasonably effective for both conditions for the selected group of stimuli.

Following the behavioral experiment, we ran an ERP pilot study using the novel paradigm and selected quiz questions. Albeit only qualitative, pilot results are showing promising disclosure of different ERP signatures for the two conditions, opening a new window into syntactic processing during language comprehension in normal reading. We believe that it may provide a way to relate observable signals in the human brain to hypothesised mechanisms in models of latching dynamics (Russo, Namboodiri, Treves, & Kropff, 2008; Russo & Treves, 2012), in particular at the transition between words (Pirmoradian & Treves, 2013).

References

Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in spanish: An erp comparison. *Journal of Cognitive Neuroscience*, *17*(1), 137–153.

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: rethinking the functional role of the p600 in language comprehension. *Brain Research*, *1446*, 127–143.

Caffarra, S., Molinaro, N., Davidson, D., & Carreiras, M. (2015). Second language syntactic processing revealed through event-related potentials: an empirical review. *Neuroscience & Biobehavioral Reviews*.

Carreiras, M., Salillas, E., & Barber, H. (2004). Event-related potentials elicited during parsing of ambiguous relative clauses in spanish. *Cognitive Brain Research*, *20*(1), 98–105.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, *44*(4), 491–505.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in cognitive sciences*, *6*(2), 78–84.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive brain research*, *1*(3), 183–192.

Friederici, A. D., & Weissenborn, J. (2007). Mapping sentence form onto meaning: The syntax–semantic interface. *Brain research*, *1146*, 50–58.

Garnier, M. (2011, septembre). Correcting erroneous N+N structures in the production of French users of English as part of the CorrecTools project (regular paper). In *EUROCALL, Nottinghma, 31/08/2011-03/09/2011* (pp. 157–165). http://www.eurocall-languages.org: Eurocall Association.

Giegerich, H. J. (2004). Compound or phrase? english noun-plus-noun constructions and the stress criterion. *English Language and Linguistics*, *8*(01), 1–24.

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: Erp effects of combining syntactic and semantic violations. *Cognitive Neuroscience, Journal of*, *15*(6), 883–899.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, *11*(2), 194–205.

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

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the n400 component of the event related brain potential (erp). *Annual review of psychology*, *62*, 621.

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

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*(307), 161–3.

Molinaro, N., Barber, H. A., & Carreiras, M. (2011). Grammatical agreement processing in reading: Erp findings and future directions. *cortex*, *47*(8), 908–930.

Molinaro, N., Vespignani, F., Zamparelli, R., & Job, R. (2011). Why brother and sister are not just siblings: Repair processes in agreement computation. *Journal of Memory and Language*, *64*(3), 211–232.

Ojima, S., Nakata, H., & Kakigi, R. (2005). An erp study of second language learning after childhood: Effects of proficiency. *Cognitive Neuroscience, Journal of*, *17*(8), 1212–1228.

Payne, J., & Huddleston, R. D. (2002). Nouns and noun phrases.

Pirmoradian, S., & Treves, A. (2013). Encoding words into a potts attractor network. In *Proceedings of the thirteenth neural computation and psychology workshop (ncpw13) on computational models of cognitive processes, world scientific press, singapore* (pp. 29–42).

Russo, E., Namboodiri, V. M., Treves, A., & Kropff, E. (2008). Free association transitions in models of cortical latching dynamics. *New Journal of Physics*, *10*(1), 015008.

Russo, E., & Treves, A. (2012). Cortical free-association dynamics: Distinct phases of a latching network. *Physical Review E*, *85*(5), 051920.

Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (elan) in syntax studies. *Brain and language*, *120*(2), 135–162.

Traxler, M., & Gernsbacher, M. A. (2011). *Handbook of psycholinguistics*. Academic Press.

van de Meerendonk, N., Kolk, H. H., Vissers, C. T. W., & Chwilla, D. J. (2010). Monitoring in language perception: mild and strong conflicts elicit different erp patterns. *Journal of Cognitive Neuroscience*, *22*(1), 67–82.