

Detecting presupposition failure and accommodation with EEG

Alice Xia (alicexia@cmail.carleton.ca)

Institute of Cognitive Science, Carleton University, 1125 Colonel By Dr
Ottawa, ON K1S 5B6 Canada

Roxana M. Barbu (roxanamariabarbu@cmail.carleton.ca)

Institute of Cognitive Science, Carleton University, 1125 Colonel By Dr
Ottawa, ON K1S 5B6 Canada

Kathleen Van Benthem (kathy.vanbenthem@carleton.ca)

Institute of Cognitive Science, Carleton University, 1125 Colonel By Dr
Ottawa, ON K1S 5B6 Canada

Daniel A. Di Giovanni (daniel.digiovanni@mail.mcgill.ca)

Montréal Neurological Institute and Hospital, 3801 Rue Université
Montréal, QC H3A 2B4 Canada

Ida Toivonen (ida.toivonen@carleton.ca)

Institute of Cognitive Science, Carleton University, 1125 Colonel By Dr
Ottawa, ON K1S 5B6 Canada

Raj Singh (raj.singh@carleton.ca)

Institute of Cognitive Science, Carleton University, 1125 Colonel By Dr
Ottawa, ON K1S 5B6 Canada

Abstract

Sentence comprehension in part involves introducing, storing, and retrieving information about individuals. Natural languages provide various means for performing this computational work. One popular idea is that indefinite noun phrases provide instructions for updating the discourse model by adding a new discourse referent, while definite noun phrases presuppose the existence of a discourse referent available in memory, as well as instructions for retrieving it. When no antecedent is available, the definite's presupposition fails to be satisfied, resulting in the so-called 'presupposition failure' and pragmatic infelicity. However, under certain conditions, definite noun phrases *can* felicitously be used even when no antecedent is available in memory. In such cases, a conversational repair strategy called 'presupposition accommodation' can rescue the discourse by adding the required referent. It is natural to expect greater processing costs for adding a discourse referent with a definite than with an indefinite: although both involve the process of adding a referent, definites go through a stage of presupposition failure and a subsequent decision to accommodate. The experimental challenge has been to apply a method sensitive enough to detect expected costs in discourse, even when the participant is unaware of the presupposition failure and repairs it rapidly. The present study addresses this challenge by using EEG to capture temporally fine-grained processing differences between definite and indefinite noun phrases when both introduce new discourse referents in plausible and implausible contexts. Our main finding is that definite noun phrases elicit the Left Anterior Negativity (LAN) effect, compared to indefinite noun phrases, both in implausible contexts where there is a sense of oddness and in perfectly coherent contexts. We take this as evidence of a specific cognitive stage at which presupposition failure is detected and when an accommodation decision occurs. This also supports the idea that, when encountering a definite, the LAN

is tightly linked to working memory processes involving the search for discourse elements that are presupposed to exist in memory. When none are found, definites are subsequently accommodated and bridged to other entities in the discourse.

Keywords: discourse; presuppositions; context; accommodation; EEG

Introduction

Presuppositions in natural language are commonly viewed as pieces of information that impose constraints on the contexts in which they are triggered. Just as pronouns like "she" require that the context furnish a (uniquely) salient female, presuppositions require that the context entail them. For example, consider a command like (1):

(1) # After you read this paper, go call the waiter.

The sentence in (1) is strange when uttered out of the blue. It is strange for the same reason that *go call her* is strange when there is no salient female in the context. The sentence has been uttered in a context that is missing something that the sentence needs - in this case, a uniquely salient waiter. We refer to such cases as 'presupposition failure': the failure is technical (i.e., the context does not entail the presupposition), and this technical failure leads to a discourse failure. Note that there is nothing inherently odd about the sentence in (1); it is odd when the context in which it is uttered fails to provide a waiter as antecedent for the definite phrase *the waiter*. If we introduce a waiter into the prior context, meaning that there

is no longer any technical presupposition failure, the oddness disappears:

- (2) A waiter and a cook came by and left a flyer. After you read this paper, go call the waiter.

Note also that the oddness disappears when we change “the” to “a”; since the latter has no presuppositions, there is no threat of presupposition failure, and hence none of the oddness that is experienced in (1).

- (3) After you read this paper, go call a waiter.

You might not know why the speaker is telling you to call a waiter, but nothing has gone wrong as far as language itself is concerned.

This connection between presupposition and the prior context is the centrepiece of the so-called “satisfaction theory” of presupposition (e.g., Karttunen, 1974; Stalnaker, 1974; Heim, 1983). Its basic assumption is that a sentence S with presupposition p , S_p , may be used in context c only if c ‘satisfies’ p , i.e., only if $c \subseteq p$. The satisfaction theory adds additional auxiliary assumptions to deduce predictions about presupposition projection, that is, about the presuppositions of complex sentences. For example, it predicts that $\neg S_p$ presupposes p but that *if T then S_p* presupposes $A \rightarrow p$ (e.g., Heim, 1983). We will limit our attention to atomic sentences here.

A *prima facie* challenge for the satisfaction theory comes from the observation that it is often possible to felicitously use S_p even when c does *not* satisfy p . In other words, there appear to be instances of technical presupposition failure without any sense of a higher discourse failure. For example, consider the following text (modified from Singh, Fedorenko, Mahowald, & Gibson, 2016):

- (4) I went to a restaurant last night. The waiter yelled at me.

In (4), the context in which *the waiter* is uttered does not entail that there was a waiter. It is plausible, of course, that there should be waiters in the restaurant, but this information is not strictly entailed by the context. Nevertheless, there is no hint of oddness here.¹

The satisfaction theory explains the contrast between (4) and (1) by appealing to what is called ‘presupposition accommodation’ (Lewis, 1979). When addressees hear or read a definite description like *the waiter* in a context that does not furnish an antecedent, they face a choice: they can either accommodate the required presupposition; that is, they

¹A reviewer raises the question about the relative appropriateness of “I went to a restaurant last night. The waiter yelled at me” and “I went to a wedding. The bride talked to me.” Weddings typically have one and only one bride, while there may be no waiters or many waiters at a restaurant, and the reviewer suggests that these considerations might lead to differences in appropriateness judgments. We are not aware of work on this, and we hope to return to it in future work. We intentionally designed contexts that would (i) allow multiple referents, such as multiple waiters – this is that indefinites could also be used in these contexts), with one being uniquely salient (e.g., the waiter who serves you).

can ‘quietly and without fuss’ (Von Stechow, 2008) adjust the context by adding the missing presupposition, or they can let the discourse come to a crashing halt. If the context is one that makes it reasonable to accommodate, say if the presupposition is unsurprising or uncontroversial, then cooperative speakers will recognize that they should keep the discourse running and will therefore simply accommodate.² Viewed in this light, accommodation is a repair mechanism that can fix a context so that technical presupposition failure – the failure to initially find the required antecedent – does not become pragmatic presupposition failure. By ‘pragmatic presupposition failure’ we mean that the context does not get amended and the discourse is interrupted because the definite noun phrase is unable to do its job. It is considered bad conversational practice to rely on accommodation when the presupposition is somehow noteworthy. If the addressee faces the choice of having to either let the discourse crash because of pragmatic presupposition failure, or accommodate a presupposition that is surprising, controversial, or otherwise hard to incorporate into the context, then the addressee would rightly feel that the speaker is asking too much of them.

The appeal to accommodation has been controversial (e.g., Gazdar, 1979; Van der Sandt, 1992; Gauker, 1998; Abbott, 2006). The satisfaction theory predicts that the addressee has passed through a stage of processing at which technical presupposition failure was detected but was then overcome by the accommodation repair. However, there is no trace of this failure detection in our conscious awareness, and it would be desirable to find a way to measure whether accommodation is real and whether it is indeed triggered by a stage of technical presupposition failure.

Previous psycholinguistic studies have found that S_p is easier to process in contexts that satisfy p than in contexts that do not (e.g., Haviland & Clark, 1974; Crain & Steedman, 1985; Burkhardt, 2006; Schwarz, 2007). This might be thought to lend support to the satisfaction theory. However, this processing difference may not be about presupposition accommodation itself; instead, it could have arisen from the fact that in contexts in which p is not satisfied, there is an extra step of adding p to the context. This additional step may have been responsible for the extra costs whether or not there is any purported technical presupposition failure from which the addressee may choose to recover using accommodation (see Singh et al., 2016 for discussion).

To control for this, we would need a minimal pair that would also involve adding p to the context, but through assertion rather than presupposition accommodation. Indefinite articles provide the required contrast:

- (5) I went to a restaurant last night. A waiter yelled at me.

²Heim (1982) argued that with definites, there must be a prior discourse referent that the definite can ‘bridge’ to (in the sense of Clark, 1975). For example, in (4) the introduced waiter can ‘bridge’ to the restaurant mentioned in the first sentence, such that *the waiter* is roughly understood as ‘the waiter at the restaurant’, and typically with further identifying features (e.g., the waiter who served you at the restaurant – see Note 1).

In both (4) and (5), the existence of a waiter is added to the discourse context (e.g., Heim, 1982). For example, suppose with Heim (1982) that discourse referents can be thought of as file cards that can be introduced, referred to, or taken out of the ‘file’ that collects the discourse information as it accumulates. An indefinite noun phrase such as *a waiter* simply adds a new file card with ‘is a waiter’ on it. A definite like *the waiter* scans the file to find a file card with ‘is a waiter’ on it. If one exists, it refers to it; if it doesn’t, then either communication fails, or the missing file card is accommodated. This is what appears to happen in (4). Thus, the processing of both (5) and (4) involves adding a file card corresponding to a waiter, but only in (4) do you also go through a stage of recognizing that something is wrong with the context (there is no antecedent file card). The addition of a (file card corresponding to a) waiter is then an accommodation response to this recognition; we will sometimes use ‘referents’ and ‘antecedents’ when we mean file cards, as they are probably more familiar, but it is worth noting that file cards are the underlying technical object that is being manipulated (see Heim, 1982). What we would like, then, is to test whether there is indeed a stage at which the processing mechanism seeks but fails to find an antecedent for *the waiter*, and then repairs for this by accommodating one. The satisfaction theory predicts that there should be such stages; presumably the performance system executes these computations demanded by the competence system, and if so, we might expect to find reflexes of them during language processing.

Singh et al. (2016) performed an online incremental stops-making-sense (SMS) task to examine participants’ appropriateness judgments about indefinite and definite noun phrases in plausible contexts like (4) and (5) as well as in implausible contexts like the following:

(6) I went to a jail last night. {A/the} waiter yelled at me.

They found a main effect of plausibility, such that implausible conditions had more and earlier SMS judgments. This was unsurprising, given that implausible information is generally harder to process than plausible information (e.g., Trueswell, Tanenhaus, & Garnsey, 1994; Gibson & Perlmutter, 1998). More interestingly, they found an interaction, such that implausible definites had earlier and more SMS judgments than implausible indefinites. This provides support for the claim that accommodation is subject to stricter requirements than assertion. In particular, it is inappropriate to force your addressee to accommodate implausible information as a presupposition; such information is better expressed as an assertion so that your addressee is at least given the opportunity to challenge it (e.g., Soames, 1989; Heim, 1992; Beaver, 2001; Von Stechow, 2008).

However, together with appropriate auxiliary assumptions about how the competence theory is realized in performance (see above), it is plausible that the satisfaction theory would expect accommodation difficulty relative to indefinite controls not only in implausible contexts but also in plausible

contexts (though cf. Stalnaker, 2002). There is technical presupposition failure in both plausible and implausible contexts. Pragmatic presupposition failure of course is more easily averted in plausible contexts than implausible ones. Thus, the enhanced difficulty of implausible definites makes sense. However, the predicted stage of technical failure in plausible contexts did not reveal itself. Perhaps the method was inappropriate for detecting such a stage, if there is one. We have seen that accommodation is not sensed as odd or costly when the presupposition is sufficiently supported in the context. Thus, it is perhaps not surprising that participants’ SMS judgments did not differentiate between plausible definites and plausible indefinites. It remains an open question, then, whether an empirical cognitive account can be found for when the absence of an expected antecedent is noticed and when we decide to accommodate in response.

We explored this question by means of an electroencephalography (EEG) study using materials from Singh et al. (2016). We give a comprehensive account of our materials and methods momentarily, but briefly our goal was to compare definites and indefinites in contexts in which both would have the effect of introducing a new discourse referent into the context. The relevant difference between the two is that definites presuppose the existence of an object in memory and aim to retrieve it while indefinites introduce a new object into memory. Ideally, we want these objects to be the same, or as close to that as possible. That way, any detected difference between the two could be plausibly attributed to the assumption that definites introduce the desired object only when the search for it fails. We wanted to see if we could find an EEG signature of this hypothesized failure and repair. Previous EEG studies comparing definites and indefinites did not isolate this difference between definites and indefinites. Experiment 2 of Anderson and Holcomb (2005) had a definite and indefinite condition but the definite in these cases *had* an antecedent and the indefinite (as it does) introduced a new discourse referent. Schumacher (2009), building on Burkhardt (2006) (which investigated definites alone), included a definite given condition and an indefinite given condition (in which an indefinite NP in the second sentence has a matching indefinite NP in the first sentence); but the latter texts are odd (for reasons we discuss shortly), and hence the definite and indefinite are not properly matched, and in any event, this ‘given’ condition breaks our desired symmetry under which definites and indefinites both introduce a new discourse referent in all conditions.

Methods

Participants

Thirty-four participants were recruited from Carleton University. As compensation for participating in the experiment, students received 3% class credit towards a first-year cognitive science course. All students were English speakers between the ages of 18 and 24.

Materials

We used shortened versions of all 128 sentence pairs from Singh et al. (2016) as our experimental stimuli. The sentence pairs were divided evenly into four blocks using a Latin Square design as illustrated in Table (1):

Table 1: Sample stimuli

Indefinite Plausible:
Philip went to a <i>pool</i> on Tuesday evening.
A <i>swim instructor</i> insulted him there.
Definite Plausible:
Philip went to a <i>pool</i> on Tuesday evening.
<i>The swim instructor</i> insulted him there.
Indefinite Implausible:
Philip went to a <i>laboratory</i> on Tuesday evening.
A <i>swim instructor</i> insulted him there.
Definite Implausible:
Philip went to a <i>laboratory</i> on Tuesday evening.
<i>The swim instructor</i> insulted him there.

Procedure

Participants sat in a Faraday cage in front of a computer monitor and were instructed to read all sentence stimuli for comprehension. Using PsychoPy, all stimuli were presented visually in the center of the monitor in white letters against a grey background. A practice session consisting of four trials was completed before beginning each session. The first sentence in the sentence pair appeared in full for 3000 ms, followed by 100 ms of a blank screen. Our critical noun phrase (e.g., “the lion”) in the second sentence of the pair then appeared on screen for 600 ms, followed by another 100 ms of a blank screen. The remaining non-critical segments of the second sentence, which had an average length of three words, appeared for 400 ms. All participants saw all items in all conditions, counterbalancing block orders.

EEG Recording

A 128-channel HydroCel Geodesic Net was used to record continuous EEG signals against Cz as reference, at a sampling rate of 250 Hz, with Net Station 4.3.1. Electrode impedance was kept below 5 kOhms.

Data Analysis

Data from two participants were excluded due to excessive noise during EEG recording. Four channels (E68, E73, E88, E94) were removed prior to preprocessing as is common for high-density electrode nets (to allow the plug-in of other external biometric devices). EEG recordings were re-referenced

offline to the average and digitally filtered with a low-pass of 0.5 Hz and a high-pass of 30 Hz. Filtered data were then epoched from 500 ms before to 1000 ms after the critical noun phrase.

Subject data were preprocessed using a combination of EEGLAB 14.1.2 (Delorme & Makeig, 2004) and custom-written MATLAB scripts. Independent component analysis (ICA) in EEGLAB was used to first remove eye-blinks and other physiological noise. The CleanLine toolbox (Mullen, 2012) was used to reduce drift. Channels that were three standard deviations away from the mean, based on a power spectrum threshold, were removed. Lastly, an automatic component rejection was performed using the MARA toolbox (Winkler, Haufe, & Tangermann, 2011).

Following previous literature, two time windows were selected for analysis: 300-500 ms and 500-700 ms after onset of the critical noun phrase. This allowed us to examine the N400/P600 complex (Burkhardt, 2006), as well as the LAN effect (Kutas & Federmeier, 2007). Using the EEGLAB Darbeliai extension, event-related potentials were computed for the 1000 ms after stimulus onset relative to a 100 ms pre-stimulus baseline for each participant, for each condition, from electrodes clustered in each of the following four regions: left anterior (F3/F7/FC3/FT7), right anterior (F4/F8/FC4/FT8), left posterior (P3/T5/CP5/T5), and right posterior (P4/T6/CP6/T6).

For statistical analyses, mean amplitude data were submitted to linear mixed-effects models using the lme4 package (Bates, Mächler, Bolker, & Walker, 2014) in R (R Core Team, 2013). Significance testing was done using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017). In both time windows of interest, our models evaluated mean amplitude as a function of a three-way interaction among Plausibility (plausible, implausible), Definiteness (definite, indefinite), and Electrode Region (left anterior, right anterior, left posterior, right posterior). Participant was included as a random factor. We performed planned comparisons between our four conditions (Definite Plausible, Indefinite Plausible, Definite Implausible, Indefinite Implausible) if significant interactions were found between region and definiteness or plausibility. Pairwise contrasts were investigated using the emmeans package in R (Lenth, 2018) and p -values were adjusted using the Bonferroni correction.

Results

300-500 ms time window: We found a significant interaction between Region and Definiteness ($F(3, 1946) = 5.64, p < .001$). In particular, the Definite - Indefinite condition contrast in the group of left anterior electrodes was significant (beta = $-0.37, t = -5.24, p < .001$) (Figure 1). This general determiner effect was further corroborated by a scalp map of the same left anterior electrodes in the 300-500 ms time window, averaged across all participants for each condition (Figure 2). We further found a significant difference in the Definite Implausible - Indefinite Implausible

contrast in the same region ($\beta = -.35, t = -3.59, p = .002$), which was reflected in greater negativity elicited by definite noun phrases (e.g., “the lion”) compared to indefinite noun phrases (e.g., “a lion”) (Figure 3). Similarly, in the plausible context, the Definite Plausible - Indefinite Plausible contrast revealed a significantly more negative deflection for definite noun phrases ($\beta = -.37, t = -3.81, p < .001$) (Figure 4).

500-700 ms time window: In the left anterior, we again found a significant interaction between Region and Definiteness ($F(1, 1946) = 22.17, p < .001$). Contrast analyses were significant for Definite - Indefinite ($\beta = -.22, t = -4.05, p < .001$) (Figure 1), as well as for Definite Implausible - Indefinite Implausible ($\beta = -.30, t = -3.86, p < .001$) (Figure 4).

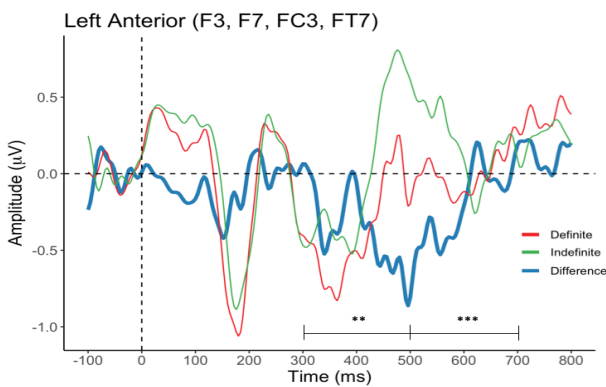


Figure 1: Significant negative-going ERP elicited by the Definite condition relative to Indefinite condition in the 300-500 ms time window in left anterior electrodes, reminiscent of the Late Anterior Negativity (LAN). Significant divergence between the two conditions continues into the 500-700 ms time window.

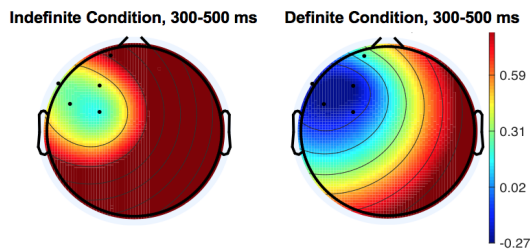


Figure 2: Scalp map of left anterior electrodes for Indefinite (left) and Definite (right) conditions in the 300-500 ms time window.

Discussion

Our study used EEG to explore the consequences of processing definite and indefinite noun phrases in plausible and implausible sentence contexts. Our goal was to test both types of determiners together, to isolate the crucial stages of presupposition failure and accommodation of a new discourse

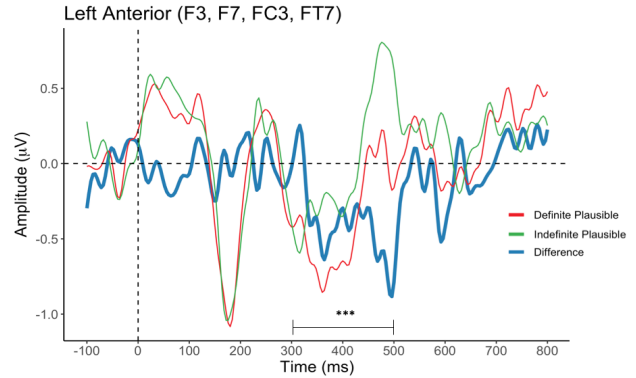


Figure 3: Greater negativity elicited by the Definite Plausible condition relative to the Indefinite Plausible condition in the 300-500 ms time window.

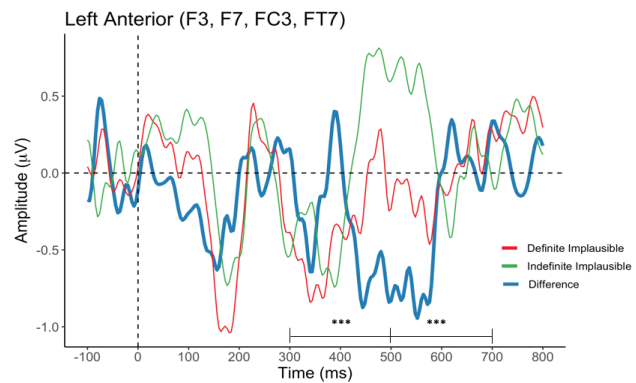


Figure 4: Greater negativity elicited by the Definite Implausible condition relative to Indefinite Implausible in the 300-500 ms time window. Significant divergence between the two conditions continues into the 500-700 ms time window.

referent (see our earlier discussion of Burkhardt, 2006 and Schumacher, 2009; see also Hirotsu & Schumacher, 2011). While Schumacher (2009) also investigated definite and indefinite phrases together, the pragmatic appropriateness of the phrase pairs in that task was not stable across contexts. For example, consider a given definite text like *Peter has recently visited a speaker in Munich. He said that the speaker had been very nice.* The text is coherent and there is no sense of oddness. This cannot be said of the indefinite given counterpart: *Peter has recently visited a speaker in Munich. He said that a speaker had been very nice.* This text is decidedly odd, confirmed by offline data reported in Schumacher (2009). Presumably, the oddness is due to so-called *Maximize Presupposition!* effects (e.g., Heim, 1991; Singh, 2011), which essentially demand that the speaker use a presuppositional alternative (like a definite) when its presupposition is satisfied instead of a non-presuppositional minimal variant (like an indefinite). *Maximize Presupposition!* has been proposed to explain the oddness of sentences like *A sun is*

shining; the definite variant is preferred because of *Maximize Presupposition!* (see also Note 1). Similarly, in the case under current consideration, the indefinite *a speaker* would be ruled out in favour of *the speaker* (if the same speaker is intended), or in favour of *another speaker* (if a different speaker is intended).

In our study, the definite/indefinite pairs are both either appropriate or both inappropriate in their given contexts (cf. stimuli norming results in Singh et al., 2016). There is no influence of *Maximize Presupposition!* because there is no ‘given’ context. All conditions required the addition of a new discourse referent, and hence were expected to not differ with respect to the P600 (given the findings in Burkhardt, 2006; Schumacher, 2009). Somewhat to our surprise, our results showed a greater positive deflection for indefinites than definites in the implausible context during the 500-700ms window. This might suggest a P600, but we are not confident that it is since the P600 is typically found over parietal lobes (e.g., Osterhout & Holcomb, 1992). Whatever this effect’s proper classification, the difference does not replicate the finding in Schumacher (2009) that definites and indefinites both generate a late positivity (a P600 in her studies) that indexes the addition of a discourse referent. The difference here might be teaching us that by this stage the accommodation for the definite has already occurred, or that the introduction of referents via assertion is different than via accommodation, and the late positivity we found for indefinites indexes only assertive updates.

Here we tentatively suggest that by that late stage the accommodation for the definite has already taken place. According to the satisfaction theory, assertions are updates to the context, and the presuppositions in a context get updated by the assertion. Thus, there is an implied temporality: presuppositional matters are resolved prior to assertive updates (hence the *pre-*). Thus, it is conceivable that the accommodation step occurs early, right after the detection of the technical presupposition failure. Perhaps this bundle of computations is what our left-lateralized frontal negativity for definite noun phrases in the 300-500ms window was indexing. This Left Anterior Negativity (LAN) has been found in previous studies of (in-)definiteness (e.g., Anderson & Holcomb, 2005; Schumacher, 2009), but for reasons discussed earlier it is hard to interpret such findings because the contrasts between definite and indefinite conditions were not quite minimal. More generally, the LAN has been linked to processes of working memory resources that involve ‘reactivating’ previous entities or forming dependencies between new and old entities (see e.g., King & Kutas, 1995; Kirsten et al., 2014, among others). We tentatively propose here that the detection of presupposition failure and accommodation are among the computations the LAN indexes. Note that definites involve the search for entities in memory (‘reactivation’), and that accommodation when no antecedent is found typically involves ‘bridging’ the new entity to a previous entity (e.g., linking ‘the waiter’ to the previously mentioned restaurant – see also

Note 2).

The design of our study was based on a previous stops-making-sense task that investigated temporal decisions during the silent reading of definite and indefinite phrases in contexts that varied in plausibility (Singh et al., 2016). Based on the results of that study, we initially expected sentences with implausible contexts to result in a semantic violation that would be captured by the N400, relative to sentences with plausible contexts. Our results did not support this expectation. There may be several reasons for this. First, our stimuli did not include traditional semantic violation phrases that are used elsewhere in the N400 literature (e.g., *He spread the warm bread with socks*, Kutas & Hillyard, 1984). The process of reading an otherwise well-formed phrase in an implausible context (e.g., “the lion” in the context of a restaurant) may not map directly onto the process of reading semantic violation phrases, which typically are incoherent. Our implausible texts are not incoherent like the traditionally studied ones; they are merely implausible, and this distinction might be relevant to the N400 component. Second, unlike our instructions, the stops-making-sense task used in Singh et al. (2016) explicitly required participants to make judgments about nonsense. This may have led participants to pay greater attention to coherence and sensibility than our instructions. Third, it is possible that there is a lag between the time at which the brain first detects implausibility/incoherence and the time at which our minds become consciously aware of this, and the N400 may be sensitive to the first and not necessarily the second.

Our results thus sharpen Schumacher’s finding that the LAN appears to be associated with failure to find an appropriate antecedent when triggered by a uniqueness presupposition, i.e., *the*. By removing ‘given’ conditions, and the need to compare context updates, we have isolated the cognitive cost associated with technical presupposition failure and accommodation: the brain registers it as a LAN effect. If this is correct, we would expect the LAN to show up in other environments that require accommodation but which are not odd. For example, *The psychology department is facing a crisis. Both of their neuroscientists left* should elicit a LAN relative to *The psychology department is facing a crisis. All of their neuroscientists left*.

References

- Abbott, B. (2006). *Unaccommodating presuppositions: A neogrician view*. (Manuscript, Michigan State University)
- Anderson, J., & Holcomb, P. (2005). An electrophysiological investigation of the effects of coreference on word repetition and synonymy. *Brain and Language*, 94(2), 200-216.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.
- Beaver, D. (2001). *Presupposition and assertion in dynamic semantics*. Stanford, CA: CSLI.
- Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related

- brain potentials. *Brain and Language*, 98(2), 159–168.
- Clark, H. H. (1975). Bridging. In R. Schank & B. Nash-Webber (Eds.), *Theoretical issues in natural language processing* (p. 169-174). New York, NY: Association for Computing Machinery.
- Crain, S., & Steedman, M. (1985). On not being led up the garden path: the use of context by the psychological syntax processor. In D. R. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives* (p. 320-358). Cambridge: Cambridge University Press.
- 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(1), 9–21.
- Gauker, C. (1998). What is a context of utterance? *Philosophical Studies*, 91, 149-172.
- Gazdar, G. (1979). *Pragmatics*. New York, NY: Academic Press.
- Gibson, E., & Perlmutter, N. (1998). Constraints on sentence comprehension. *Trends in Cognitive Science*, 2, 262-268.
- Haviland, S., & 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.
- Heim, I. (1982). *On the semantics of definite and indefinite noun phrases*. (Doctoral Dissertation, University of Massachusetts at Amherst)
- Heim, I. (1983). On the projection problem for presuppositions. In *WCCFL 2* (p. 114-125).
- Heim, I. (1991). Artikel und definitheit. In *Semantics: An international handbook of contemporary research*. Berlin: de Gruyter.
- Heim, I. (1992). Presupposition projection and the semantics of attitude verbs. *Journal of Semantics*, 9, 183-221.
- Hirotsu, M., & Schumacher, P. B. (2011). Context and topic marking affect distinct processes during discourse comprehension in Japanese. *Journal of Neurolinguistics*, 24(3), 276–292.
- Karttunen, L. (1974). Presupposition and linguistic context. *Theoretical Linguistics*, 1, 181-193.
- King, J. W., & Kutas, M. (1995). Who did what and when? using word-and clause-level ERPs to monitor working memory usage in reading. *Journal of cognitive neuroscience*, 7(3), 376-395.
- 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(9), 1147-1162.
- Kutas, M., & Federmeier, K. (2007). Event-related brain potential (ERP) studies of sentence processing. In G. Gaskell (Ed.), *Oxford handbook of psycholinguistics*. Oxford: Oxford University Press.
- Kutas, M., & Hillyard, S. A. (1984). Event-related brain potentials (ERPs) elicited by novel stimuli during sentence processing. *Annals of the New York Academy of Sciences*, 425(1), 236–241.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, 82(13).
- Lenth, R. (2018). Emmeans: Estimated marginal means, aka least-squares means. *R Package Version*, 1(2).
- Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, 1, 339-359.
- Mullen, T. (2012). *NITRC: Cleanline: Tool/resource info*.
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of memory and language*, 31(6), 785-806.
- R Core Team. (2013). *R: A language and environment for statistical computing* [Computer software manual]. Vienna, Austria.
- Schumacher, P. B. (2009). Definiteness marking shows late effects during discourse processing: evidence from ERPs. In *Discourse anaphora and anaphor resolution colloquium* (pp. 91–106).
- Schwarz, F. (2007). Processing presupposed content. *Journal of Semantics*, 24, 373-416.
- Singh, R. (2011). *Maximize Presupposition!* and local contexts. *Natural Language Semantics*, 19, 149-168.
- Singh, R., Fedorenko, E., Mahowald, K., & Gibson, E. (2016). Accommodating presuppositions is inappropriate in implausible contexts. *Cognitive Science*, 40, 607-634. (DOI: 10.1111/cogs.12260)
- Soames, S. (1989). Presupposition. In D. Gabbay & F. Guenther (Eds.), *Handbook of philosophical logic, vol. iv*. Dordrecht: Reidel.
- Stalnaker, R. (1974). Pragmatic presuppositions. In M. Munitz & P. Unger (Eds.), *Semantics and philosophy*. New York, NY: NYU Press.
- Stalnaker, R. (2002). Common ground. *Linguistics and Philosophy*, 25, 701-721.
- Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, 33, 285-318.
- Van der Sandt, R. (1992). Presupposition projection as anaphora resolution. *Journal of Semantics*, 9, 333-377.
- Von Stechow, K. (2008). What is presupposition accommodation, again? *Philosophical Perspectives*, 22, 137-170.
- Winkler, I., Haufe, S., & Tangermann, M. (2011). Automatic classification of artifactual ICA-components for artifact removal in EEG signals. *Behavioral and Brain Functions*, 7(1), 30.