The effect of local context growth and global lexical antecedent frequency on the EEG response to anaphoric pronouns during the comprehension of auditory stories.

Ingmar Brilmayer[1], Ina Bornkessel-Schlesewsky[2], and Matthias Schlesewsky[2]

[1]Institute of German Language and Literature I, University of Cologne, Cologne, Germany
[2]School of Psychology, Social Work & Social Policy, University of South Australia, Adelaide, Australia

Definite referents are usually assumed to be in the focus of attention during discourse comprehension (e.g. Ariel, 2013; Gundel & Fretheim, 2008; Gildea, 2012; Grosz, Weinstein, & Joshi, 1995), and recent evidence points to a critical involvement of definiteness in the prediction of future discourse topics (e.g. Von Heusinger & Chiriacescu, 2013). However, since most experimental research has been carried out with rather small contexts, it remains largely unclear how dynamic aspects of natural language use influence the discourse predictive potential of definiteness marking. For instance, how does the global information provided by lexical frequency influence the accessibility of discourse referents, and how does this global information behave in a constantly growing, local discourse context? Here, we present results from an event-related potential (ERP) study on pronoun resolution which tested the interaction of local and global information using auditory short stories. To foreshadow, the results, on the one hand, support the predictive function of definiteness marking as assumed in the literature, but, on the other, call for an incorporation of more dynamic aspects into linguistic theories of reference.

We embedded sentence pairs into auditory short stories (N = 20, ~2 min duration). The first sentence of each pair introduced a new referent with either an indefinite or a definite subject noun phrase (e.g. 'the mayor' or 'a mayor'). These referents were re-mentioned by a personal pronoun with unambiguous reference (definite/indefinite pronouns) in the second sentence. Each story contained 2 pairs with indefinite and 2 pairs with definite antecedent. We varied the position of the sentence pair (*index*; local context) on a continuous scale across stories, as well as lexical frequency (*frequency*; global context) of the antecedent. Participants (N=40) listened to the short stories while we recorded their EEG. For the anaphoric pronouns, we computed and statistically assessed event-related potentials (ERPs). Using linear mixed-effect modeling, we assessed the influence of an antecedent's lexical frequency and its position in a text on the predictive potential of definite and indefinite marked subjects as reflected in the ERPs between 250-350 ms and 350-500 ms following the onset of the pronoun.

Our results reveal an early negativity for indefinite pronouns (200-350 ms, fig. 1a), reflecting a prediction mismatch response (Friston, 2005) upon the detection of a co-referential pronoun, as well as a fronto-centrally distributed positivity (350-500 ms, fig. 1b) for indefinite pronouns, which we assume to reflect attentional reorienting towards the unexpected referent. However, we found that the ERPs to definite and indefinite pronouns are differentially and critically affected by *index* and *frequency*: lower frequency elicits a graded positive effect between 200-350 ms in both conditions (fig. 1c), i.e. lower frequency increases the predictability of discourse referents, thereby acting similar to a linguistic prominence cue. In the later time window, the ERPs at anterior electrodes, corresponding to the fronto-central positivity in the grand averaged ERPs, show a classic frequency effect in the indefinite condition (stronger negativity for low frequent words), but not in the definite condition (fig. 1d). Furthermore, in the later time window only the ERPs following definite pronouns are significantly affected by *index*, while the responses following indefinites are stable throughout the course of a story (fig. 1e).

To sum up, the present results generally support the assumption that definiteness marking has discourse predictive function and that definite referents are currently in the focus of attention. Yet, they also suggest that definite and indefinite marking differ in their dependence on local and global context. While the ERP effects for definite pronouns change throughout the course of a story as a function of index (local feature) and frequency class (global feature), they remain relatively stable for indefinite pronouns. This clearly calls for more dynamic notions of definiteness in linguistic theorizing, in combining functional linguistic accounts with accounts of attention and prediction borrowed from neuroscientific knowledge about dynamicity in the processing systems of the human brain (e.g. Friston, 2005; Hasson, Chen, & Honey, 2015).

References

Ariel, M. (2013). Centering, accessibility and the next mention. *Theoretical Linguistics*, *39*(1–2), 39–58.

Friston, K. (2005). A theory of cortical responses. *Philosophical transactions of the Royal Society B: Biological sciences*, *360*(1456), 815–836.

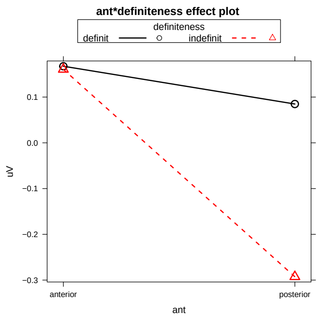
Gildea, S. (2012). The referential hierarchy and attention. *Faits de Langues*, *38*, 33–47.

Grosz, B. J., Weinstein, S., & Joshi, A. K. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational linguistics*, *21*(2), 203–225.

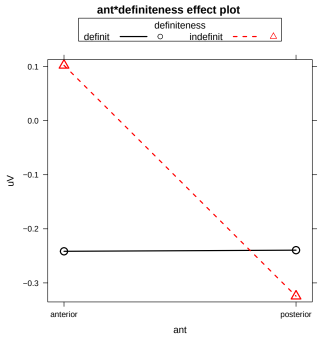
Gundel, J. K., & Fretheim, T. (2008). Topic and focus. In L. Horn & G. Ward (Eds.), *The handbook of pragmatics* (Vol. 26). John Wiley & Sons.

Hasson, U., Chen, J., & Honey, C. J. (2015). Hierarchical process memory: memory as an integral component of information processing. *Trends in cognitive sciences*, *19*(6), 304–313.

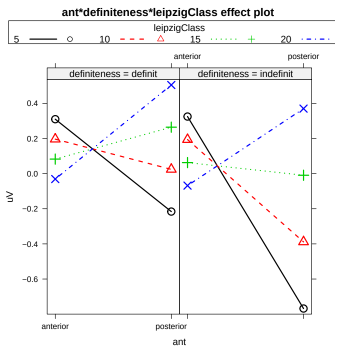
Von Heusinger, K., & Chiriacescu, S. (2013). The discourse structuring potential of differential object marking, the case of indefinite and definite direct objects in Romanian. *Revue Roumaine de Linguistique–Romanian Review of Linguistics*, *58*(4), 439–456.



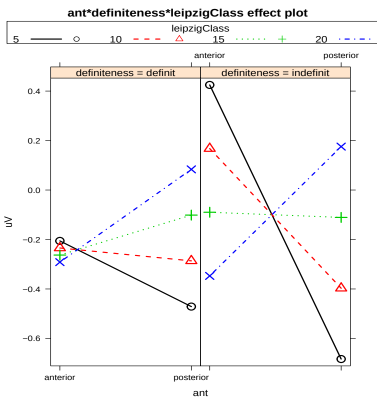
(a) Interaction of anteriority and definiteness between 200–350 ms.



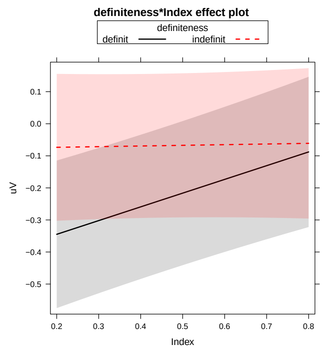
(b) Interaction of anteriority and definiteness between 350–500 ms.



(c) Interaction of anteriority, definiteness and frequency class between 200–350 ms.



(d) Interaction of anteriority, definiteness and frequency class between 300–500 ms.



(e) Interaction of definiteness and index between 300–500 ms.

Incremental & predictive pragmatic interpretation: A rational analysis

Michael Franke & Petra Augurzky (Tübingen)

The Rational Speech Act (RSA) model [e.g. 1] is a recently popular approach that models Gricean pragmatic inference probabilistically. While applications have so far only focused on interpretation of whole utterances, we here explore an extension of RSA that yields predictions about incremental processing. We assess the predictions of incremental RSA for data from an EEG-study on scalar implicature processing.

Study. 25 native speakers of German took part in the experiment. Each trial presented one out of four picture types (see Figure 1), followed by a sentence of type (1) presented word-by-word. After each sentence participants gave a truth-value judgement. ERPs were measured at the three positions indicated in (1): the quantifier, the color adjective and the shape noun in the postfixed relative clause. While the experiment's design was motivated by previous investigations [e.g. 3, 4] to test hypotheses about processing costs associated with scalar implicature calculation, suppression or cancellation in context, we here focus on a novel approach of probabilistic modeling of the previously collected data.

(1) Alle/Einige$_1$ Punkte sind blau/rot$_2$, die im Kreis/Quadrat$_3$ sind
all/some dots are blue/red, that in the circle/square are
All/some of the dots in the circle/square are blue/red

Model. To capture pragmatic interpretation, the RSA model assigns probability $P_L(t \mid m) \propto P(t)\, P_S(m \mid t)$ to meaning $t \in T$ given message $m \in M$ via Bayes rule by combining prior $P(t)$ and likelihood $P_S(m \mid t)$. The likelihood $P_S(m \mid t)$ usually implements the standard Gricean assumptions concerning proper speaker behavior, e.g., $P_S(m \mid t) \propto \exp(\lambda\, P(t \mid [\![m]\!])$, where λ is a soft-max parameter and $[\![m]\!] \subseteq T$ is m's semantic meaning. RSA's interpretation rule assigns an interpretation t to a whole utterance m. The following captures a rational interpreter's beliefs during incremental processing. Let utterance $m = w_1, \dots, w_n$ be a sequence of words. Define $m_{\rightarrow i} = w_1, \dots w_i$ as the initial subpart of m up to and including word w_i. Write $m_{\rightarrow i}, w_{i+1}$ for the continuation of $m_{\rightarrow i}$ with word w_{i+1}. Let $M(m_{\rightarrow i}) = \left\{m' \in M \mid m'_{\rightarrow i} = m_{\rightarrow i}\right\}$ be the subset of M which have the same initial subpart up to word w_i as m. A rational pragmatic interpreter has beliefs about the likely next word w_{i+1}:

$$P_L(w_{i+1} \mid m_{\rightarrow i}) \propto \sum_t P(t) \sum_{m' \in M(m_{\rightarrow i}, w_{i+1})} P_S(m' \mid t) \qquad [1]$$

Similar measures of next-word probability, e.g., derived from corpus frequencies, predict quantitative patterns in reading times related to syntactic parsing or N400 amplitudes associated with lexical-semantic processing [e.g. 5, 6]. Our approach instead quantifies expectations of *pragmatic adequacy in context* beyond syntax or lexeme associations for sentences of identical syntactic complexity and lexical cohesion.

Fit. The above model has one free parameter λ and requires specification of T and M. We here assume that M contains the sentences in (1), and that T is the set of all picture types from Figure 2 paired with the shape in each picture that the speaker wants to talk about. The listener knows the context but not the shape within each context that the speaker will talk about. With these assumptions, the incremental RSA model predicts lexical expectations for each recording position in (1). Plotted below are model predictions against observed grand-average N400 amplitude in an early time window (300-400ms) following Augurzky et al. [7] for a hand-picked favorable value of $\lambda = 7$. There appears to be a general positive linear relationship between the model's predicted next-word probability and the N400 amplitude ($r \approx 0.44$, $p < 0.01$ for all data points; see Figure 2). Interestingly, a cluster of data points are not well predicted by the model: the model predicts an equal expectation for shape nouns in the relative clause when the preceding sentence gives enough information for a truth value judgement already (e.g., hearing "All dots are blue ..." in context C). ERPs on functionally redundant relative clauses have the largest N400 amplitude, but the model predicts .5 expectability. Leaving out data points from shape nouns in functionally redundant relative clauses, the model's predictions are excellent ($r \approx 0.81$, $p < 0.001$; see gray line in Figure 2). We hypothesize that the fault in the model may be the assumption that M does not contain sentences which stop before the relative clause. Further work is necessary to investigate the relation between a pragmatically motivated next-word expectation and the ERP response. Still, our findings indicate that incremental RSA may turn out to be a promising alternative to standard interpretations of ERP data on pragmatic processing, since it makes quantitative predictions and resounds well with successful modeling elsewhere in psycholinguistics.

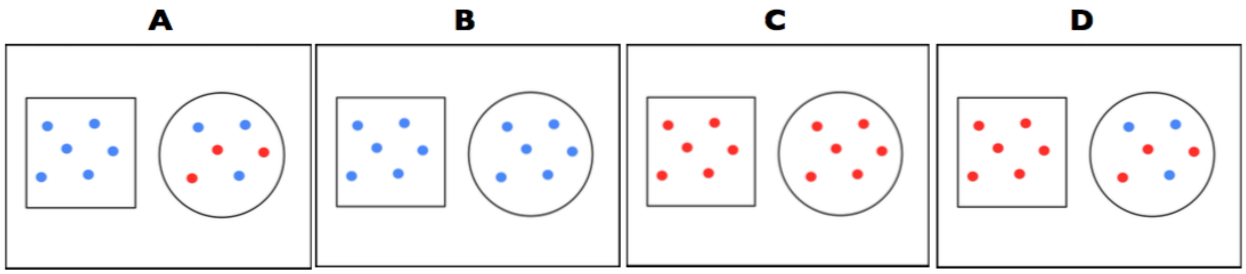


Figure 1: Examples of picture material.

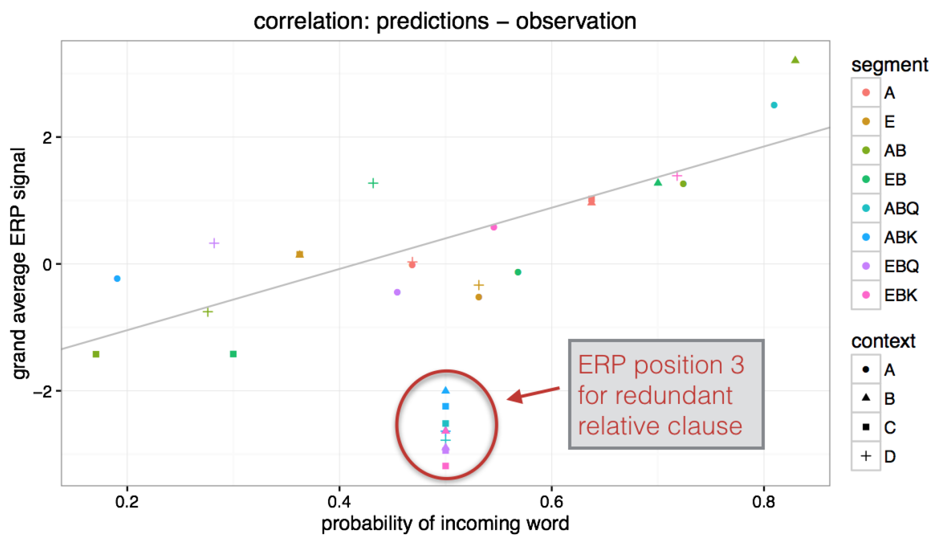


Figure 2: Model predcitions (probability of next word) plotted against observations (grand-average early N400).

References

[1] Noah D. Goodman and Michael C. Frank. "Pragmatic Language Interpretation as Probabilistic Inference". In: *Trends in Cognitive Sciences* 20.11 (2016), pp. 818–829.

[2] Noah D. Goodman and Andreas Stuhlmüller. "Knowledge and Implicature: Modeling Lanuage Understanding as Social Cognition". In: *Topics in Cognitive Science* 5 (2013), pp. 173–184.

[3] Lamar Hunt III et al. "Pragmatic inferences modulate N400 during sentence comprehension: Evidence from picture–sentence verification". In: *Neuroscience Letters* 534 (2013), pp. 246–251.

[4] Maria Spychalska, Jarmo Kontinen, and Markus Werning. "Investigating scalar implicatures in a truth-value judgement task: evidence from event-related brain potentials". In: *Language, Cognition and Neuroscience* 31.6 (2016), pp. 817–840.

[5] Roger Levy. "Expectation-Based Syntactic Comprehension". In: *Cognition* 106 (2008), pp. 1126–1177.

[6] Stefan L. Frank et al. "The ERP response to the amount of information conveyed by words in sentences". In: *Brain and Language* 140 (2015), pp. 1–11.

[7] Petra Augurzky et al. "Are all the triangles blue? — ERP evidence for the incremental processing of German quantifier restriction". In: *Language and Cognition* (online first), pp. 1–34.

Viviana Haase, Institute for Philosophy II, Ruhr University Bochum: An alternative approach to the comprehension of negated sentences: evidence from event-related potentials

Negative sentences haven been shown to elicit higher error rates and longer response times than their affirmative counterparts (e.g. Just & Carpenter, 1971). Accordingly, it has been argued that negative sentences require more cognitive resources than affirmative sentences due to (i) their higher (morpho-)syntactical complexity and (ii) the need to suppress positive information and to eventually represent the negated state of affairs (i.e. the affirmative counterpart) on a first step before representing the actual state of affairs (e.g. Just & Carpenter, 1971; Kaup et al, 2006). The latter argument seems to find further support from the observations that true negative sentences seem to be harder to process than false negatives, while the pattern is reversed for affirmative sentences (Just & Carpenter, 1971; Kaup et al., 2006; Lüdtke et al. 2008). In line with this observation, an early event-related potential (ERP) study by Fischler et al. (1984) showed that true negatives elicited larger N400 amplitudes than false negative sentences did. However, this effect disappears as soon as the negative sentence is pragmatically licensed (Nieuwland & Kuperberg, 2008) or as soon as no truth-value judgement task is employed, i.e. as soon as subjects are not asked to explicitly evaluate the sentences as true or false (Wiswede et al., 2013).

My talk features two goals. First I would like to present data from an ERP-study investigating the time-course of the comprehension process underlying negated sentences, thereby addressing the questions (i) whether negation can be processed incrementally or whether a multistep process is necessary and (ii) whether processing differences between negative and affirmative sentences are correlated with memory capacities. Since the experiment did not employ any explicit sentence-picture verification or truth evaluation, any resulting effects should be related to the mere comprehension of the sentences, i.e. to the combinatorial integration of various sorts of linguistic information such as for example lexical meaning and structural dependencies. Furthermore, other than in Fischler et al. (1984), the subject and the object of the sentence material used in this study are both from related semantic fields in order to increase their plausibility.

Second, I would like to present a number of arguments on why we should shed light on the theoretical aspects of negation comprehension taking into account a broader perspective. I will argue that (i) in order to attempt an understanding of the combinatorial processes underlying the comprehension of negated sentences we should implement the aspect of negation comprehension into neurobiologically plausible models of language and (ii) we have to take into account the various forms and functions of negations rather than predicting a universal processing mechanism for negation in general. Therefore, we should take into account language-specific differences, as different languages vary with regard to their (default) relative weighting of top-down and bottom-up information sources (e.g. MacWhinney et al 1984, Tune et al 2014). Given the diversity of results of different studies investigating various forms of negation in different languages with the help of various experimental methods, the interpretation of these data and the resulting theoretical implications need to be revised.

In an attempt to do so, I combine two rather recent neurolinguistic and neurocomputational approaches: (1) Schumacher & Hung (2012) and Schumacher (2014) predict the N400 amplitude to be gradually dependent to the information status of an entity while processing demands arising from inferencing and discourse representation are assumed to be reflected in a late positivity. (2) Instead, Brouwer & Hoeks (2013) predict the N400 to reflect the ease of retrieval of conceptual information from memory while the P600 is assumed to reflect the integration of the retrieved meaning with the word's preceding context. I will argue that it is the retrieval from memory that results in a N400 effect for negative compared to affirmative information, while it is an entities information status that might lead to the observed differences between true and false negative sentences. Furthermore, depending on the context and task, negative sentences may elicit a late positivity when compared to their affirmative counterparts, reflecting some sort of pragmatic integration and pragmatic enrichment.

Literature:

Brouwer, H. and Hoeks, C. (2013). A time and place for language comprehension: mapping the n400 and the p600 to a minimal cortical network. *Frontiers of Human Neuroscience*, 7(758).

Fischler, I., Bloom, P.A., Childer, D.G., Arroyo, A.A., Perry, N.W. (1984): Brain potentials related to stages of sentence verification. *Neuropsychologia* 22(5), 559-568.

Just, M. A., & Carpenter, P. A. (1971). Comprehension of negation with quantification. *Journal of verbal Learning and Verbal Behavior*, 10, 244-253

Kaup, B., Lüdtke, J., & Zwaan, R. A. (2006). Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed? *Journal of Pragmatics*, 38(7), 1033-1050.

MacWhinney B, Bates E, Kliegl R. (1984): Cue validity and sentence interpretation in English, German and Italian, *Journal of Verbal Learning and Verbal Behavior*;23:127–150.

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

Schumacher, P. B. (2014). Content and context in incremental processing: "the ham sandwich" revisisted. *Philosophical Studies*, 168(151- 165).

Schumacher, P. B. and Hung, Y. (2012). Positional influences on information packaging: insights from topological fields in german. *Memory and Language*, 67(2):295{310.

Tian, Y., Ferguson, H. & Breheny, R. (2016): Processing negation without context - why and when we represent the positive argument. *Language, Cognition and Neuroscience,, 31*, 683-698

Wiswede, D., Koranyi, N. Müller, F., Langner, O., Rothermund, K. (2013): Validating the truth of propositions: behavioral and ERP indicators of truth evaluation processes. *Social Cognitive and Affective Neuroscience* 8. 647-653.

Bilingual novel and conventional metaphor comprehension: the N400 and LPC

Katarzyna Jankowiak[1,2], Karolina Rataj[1], Ryszard Naskręcki[3]

[1]Faculty of English, Adam Mickiewicz University in Poznań, Poland
[2]NanoBioMedical Centre, Adam Mickiewicz University in Poznań, Poland
[3]Laboratory of Vision Science and Optometry, Faculty of Physics, Adam Mickiewicz University in Poznań, Poland

Previous research into novel and conventional metaphor comprehension has repeatedly shown more pronounced N400 amplitudes for novel than conventional metaphors (Lai et al. 2009; Arzouan et al. 2007). Within the LPC time frame, conventional metaphors have elicited larger LPC amplitudes relative to anomalous utterances (De Grauwe et al. 2010; Arzouan et al. 2007), while novel metaphors have evoked an attenuated LPC response (Arzouan et al. 2007), thus suggesting ongoing difficulty in meaning processing or access to the non-literal route during novel metaphor comprehension. Importantly, ERP research into novel and conventional metaphor that has so far been conducted has been restricted to the monolingual context. The current study was therefore aimed at testing how bilingual speakers compute novel and conventional metaphors in their native (L1) and non-native (L2) tongue. 23 Polish (L1) – English (L2) bilinguals performed a semantic decision task to Polish and English novel metaphoric, conventional metaphoric, literal, and anomalous word pairs. ERP findings showed a language-independent effect of utterance type within the late N400 time window (400-500 ms), where a graded effect was observed for both languages, with most pronounced N400 amplitudes in response to anomalous, followed by novel metaphoric, conventional metaphoric, and finally literal word dyads. Interestingly, within the LPC time frame (500-800 ms), an interaction between language and utterance type was found. In the native tongue, novel metaphors evoked attenuated LPC amplitudes relative to conventional metaphoric and literal word pairs. In the non-native language, on the other hand, a reduced LPC response was found to not only novel but also conventional metaphoric word pairs. Such findings might be interpreted as indicative of similar mechanisms engaged in lexico-semantic processing during metaphor comprehension in both languages, as indexed by the N400, and more effortful processes involved in novel metaphor in L1, and in both novel and conventional metaphor in L2, as reflected in LPC patterns. The observed findings will be discussed with reference to the Career of Metaphor Model (Bowdle and Gentner 2005) as well as the Bilingual Interactive Activation Plus Model (BIA+; Dijkstra and van Heuven 2002).

References:

Arzouan, Yossi, Abraham Goldstein and Miriam Faust. 2007. “Brainwaves are stethoscopes: ERP correlates of novel metaphor comprehension”, *Brain Research* 1160: 69–81.

Bowdle, Brian F. and Dedre Gentner. 2005. “The career of metaphor”, *Psychological Review* 112, 1: 193–216.

De Grauwe, Sophie, Abigail Swain, Phillip J. Holcomb, Tali Ditman and Gina R. Kuperberg. 2010. “Electrophysiological insights into the processing of nominal metaphors”, *Neuropsychologia* 48, 7: 1965–1984.

Dijkstra, Ton and Walter J.B. van Heuven. 2002. “The architecture of the bilingual word recognition system: From identification to decision”, *Bilingualism: Language and Cognition* 5: 175–197.

Lai, Vicky Tzuyin, Tim Curran and Lise Menn. 2009. “Comprehending conventional and novel metaphors: An ERP study”, *Brain Research* 1284: 145–155.

Brain signature of planning for production: An EEG study

Suzanne R. Jongman[1], Antje S. Meyer[1], & Vitória Piai[2,3]
[1]Max Planck Institute for Psycholinguistics, Nijmegen, [2]Radboud University, Donders Centre for Cognition, Nijmegen, [3]Radboudumc, Department of Medical Psychology, Nijmegen

In conversation, turns from one speaker to the next happen in rapid succession. It has been suggested that planning for production happens simultaneously with listening to keep gaps between turns minimal (Levinson & Torreira, 2015). There is some experimental evidence corroborating this hypothesis. Both Bögels, Magyari, and Levinson (2015), using a quiz study, and Barthel, Sauppe, Levinson, & Meyer (2016), using a list-completion paradigm, found shorter naming latencies for trials where the response could be known early in the sentence than late. Levinson and Torreira argued that planning happens as soon as there is enough information and the phonological code is then held in working memory till the turn end. However, whether planning actually proceeds all the way to phonological encoding has not been proven. EEG might be able to provide us with an answer to this question. In the study by Bögels et al., EEG was measured in addition to naming latencies. The authors found a late positivity starting 500ms after the onset of the information that allowed participants to prepare their answer. They localized this effect to the middle and superior temporal gyrus and the inferior frontal gyrus, areas previously found to be involved in language production (Indefrey & Levelt, 2004). The authors concluded that planning progressed to phonological encoding.

We believe that in order to investigate what planning for production entails, we need a bottom-up approach starting with a very simple paradigm to finally result in an experiment that simulates natural conversation. In the quiz paradigm planning for production co-occurred with language comprehension, making it extremely difficult to disentangle the two processes. In the current study we solely tested what happens when a participant must withhold speech until a go signal is presented, using a paradigm developed by Piai, Roelofs, Rommers, Dahlslätt, and Maris (2015). Participants were presented with five non-words in a row, with the fourth or fifth being pronounceable (see Figure 1). They were instructed to read aloud the pronounceable non-word, but only after the fifth non-word was shown. Thus, speech had to be withheld or speech was immediate. The 800ms time-window between the fourth and fifth non-word was compared between the withhold and immediate condition using cluster-based permutation tests.

The ERP results indicated three significant clusters in posterior regions, with a larger positivity for the withhold condition than for immediate speech (Figure 2). In the power analysis, two significant clusters were found for the frequency range 15 to 27Hz in posterior regions, indicating that beta power over posterior regions was reduced for the withhold condition (Figure 3). As in Bögels et al., we find a positivity for the withhold condition, but it seems to be localized in a different region. Our results indicate that withholding speech causes a beta decrease over posterior regions. This posterior beta was found previously (Piai et al., 2015), and could reflect attention or working memory processes. When we know what we should say and can plan it to the phonological level (as is the case with pronounceable non-words) but we must wait to pronounce the phonological code, posterior beta power is modulated. The next step is to add a level of complexity to the task. This paradigm can for instance be used to find out what happens when you have to generate your own message instead of just repeating what is just presented? And what happens when this co-occurs with heard speech? By adding a piece of the puzzle one by one it will become clear what exactly happens when we comprehend and produce simultaneously, just as in conversation.

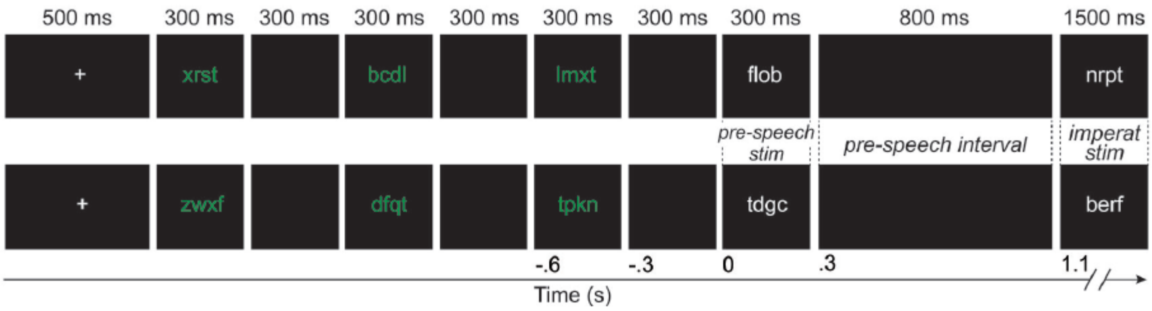


Figure 1. An example of a withhold trial (top panel) and an immediate trial (bottom panel).

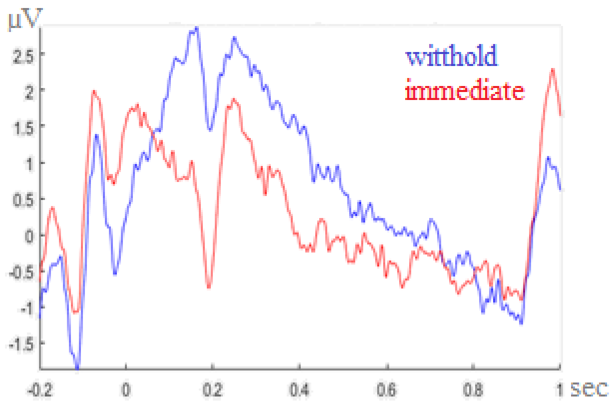


Figure 2. Grand average event-related potentials for the two conditions averaged over all posterior electrodes. Positivity plotted upwards.

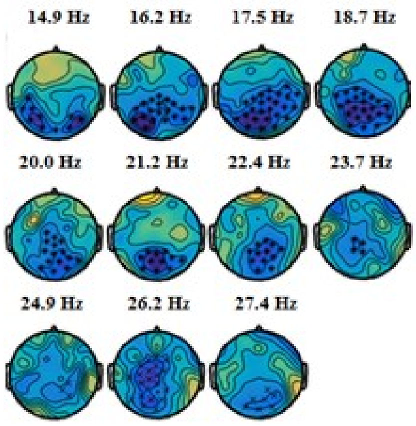


Figure 3. Group-level topographical maps of the relative power change.

References

Barthel, M., Sauppe, S., Levinson, S. C., & Meyer, A. S. (2016). The timing of utterance planning in task-oriented dialogue: Evidence from a novel list-completion paradigm. *Frontiers in Psychology, 7*.

Bögels, S., Magyari, L., & Levinson, S. C. (2015). Neural signatures of response planning occur midway through an incoming question in conversation. *Scientific reports, 5*.

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*(1–2), 101-144.

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in psychology, 6*.

Piai, V., Roelofs, A., Rommers, J., Dahlslätt, K., & Maris, E. (2015). Withholding planned speech is reflected in synchronized beta-band oscillations. *Frontiers in human neuroscience, 9*.

Electrophysiological signatures of negative and positive polarity processing in German sentence comprehension

Mingya Liu, Peter König and Jutta L. Mueller (University of Osnabrück)

I. Introduction The lexical inventory of many languages consists of expressions that are sensitive to the polarity of context. Negative polarity items (NPIs) such as German *jemals* 'ever' tend only to occur in negative contexts (1a), whereas positive polarity items (PPIs) such as German *schon* 'already' tend only to occur in positive contexts (1b). Polarity sensitivity has been a key research field in generative linguistics, as it is revealing w.r.t. the internal structure of language, i.e. how different aspects of grammar (syntax, semantics, pragmatics) interact with one another [1,2].

(1) a. Peter hat {**keinen Kuchen** / #**den Kuchen**} <u>jemals</u> oft gebacken.
b. Peter hat {#**keinen Kuchen** / **den Kuchen**} <u>schon</u> oft gebacken.

One unresolved related question in theoretical linguistics concerns whether and to what extent NPIs/PPIs are parallel, i.e., whether their requirements on context and violations of these are of a similar syntactic/semantic nature [1,2,6]. In terms of processing, previous ERP studies report on different ERP (i.e. N400/P600) components elicited by unlicensed NPIs (i.e. NPIs in an affirmative context) and antilicensed PPIs (i.e. PPIs in a negative context) and their results do not match [5,7]. The present ERP study aims at resolving this inconsistency using the German NPI/PPI *jemals*/*schon* and the polarity-insensitive adverb *sehr* 'very' as an additional control item, henceforce nonPI.

II. Method

Materials: We used 120 items in 6 critical conditions (with a 2×3 factorial design), e.g. (2a-f), with two additional control conditions involving semantic or syntactic violations, e.g. (2g-h). The total of 960 sentences are divided into 6 sets, each set containing 320 sentences with 40 ones per condition.

(2) a. #Peter hat **den Kuchen**, der viele Nüsse enthielt, <u>jemals</u> oft gebacken. **(Cond1)**
b. Peter hat **keinen Kuchen**, ... <u>jemals</u> oft gebacken. **(Cond2)**
c. Peter hat **den Kuchen**, ... <u>schon</u> oft gebacken. **(Cond3)**
d. #Peter hat **keinen Kuchen**, ... <u>schon</u> oft gebacken. **(Cond4)**
e. Peter hat **den Kuchen**, ... <u>sehr</u> oft gebacken. **(Cond5)**
f. Peter hat **keinen Kuchen**, ... <u>sehr</u> oft gebacken. **(Cond6)**
g. #Peter hat **keinen Kuchen**, ... sehr oft <u>gelernt</u>. **(Cond7)**
h. #Peter hat **den Kuchen**, ... sehr oft <u>backen</u>. **(Cond8)**

Participants and procedure: 24 native German speakers (12 female; mean age 21.8, SD=2.65) read and rated the naturalness of the stimuli (cf. **Fig. 1**) during the EEG recording with a 64 channel system.
Data pre-processing and analysis: We conducted ANOVAs of the behavioral data by subjects and by items. EEG-recorded data were pre-processed using EEGlab and FASTER [2] for automatic artifact rejection (cf. **Fig. 2**). We performed cluster-based permutation tests for the time windows of N400 (300-500ms) and P600 (500-700ms) in FieldTrip [3].
Results: The behavioral data are by and large in line with the literature (cf. **Table 1**). The ANOVAs show a main effect of context (by subjects and items), a main effect of polarity item (only by items) and an interaction between context and polarity item by subjects and items (statistical details to be presented in the talk). W.r.t. the ERP data, unlicensed NPIs and antilicensed PPIs (Cond1 vs. Cond4) elicited N400 of similar amplitude (**Fig. 4**). Unlicensed NPIs (Cond1) elicited P600 vs. nonPI in affirmation (Cond5), but not vs. licensed NPIs (Cond2), **Fig. 2**; antilicensed PPIs (Cond4) elicited P600 vs. licensed PPIs (Cond3) but not vs. nonPI in negation (Cond6), cf. **Fig. 3**; unlicensed NPIs (Cond1) elicited P600 vs. antilicensed PPIs (Cond4), **Fig. 4**. The comparison between nonPIs in negation vs. affirmation did not yield significant effects (statistics and graphs to be presented in the talk).
III. Discussion and conclusion: Overall, our study shows that NPIs and PPIs are not entirely parallel. The N400 results indicate that the online semantic processes of NPIs/PPIs are similar (in line with the result of [5], contra [7]); this confirms semantic (or pragmatic) theories of NPIs/PPIs such as [1]. The P600 results (contra [5,7]) indicate that the online syntactic processes of NPIs/PPIs are different; this casts doubt on syntactic theories that treat NPIs/PPIs as parallel (cf. [2,6]).

I. Procedure

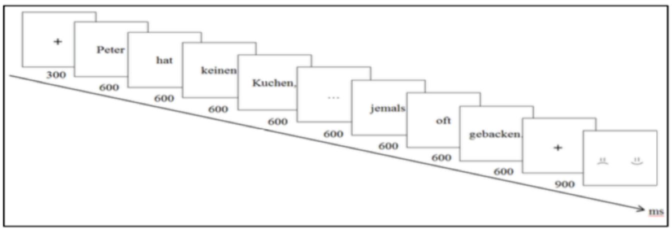


Figure 1. Time sequence of the stimuli presentation with an open-end response period

II. Behavioral data

Condition	Mean	SD
Cond1	.20	.15
Cond2	.79	.20
Cond3	.94	.07
Cond4	.63	.28
Cond5	.95	.07
Cond6	.84	.17
Cond7	.08	.20
Cond8	.03	.03

Table 1. Descriptive statistics (0=unnatural, 1=natural)

III. ERP data A. N400 (p=0.002) but no P600 (p=0.062) for unlicensed NPI (Cond1, red line) vs. licensed NPI (Cond2, blue line); N400 (p=0.032) and P600 (p=0.04) for unlicensed NPI (Cond1, red line) vs. nonPI in affirmation (Cond5, green line)

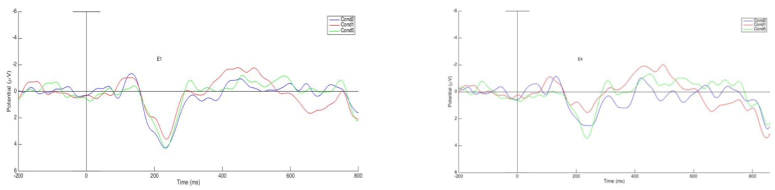


Figure 2: Grand averaged ERPs time-locked to the onset of NPI/nonPI including a baseline of 200 ms at two representative electrodes E1 and E4.

ERP data B. N400 (p=0.026) and P600 (p=0.044) for antilicensed PPI (Cond4, red line) vs. licensed PPI (Cond3, blue line); N400 (p=0.038), but no P600 (p>0.05) for antilicensed PPI (Cond4, red line) vs. nonPI in negation (Cond 6, green line)

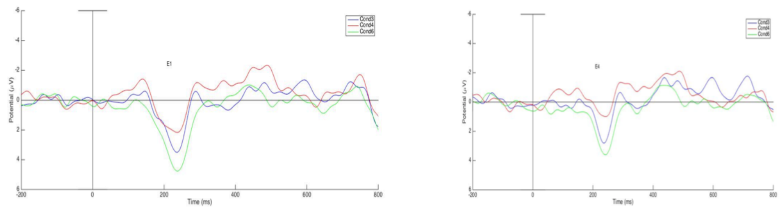


Figure 4: Grand averaged ERPs time-locked to the onset of PPI/nonPI including a baseline of 200 ms at E1 and E4.

ERP data C. No N400 (p=1) but P600 (p=0.018) for unlicensed NPI (Cond1, blue line) vs. antilicensed PPI (Cond4, red line)

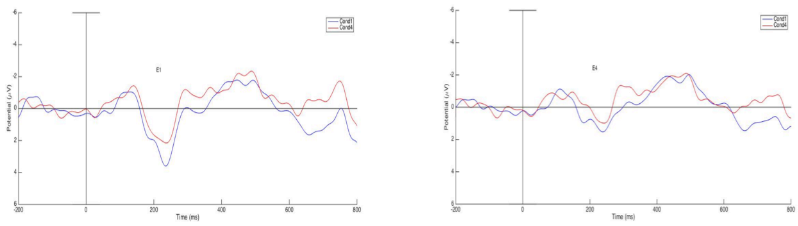


Figure 4: Grand averaged ERPs time-locked to the onset of NPI/PPI including a baseline of 200 ms at E1 and E4.

Selected references: [1] Giannakidou, A. (2011). Negative and positive polarity items. In: von Heusinger, K., Maienborn, C., & Portner, P. (Eds.). *Semantics* (HSK 33.2). Berlin, Boston: De Gruyter Mouton, 1660–1712. [2] Klima, E. S. (1964). Negation in English. In: Fodor, J. & Katz, J. (Eds.). *The structure of language*. New Jersey: Prentice Hall, 246-323. [3] Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: fully automated statistical thresholding for EEG artifact rejection. *Journal of neuroscience methods*, *192*(1), 152-162. [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. *Computational intelligence and neuroscience*, *vol. 2011*. [5] Saddy, D., Drenhaus, H., & Frisch, S. (2004). Processing polarity items: Contrastive licensing costs. *Brain and language* 90(1), 495-502. [6] Szabolcsi, A. (2004). Positive polarity – negative polarity. *Natural Language & Linguistic Theory* 22(2), 409-452. [7] Yurchenko, A., den Ouden, D. B., Hoeksema, J., Dragoy, O., Hoeks, J. C. & Stowe, L. A. (2012). Processing polarity: ERP evidence for differences between positive and negative polarity. *Neuropsychologia* 51(1), 132-141.

Early ERP evidence for children's and adult's sensitivity to scalar implicatures

Daniele Panizza (U. of Goettingen), Edgar Onea (U. of Goettingen), Nivedita Mani (U. of Goettingen)

The aim of the present work is to investigate whether and how quickly 3-year-old German-speaking children derive scalar implicatures. Scalar implicatures are pragmatic inferences (cf. Levinson, 2000) that are triggered by scalar items like the English quantifier *some* (i.e. some → some but not all) and the German quantifier *ein paar* (English: *some*, *a few*). Experimental studies demonstrated that children, in contrast to adults, are generally poor at deriving scalar implicatures (Noveck, 2001) until relatively late (i.e. between 7 and 10 years, cf. Katsos & Bishop, 2011). However, previous results can be accounted for by stating that children are indeed sensitive to scalar implicatures but that they are more tolerant regarding pragmatic violations compared to adults (Katsos & Bishop, 2011; Foppolo, Guasti & Chierchia, 2012). These pragmatic violations can be provoked by underinformative sentences given a present context. For instance, children would accept a statement such as (1) in a context where there is a hedgehog that has all of the relevant keys in the scenario.

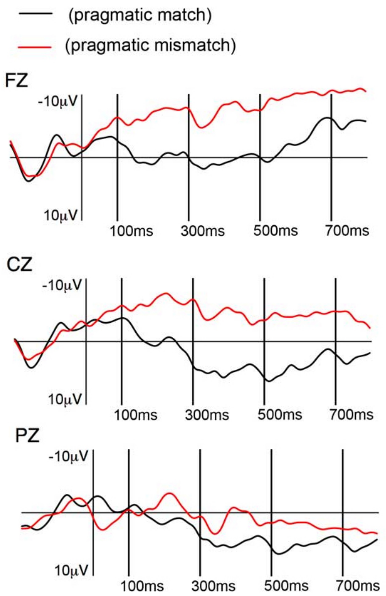
(1) The hedgehog has some of the keys.

The current study evaluates the claim that children possess full competence for deriving scalar inferences at a very young age using the ERP methodology. In particular, we examined the processing of scalar implicatures at a very young age, namely at three-years of age. To test their implicit knowledge, we employed ERP (event-related potentials) measurements because this methodology has been proven to display sensitivity to scalar implicatures (cf. Polizer-Ashles et al., 2013), and it does not require explicit responses from participants, which makes it optimal for testing very young children.

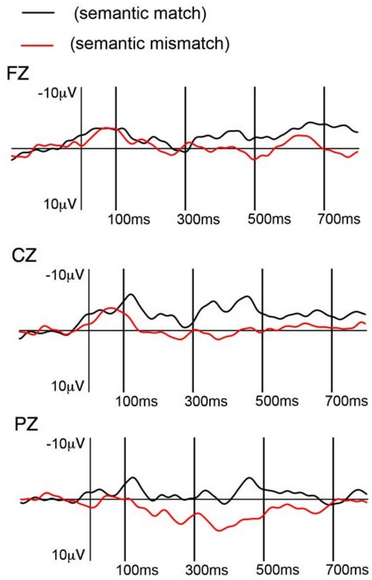
Participants were 24 German-speaking children (35 to 42 months, M = 37.29 months) and 24 adult controls. 20 videos were presented that either displayed a SOME scenario, in which a subset of objects moved from one character to the other, or an ALL scenario, in which all subjects moved. Afterwards, participants were presented auditorily with the critical sentence that was composed of a question like 'Has the hedgehog **all** the keys?' and an answer like 'He had **some** of them'. These question-answer pairs enhance derivation of scalar implicatures by putting the informational focus on the scalar item *ein paar.* Furthermore, it allows us an opportunity to contrast children's responses to semantic versus pragmatic violations. That is, hearing the word *all,* which occurs in the question that precedes the scalar item, when attending a SOME scenario triggers a temporary semantic mismatch, given that the meaning of the quantifier *all* mismatches the information provided by the visual context. In contrast, hearing the word *some* included in the answer when attending an ALL scenario should trigger a pragmatic mismatch. Thus, identical ERP effects elicited by *all* and *some* in the related mismatching conditions would index a violation of the probabilistic (i.e. lexical or phonological) expectation generated by hearing a quantifier incompatible with the visual context. Instead, different ERP profiles would suggest that the two violations are treated differently by the brain, and in particular, that the ERP effects elicited by the pragmatic mismatch are not due to the violation of phonological or lexical predictions.

The results from this study show that *some* in ALL vs. SOME scenarios (pragmatic mismatch) elicited negative brain potential starting as soon as 100 ms after the onset of the scalar term in children and at 300ms in adults (Picture 1). While with children this effect was more pronounced and centrally distributed (i.e. with N400-like topography), in adults it was weaker and more frontally distributed (Picture 3). Instead, hearing *all* in SOME vs. ALL scenarios (semantic mismatch) elicited a general pattern of increased positive activity in both children and adults (see Picture 2 and 4). In adults, however, the positivity was distributed on the left frontal electrodes, while a N400-like wave at about 300 ms was found in the centro-parietal sites (see Picture 5).

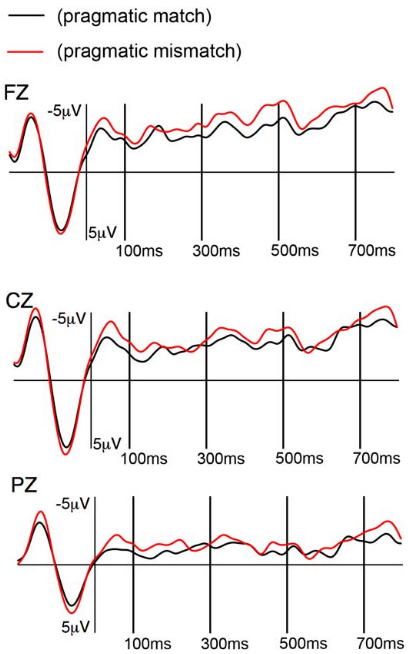
Our results suggest that both children (aged three-years) as well as adults show early and immediate sensitivity to pragmatic and semantic mismatches, despite the fact that the visual stimuli presented in both contexts were physically identical. This suggests that a) both children and adults were sensitive to the interpretation of these terms very early, and the violation thereof, and b) a difference in the reaction to different types of infelicity (i.e. semantic vs. pragmatic mismatch) as previously found in other studies with adults (cf. Politzer-Ahles et al., 2013). Moreover, the fact that our results displayed a rich pattern composed of diametrically different ERPs to *some* vs. *all* in matching and mismatching visual contexts allows us to rule out the hypothesis that our results are due to general task-related strategies, such as merely associating a video with a word, or other underlying cognitive mechanisms, such as lexico-phonological association between words and contexts. Lastly, small but significant differences emerged between the two groups of participants (i.e. children vs. adult) with the adults also displaying a N400-like negativity elicited by *all* in the SOME vs. ALL condition and a positivity with a different distribution as compared to children. Finally, there results support the hypothesis that the poor performance displayed by children at deriving implicatures in overt judgment task is due to conflict monitoring problems or pragmatic tolerance rather than inability at deriving scalar inferences, confirming what suggested in recent works (Shetreet, Cherchia & Gaab, 2014) employing neuroimaging methodologies (fMRI).



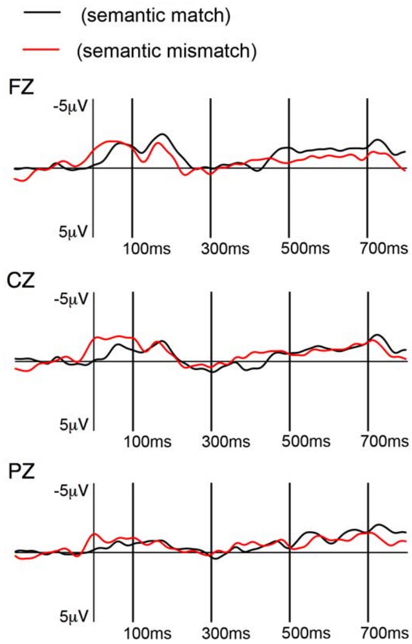
Picture 1. ERP elicited by *einpaar* (some) in children (central line electrodes)



Picture 2. ERP elicited by *alle* (all) in children (central line electrodes)



Picture 3. ERP elicited by *einpaar* (some) in adults (central line electrodes)



Picture 4. ERP elicited by *alle* (all) in adults (central line electrodes)

References. **Levinson, S.** (2000). *Presumptive meanings*. Cambridge, MA: MIT Press. **Foppolo, F., Guasti, M. T., & Chierchia, G.** (2012). Scalar implicatures in child language: Give children a chance. *Language Learning and Development*, 8, 365–394. [6] **Katsos, N.& Bishop**, D. V. M. (2011) Pragmatic Tolerance: Implications for the Acquisition of Informativeness and Implicature. *Cognition*, 20: 67-81 **Noveck, I. A.** (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, 78(2), 165-188. **Politzer-Ahles, S., Fiorentino, R., Jiang, X., & Zhou, X.** (2013). Distinct neural correlates for pragmatic and semantic meaning processing: An event-related potential investigation of scalar implicature processing using picture-sentence verification. *Brain Research*, 1490, 134–52. **Shetreet, E., Chierchia, G. and Gaab, N.** (2014). Linguistic inability or poor performance: Dissociating scalar implicature generation and mismatch in the developing brain. Dev.Psychol. 50, 2264–2275.

Introducing Virtual Reality as the Method of Choice for Experimental Pragmatics

David Peeters
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Natural languages are primarily designed for face-to-face interaction (Levinson, 1983) and much of our everyday talk takes place in dynamic, communicative, audiovisual, 3D, multimodal environments. The experimental study of the cognitive and neural underpinnings of human linguistic and communicative capacities, however, traditionally happens in strictly controlled static, non-communicative lab settings in which unimodal stimuli are commonly presented to individual participants, out of context, via headphones or in 2D on a small computer monitor. Not surprisingly, such "passive spectator science" (Hari et al., 2015) has led to neurocognitive theories of language processing that are highly language-centric and thereby do not do justice to the richness of everyday communication (Knoeferle, 2015). Having strict experimental control has clear benefits, as it allows the researcher to make causal inferences about the role of a specific independent variable in a particular process. A large discrepancy between the natural habitat of a phenomenon of interest and the experimental test setting, however, questions the ecological validity of obtained research findings and thereby the robustness and pertinence to everyday interaction of subsequently generated theories (De Ruiter & Albert, 2017). Ideally, therefore, experimental work investigating the cognitive and neural underpinnings of our pragmatic abilities should combine strict experimental control with high ecological validity.

I will introduce immersive Virtual Reality (VR) as a novel research method in the language sciences and will argue that this method is particularly suited to test pragmatic theories at a neurocognitive level in naturalistic, ecologically valid environments. In VR experiments, participants are immersed in virtual environments that closely resemble everyday communicative settings. Nevertheless, the researcher retains full experimental control over the linguistic and audiovisual input that participants receive. Furthermore, having participants communicate with virtual agents solves many of the problems that arise when one uses confederates in experimental research (Kuhlen & Brennan, 2013). I will present examples of ongoing VR-EEG studies in the field of experimental pragmatics. Specifically, I will discuss ongoing work from our lab on the concurrent processing of facial expressions and speech, and the comprehension of deictic, metonymic referential acts in virtual environments. These studies show that ecological validity and experimental control are not two extremes on a continuum but may go hand in hand when studying the neurocognitive basis of the pragmatic skills that allow us to communicate efficiently and successfully.

References

De Ruiter, J. P., & Albert, S. (2017). An appeal for a methodological fusion of conversation analysis and experimental psychology. *Research on Language and Social Interaction, 50,* 90-107.

Knoeferle, P. (2015). Language comprehension in rich non-linguistic contexts: combining eye tracking and event-related brain potentials. In R. M. Willems (Ed), *Cognitive Neuroscience of Natural Language Use* (pp. 77-100). Cambridge: Cambridge University Press.

Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: when confederates might be hazardous to your data. *Psychonomic bulletin & review*, *20*(1), 54-72.

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

Hari, R., Henriksson, L., Malinen, S., & Parkkonen, L. (2015). Centrality of social interaction in human brain function. *Neuron*, *88*(1), 181-193.

Meaning construction and novel metaphor comprehension: insights from electrophysiological data

Karolina Rataj [1, *], Anna Przekoracka-Krawczyk [2, 3], Rob van der Lubbe [4]

[1] Faculty of English, Adam Mickiewicz University in Poznań, Poland

[2] Laboratory of Vision Science and Optometry, Faculty of Physics, Adam Mickiewicz University in Poznań, Poland

[3] Vision and Neuroscience Laboratory of the NanoBioMedical Centre, Adam Mickiewicz University in Poznań, Poland

[4] Department of Cognitive Psychology and Ergonomics, University of Twente, Enschede, the Netherlands

Results of even-related potential (ERP) studies on metaphor comprehension have repeatedly demonstrated that semantic novelty modulates the N400 amplitudes (Coulson and Van Petten, 2002; Arzouan et al., 2007; Goldstein et al., 2012). These modulations have been interpreted as an index of semantic novelty and the complexity of mapping processes involved in literal and metaphoric language comprehension. Somewhat inconsistent results have been observed for the Late Positive Complex (LPC), with reports of both increased (De Grauwe et al., 2010) and attenuated (Arzouan et al., 2007; Goldstein et al., 2012) LPC amplitudes recorded in response to metaphoric utterances. To further test how semantic novelty, metaphoricity, as well as task type modulate the LPC amplitudes, two experiments were conducted, in which novel metaphors in Polish were used together with literal and anomalous sentences. In Experiment 1, 30 participants performed a semantic decision task, and in Experiment 2, 24 participants performed a reading task in response to the same set of experimental stimuli. In line with previous research, the N400 amplitudes for novel metaphors fell in between those for literal (smallest) and anomalous sentences (largest) in Experiment 1. The same although less pronounced trend was observed in Experiment 2. This effect was interpreted as an index of increased demands on semantic information retrieval related to the processing of novel metaphoric meanings. The analysis of the later time window (500-800ms) revealed smaller LPC amplitudes for novel metaphoric as compared to literal and anomalous utterances in both experiments. This finding remains difficult to interpret in light of current theories of the LPC and might reflect a negativity overlapping with late positivity, which could index continued difficulty in performing novel mappings of semantically distant concepts. Additionally, topographical between-task differences were observed, with a broad parietal distribution of the LPC effect in the semantic decision task, and a left-lateralized distribution of this effect in the reading task. These findings will be further discussed in reference to current models of metaphor comprehension and the theories of the N400 and LPC components.

References

Arzouan, Y., Goldstein, A., & Faust, M. (2007). Brainwaves are stethoscopes: ERP correlates of novel metaphor comprehension. *Brain Res, 1160*, 69-81.

Coulson, S., & Van Petten, C. (2002). Conceptual integration and metaphor: an event-related potential study. *Mem Cognit, 30*(6), 958-968.

De Grauwe, S., Swain, A., Holcomb, P. J., Ditman, T., & Kuperberg, G. R. (2010). Electrophysiological insights into the processing of nominal metaphors. *Neuropsychologia, 48*(7), 1965-1984.

Goldstein, A., Arzouan, Y., & Faust, M. (2012). Killing a novel metaphor and reviving a dead one: ERP correlates of metaphor conventionalization. *Brain Lang, 123*(2), 137-142.

* Corresponding author: krataj@amu.edu.pl

Presupposition Processing: When backgrounded information is as expected as foregrounded information – The case of factive verbs

Jacques Jayez[1,2] & Robert Reinecke[1,2]
Institut des Sciences Cognitives Marc Jeannerod (Lyon), [2]Ecole Normale Supérieure de Lyon

Presupposition triggers jointly convey two pieces of information, a main content and a backgrounded content, the so-called presupposition. For instance, the aspectual verb *stop* in the following sentence *Mary stopped smoking* communicates the main content that *Mary no longer smokes* and presupposes that *Mary used to smoke*.
Previous theoretical (Ducrot, 1972) and empirical (Jayez, 2010) research has highlighted that the main content and the presupposition have different discourse attachment properties. It is typically infelicitous to attach a discourse constituent to a presupposition. More precisely, using an acceptability judgment task for consequence discourse markers (*so*, *therefore*) and justification subordination conjunctions (because and *since*), Jayez' results indicate that discourse continuations linked to the main content (1a) are significantly more acceptable than discourse continuations to the presupposed content (1b). However, little is known about the cognitive correlates of discourse continuation processing targeting either the main content or the presupposition.

(1a) Mary stopped smoking because she got pregnant.
(1b) # Mary stopped smoking because she liked it.

Taking the heterogeneity of distinct presupposition triggers into account, the present study investigates the on-line processing of discourse continuation to the main content and the presupposition in *factive verb constructions*. Factive verbs like *know* or *realize* are verbs that presuppose the truth of the complement clause (Karttunen, 1971). These verbs have the unique property that both contents, that is the main content and presupposition are communicated by explicit material.
In order to test whether continuations to the less relevant content, i.e. the presupposition are cognitively more demanding than continuations to the more relevant one, i.e. is the main content, we were interested in the N400 component of language processing. Previous research indicates that information structured unexpectedness elicits an N400 (e.g. Cowles et al., 2007; Wang & Schumacher, 2013). In this line, it is hypothesized that a continuation to the less expected condition, that is the presupposition (2b) elicits an N400 in comparison to the more expected condition, that is the main content.

Context:
(2) Michel sait que Pierre prend le bus. (*Mary knows that Peter takes the bus.)*

Target sentence :
(2a) Pierre aussi le sait (More expected continuation)
Peter also it knows.
(2b) Pierre aussi le prend (Less expected continuation – N400)
Peter also takes it.

Using a rapid serial word paradigm, two distinct event-related potential studies (*Study 1* 29 participants (M = 20.68, SD = 2.33) and *Study 2* 30 participants (M = 20.94, SD = 2.32)) investigated the time course of semantic-thematic processes with respect to the main content and presupposition of factive verbs in an additive discourse relation scenario (see example 2 – the experimental design of study 1). In total, 88 stimuli sentences using 22 distinct factive verbs and 44 fillers were used. Half of the stimuli targeted the main content in a continuation scenario (2a), whereas the other half targeted the presupposed content (2b).
The results of study 1 (Figure 1, left) indicate that there is no evidence for an N400 for the presupposed content. Surprisingly, the main content elicited a P600 at frontal electrodes, which could be linked to the anaphoric reference *le*, which may appear more complex linking information to the abstract entity of the entire proposition, whereas a linkage to the direct object (*le bus*) may have been less demanding. In study 2, the ambiguous anaphoric reference *le* was removed. The results (Figure 1, right) show that there is no longer a frontal P600 between both contents. In addition, there is also no significant difference between both contents in the time course between 350 – 500ms.

Overall, the results indicate that discourse continuations for both contents in factive verb constructions are equally relevant. Such a result is in line with the recent literature, in which the backgrounded character of the presupposed content in factive verb constructions is questioned (see e.g. Beaver, 2010; Simons et al., 2015; Spenader, 2001). Further research should investigate the impact of prosody in order to better understand whether accenting the main content or accenting the presupposition can change the status of both contents.

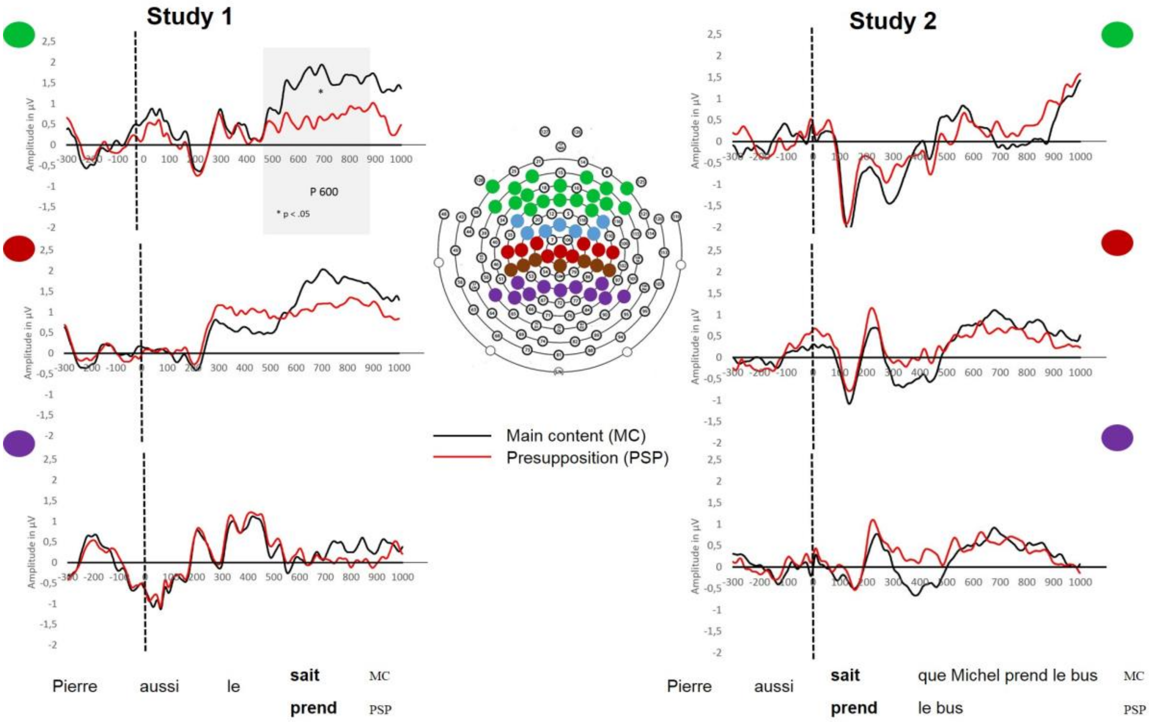


Figure 1. Processing of the main and presupposed content of factive verb constructions in study 1 (with anaphoric reference *le*) and study 2 (without anaphoric reference le).

References

[1] Beaver, D. (2010). Have you noticed that your belly button lint colour is related to the colour of your clothing. *Presuppositions and discourse: Essays offered to Hans Kamp*, 65-99.

[2] Ducrot, Oswald (1972). Dire et ne pas Dire. Paris: Hermann.

[3] Jayez, J. (2010). Projective meaning and attachment (2010). In M. Aloni et al., Logic, Language and Meaning. Revised Selected Papers of the 17thAmsterdam Colloquium, Amsterdam 2009, LNAI 6042. Springer.

[4] Karttunen, L. (1971). Some observations on factivity. Papers in Linguistics 4.

[5] Spenader, J. (2001, August). Presupposition or Abstract Object Anaphora?: Constraints on Choice of Factive Complements in Spoken Discourse. In *Information Structure, Discourse Structure AND Discourse Semantics Workshop Proceedings* (p. 161).

[6] 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.

Novel methodological approaches to perspective taking during referential communication

Maria Richter, Choonkyu Lee, Barbara Höhle, Isabell Wartenburger U Potsdam

In referential communication, listeners have to base their interpretation of an utterance on common ground (CG) that is perceptually or mentally shared information by both interlocutors. What lies in CG is monitored by an ability called perspective taking. Yet it is an open question when CG information is integrated during utterance comprehension. Listeners may integrate CG information rapidly (*Constraint-based Account*, e.g., Nadig & Sedivy, 2002; Hanna et al., 2003) or rather late with effort (*Egocentrism Account,* e.g., Keysar et al., 2000; Apperly et al., 2010). Alternatively, listeners may attempt to consider a speaker's perspective in anticipation of a linguistic expression, but then fail to fully integrate CG information due to autonomous activation of privileged information (*Autonomous Activation Account*, Epley et al., 2004; Barr, 2008). Besides these controversial accounts of the time line of CG integration, little is known about the neuronal underpinnings of perspective taking in referential communication. In the present study we therefore use the electroencephalography (EEG), a method with a very high temporal resolution, to investigate the temporal dynamics of CG information in utterance processing. In addition, we aim at revealing the electrophysiological underpinnings of the integration of CG information by means of event-related potentials (ERPs) and time frequency analysis (TFA).

In our study, 34 adults played a computerized version of the referential communication game (e.g., Keysar et al., 2000). We used a virtual 4x4 grid containing two sets of three different sized objects (i.e., a small, a medium-sized, and a big star) and two single distracters. The crucial feature of this game was the manipulation of visual access to certain objects placed in a grid: some objects were placed in CG (i.e., visible for both interlocutors), and some in privileged ground (i.e., only visible for the participant). A virtual confederate behind the grid provided auditory instructions (e.g., "Move the big star to the top."). In conflict trials, the object that fit the confederate's request best from the perspective of the participant ("competitor"; e.g., "the big star") was occluded from the confederate's view. Thus, participants had to consider CG information to select the correct object ("target"; e.g., the medium-sized star). The experiment also entailed a no-conflict condition (targets visible for confederate and participant), a no-hidden condition (no occlusions at all), and filler trials. The EEG was recorded and reaction times (RTs) and accuracy rates were measured.

The results showed, that, although overall accuracy was at ceiling (100%), perspective taking had its costs: participants were on average 195ms (SD ±7.2ms) slower in the conflict condition compared to the no-conflict and the no-hidden conditions. These increased processing costs were also reflected in the ERPs and TFA data. By focusing on the comparison between the conflict vs. the no-conflict condition, a cluster-based permutation analysis (Maris & Oostenveld, 2007) revealed enhanced late positivities (ERPs) in posterior and anterior brain regions (centro-parietal brain areas: 800-850 ms, fronto-central areas: 850-950 ms/1050-1150 ms) as well as a power increase (TFA) in theta (4-7 Hz) and alpha bands (8-12 Hz) 400-1050 ms relative to noun onset (e.g., "star"). While late positivities in the ERPs have been linked to the modification of discourse representations and to conflict resolution (e.g., van Herten et al., 2005; Bornkessel-Schlesewsky & Schlesewsky, 2008; Schumacher, 2009; McCleery et al., 2011), increased power in the theta band has been found to be associated with slow potentials in the ERPs (Herrmann et al., 2005), reflecting, for instance, the processing of complex events (Başar et al., 2001), or semantic violations (van den Brink et al., 2012). Accordingly, our ERP and TFA findings most likely support accounts that consider the integration of CG information as an effortful and relatively late process (i.e., the *Egocentrism Account*, Keysar et al., 2000, or the *Autonomous Activation Account*, Barr, 2008). However, as these accounts are based on eye tracking results, we have recently conducted an additional eye-tracking study (n=29) with an almost identical design to allow for a better discussion of the ERP and TFA data and to draw a comprehensive picture of the mechanisms underlying perspective taking in referential communication. These results will also be discussed.

References

Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A., & Moffitt, G. (2010). Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker. *Quarterly journal of experimental psychology (2006)*, *63*, 1201–1217.

Barr, D. J. (2008). Pragmatic expectations and linguistic evidence: Listeners anticipate but do not integrate common ground. *Cognition*, *109*, 18–40.

Başar, E., Başar-Eroglu, C., Karakaş, S., & Schürmann, M. (2001). Gamma, alpha, delta, and theta oscillations govern cognitive processes. *International Journal of Psychophysiology*, *39*, 241–248.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain research reviews*, *59*, 55–73.

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, *40*, 760–768.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, *49*, 43–61.

Herrmann, C. S., Grigutsch, M., & Busch, N. A. (2005). EEG oscillations and wavelet analysis. In T. C. Handy (Ed.), *Event-Related Potentials: A Methods Handbook*. Cambridge, Massachusetts: MIT Press.

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking Perspective in Conversation: The Role of Mutual Knowledge in Comprehension. *Psychological Science*, *11*, 32–38.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of neuroscience methods*, *164*, 177–190.

McCleery, J. P., Surtees, Andrew D R, Graham, K. A., Richards, J. E., & Apperly, I. A. (2011). The neural and cognitive time course of theory of mind. *The Journal of neuroscience the official journal of the Society for Neuroscience*, *31*, 12849–12854.

Nadig, A. S., & Sedivy, J. C. (2002). Evidence of Perspective-Taking Constraints in Children's On-Line Reference Resolution. *Psychological Science*, *13*, 329–336.

Schumacher, P. B. (2009). Definiteness Marking Shows Late Effects during Discourse Processing: Evidence from ERPs. In S. L. Devi, A. Branco, & R. Mitkov (Eds.), *DAARC 2009: LNAI 5847* (pp. 91–106). Berlin Heidelberg: Springer-Verlag.

van den Brink, D., Van Berkum, Jos J A, Bastiaansen, M. C. M., Tesink, Cathelijne M J Y, Kos, M., Buitelaar, J. K., & Hagoort, P. (2012). Empathy matters: ERP evidence for inter-individual differences in social language processing. *Social cognitive and affective neuroscience*, *7*, 173–183.

van Herten, M., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Brain research. Cognitive brain research*, *22*, 241–255.

Speaker's competence, hearer's prediction and the processing of scalar implicatures.

Maria Spychalska
University of Cologne & Ruhr University Bochum

It is considered underinformative to say *Some cards contain cats* if all cards contain cats, even though semantically it is true. This phenomenon is described in terms of scalar implicature: If the speaker uses a semantically weak quantifier *some*, the listener may infer that the speaker is not in a position to use the stronger expression all. Assuming that the speaker is informed regarding the stronger alternative (*competence assumption*), the listener may infer that the stronger alternative is believed by the speaker to be false. From the psycholinguistic perspective the main question has been whether this implicature is processed incrementally. Most experiments investigating this issue have involved paradigms where full information relevant for the sentence evaluation is available to all parties involved. In such contexts, underinformative sentences tend to trigger divergent truth-value judgments. Using event-related brain potentials, Spychalska and colleagues (2016) showed that this intuitive truth-value evaluation determines the way the implicature is processed: underinformative sentences were associated with larger N400 ERPs relative to informative sentences only for subjects who evaluated them as false (so-called pragmatic responders), whereas no such effect was observed for those participants who evaluated underinformative sentences as true. The N400 amplitude was further shown to be modulated by the number of alternative sentence completions that can be predicated by the hearer based on the context scenario.

Up to date, there is relatively little evidence regarding the role of the speaker's competence assumption for the implicature processing. In our current ERP experiment we investigated the processing of the scalar implicature in the context of partial information, i.e. when the assumption of the speaker's competence is violated. The experiment uses a paradigm in which participants evaluate appropriateness of the speaker's utterances about a card game situations. The target scenarios consist of (i) the speaker's avatar; (ii) four open cards placed on the table; and (iii) two cards face down (whose content cannot be seen) that are placed on the side of the speaker. The subject is informed that the speaker doesn't know what is on the face-down cards. The speaker's utterances (auditory stimuli) either refer to the cards in the game, i.e. all cards including the face-down cards (game-sentences: *Some cards in the game contain As*), or to the cards on the table only (table-sentences: *Some cards on the table contain A*s). By manipulating whether the critical noun A refers to (i) the object category contained by every visible card; (ii) the object category contained by a subset of visible cards; (iii) another object category not presented at the screen, we compare cases where the sentence's truth-value and pragmatic felicity is either known or unknown to the speaker. For the partial information context, we observe that sentences that are *known to be informative* form a significant negativity relative to *potentially underinformative* sentences, as well as relative to known underinformative sentences. In my talk I compare these results to the results of Spychalska and colleagues (2016), and discuss potential factors that determined the different processing patterns. I suggest that the context of partial information endorses the logical interpretation of *some*.

Maria Spychalska, Jarmo Kontinen, and Markus Werning. Investigating scalar implicatures in a truth-value judgment task: Evidence from event-related brain potentials. Language, Cognition and Neuroscience, 31(6):817–840, 2016.

Combining Virtual Reality and EEG to study semantic and pragmatic processing in a naturalistic environment.

Johanne Tromp, David Peeters, Antje S. Meyer, & Peter Hagoort

In natural conversation, we often process language in visually rich environments, such as a restaurant, where many cues are available to understand what someone is saying (Knoeferle, 2015). In addition, in everyday settings the communicative intention of the speaker, or *speaker meaning*, is frequently not lexically encoded (e.g. Levinson, 1983). For example when someone tells a waiter: "My soup is cold.", they may actually indirectly request the waiter to go warm up the soup. How we process this type of everyday language in contextually rich settings is not yet well understood. This is in part due to the fact that it has been difficult to design experiments with rich three-dimensional everyday contexts, while maintaining experimental control.

We used Virtual Reality (VR) and Electroencephalography (EEG) to overcome this problem and investigated audiovisual processing (Exp. 1) and indirect request comprehension (Exp. 2 & 3) in a naturalistic three-dimensional virtual environment. In all experiments, participants were immersed in a virtual restaurant (see Fig. 1), where they encountered virtual restaurant guests, seated at separate tables, each with an object in front of them.

The aim of Experiment 1 was to test the reliability of the combined use of VR and EEG to study the simultaneous processing of auditory and visual information. The restaurant guests produced sentences (e.g. "I just ordered this salmon."). The noun in this sentence could either match ("salmon") or mismatch ("pasta") with the object on the table, creating a situation where the auditory information was either appropriate or inappropriate in the visual context. We observed a reliable N400 effect as a consequence of the mismatch, which suggests that EEG and VR can be combined to study language processing in more naturalistic settings.

The aim of Experiments 2 and 3 was to investigate the processing of non-conventional indirect requests like "My soup is cold." under the same naturalistic conditions. For this type of utterance, the context, including the beliefs and inferences on the part of the listener, contribute greatly to the meaning of the utterance. In Experiment 2, participants were presented with possible indirect requests (IRs; e.g. "My soup is cold.") and control statements (e.g. "My soup is nice."). To create a naturalistic bias in the way in which the possible IRs would be interpreted, participants were assigned a role before the start of the experiment, namely to be a waiter or a restaurant critic. This was done by means of a short instruction and a mirror in the virtual restaurant, in which participants saw themselves in the outfit appropriate for the role. The task of the participant was to briefly reply to what the restaurant guest said. Replies were coded into two categories; as indicating that the participant understood the sentence ("My soup is cold.") as a request (e.g. "Sorry, I will go warm up your soup. ") or as a statement (e.g. "That is not good. "). As can be seen from Table 1, waiters provided more 'request' answers than restaurant critics. Furthermore, waiters indicated more often that they understood a request when they heard a possible request as compared to a statement, but this was not the case for the critics. Thus, we created a naturalistic context in which the same sentence ("My soup is cold.") was processed differently depending on the role of the participant. In Experiment 3 we used the same task and manipulations, while recording EEG to investigate the neural underpinnings of indirect request comprehension in real time. We are currently in the process of analyzing these data.

To conclude, we demonstrated the feasibility of combining VR and EEG to study language processing. In addition, we have developed a realistic paradigm where we can investigate pragmatic processing in a rich context without extensive explicit instruction or reference to the experimental manipulation. These initial results confirm that VR provides a way to study pragmatic aspects of language and communication in a more natural and dynamic way, even when combined with electrophysiological recordings.

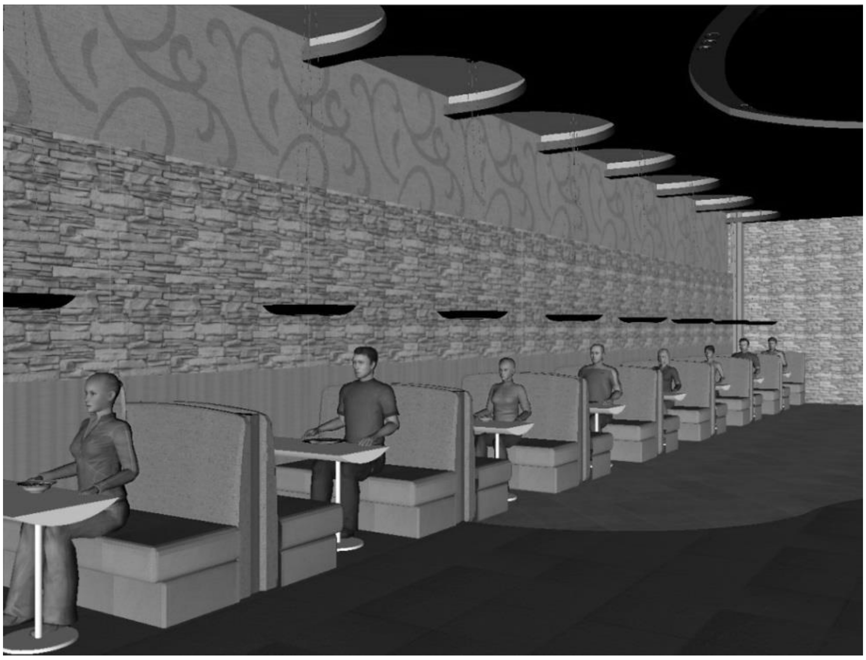


Figure 1. Screenshot of the Virtual Environment (VE).

Table 1. Percentage of 'request answers' per sentence condition (possible request versus statement) and per role (waiter versus critic).

Role	Condition			
	Possible Request		Statement	
	M	*SE*	*M*	*SE*
Waiter	84.99	1.42	4.10	0.79
Restaurant critic	3.46	0.73	0.16	0.16

References

Knoeferle, P. (2015). Language comprehension in rich non-linguistic contexts: combining eye tracking and event-related brain potentials. In R.M. Willems (Ed), *Cognitive Neuroscience of Natural Language Use* (pp. 77-100). Cambridge: Cambridge University Press.

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