Pragmatic inferencing and expert knowledge

Petra B. Schumacher and Jörg Meibauer

1. Introduction

In experimental pragmatics, there is a need to strictly separate lexical knowledge from contextual or encyclopedic knowledge. We therefore investigated the contribution of expert language – as one instantiation of encyclopedic knowledge – to a particular area of pragmatic reasoning, namely inferencing. We sought to determine whether pragmatic inferencing relies on encyclopedic knowledge or is a highly automatic operation, and which cognitive processes are affected by the computation of an expertise-based inferential relation. In the following, we first elaborate on the role of different knowledge sources for pragmatic processing, before we present a time-sensitive event-related brain potentials (ERPs) study.

The division between lexical, contextual, and encyclopedic information become for instance apparent, when testing children’s competence related to promises, where the lexical meaning of the verb *to promise*, contextual knowledge (who promises what to whom), and encyclopedic knowledge (it is bad to break a promise) has to be taken into account (Bernicot and Laval 2004). When interpreting experimental results, certain effects could have to do with insufficient lexical learning, or with a lack of contextual or encyclopedic knowledge. The representation of a context, more often than not in the form of a little story, certainly is helpful in making the task easier for the participants, but is risky in presupposing certain knowledge that may not be accessible to participants independently of the story context. Moreover, it appears that where no explicit context is given, test persons tend to construe a more or less fitting context, and accordingly, it is hard to control for all aspects of knowledge that may contribute to a judgment.

With regard to several experimental designs used in experimental pragmatics, it can be shown that contextual or encyclopedic knowledge related to lexical knowledge matters. For instance, in experimental research dealing with scalar implicatures, children and adult controls were asked whether the utterance “Some giraffes have long necks” is true (cf. Noveck 2001; Noveck and Posada 2003). “Logical” children reacted with “yes” (even if all giraffes had long necks, it is at least true that some have
long necks), whereas "pragmatic" adults answered "no" (it would be underinformative to answer yes, because, as far as they know, all giraffes have long necks). But it seems that it is encyclopedic knowledge that matters here. What, if children were not more logical, but more cautious in comparison to the adults? After all, there might exist some giraffes (e.g., a certain species or baby giraffes) that have short necks indeed. Even some adults were apparently unsure about the pragmatic adequacy of the utterances. While 89% of the responses made by children aged 7-8, and 85% of the responses by children aged 10-11 revealed "logical" agreement, 59% of the responses of the "pragmatic" adults reflected disagreement. However, 41% of the adult responses indicated child-like behavior (i.e., agreement). More generally, then, one might want to ask whether and to what extent expert knowledge influences the generation of pragmatic inferences.

In recent discussions between Neoricanists and Relevance Theorists, the question of how context influences pragmatic inferencing is crucial. Part of my experimental research focusing on scalar implicatures is motivated by the attempt to decide between two competing pragmatic models, the theory of Generalized Conversational Implicature (GCI) (Default view) and Relevance theory (RT) (Context Driven view) (cf. Noveck and Sperber 2007). Geurts 2010: 83–103). While researchers such as Bott and Noveck (2004) conclude that the derivation of scalar implicatures is highly context-dependent and does not show the autonomy claimed by default approaches such as Levinson’s (2000), this does not necessarily render the notion of a GCI totally superfluous. After all, a GCI might be a type of conversational implicature that is recurrently derived in a number of contexts, as perceived by hearers. There may still be room for a pragmatic theory that views GCIs as “regularities of use that, despite being systematic, should not be confused with linguistic meanings”, as Bach (2007: 27) put it.

Furthermore, when reasoning about the automaticity (or default) of deriving pragmatic inferences vis-à-vis the influence of context and contextually evoked knowledge, the problem of identifying and separating lexical and associated encyclopedic knowledge of the participants has widely been neglected. Usually, it is simply presupposed that participants have the necessary contextual and encyclopedic knowledge that enables them to derive pragmatic inferences. If they succeed, the results are always taken to be related to their lexical knowledge (what they know about the meaning of a lexeme, for instance some and all) plus the ability to use Gricean maxims, but in most cases not as an effect of applying contextual or encyclopaedic knowledge to the respective task. We suspect, however, that the relevant knowledge could very well matter.

We therefore designed a study where expert knowledge is directly investigated with regard to pragmatic inferencing. We decided not to replicate designs dealing with scalar implicature where the (often subtle) literal meaning of the respective trigger is at stake, but to focus on lexemes of an expert language where the literal meaning of target lexemes is clearly defined within a scientific expert system. In particular, we present data from event-related brain potentials (ERPs) to examine how rich context must be and whether expert knowledge can impact pragmatic inferencing. To this end, we investigated inferential processing in short passages consisting of two sentences and varied the validity of the inferential relation (correct vs. false) and expert-specific knowledge (from linguistics vs. chemistry).

An advantage of this design is that the contextually relevant knowledge is clearly related to the lexical meanings of certain lexemes. Experts have a direct access to the relevant lexical knowledge. For non-experts, there is a lack of respective knowledge. Since even experts may go wrong, and even non-experts may be happy in deriving correct relations, we are able to study how lexical knowledge becomes active in the task of deriving correct inferential relations. A further advantage of this study, as compared with research related to testing GCI- versus RT-approaches, is that there is no direct connection to Gricean maxims, e.g., the Maxim of Quantity (see, however, Huang 2010). While the Maxim of Relevance may play a role, this assumption is not necessary for anaphoric binding.

1.1. Electrophysiological measures

Previous investigations of discourse processing, and pragmatic inferencing in particular, have identified two central electrophysiological signals (for an overview see Schumacher 2009): First, a negative brain potential peaking around 400 ms after stimulus-onset (N400) indexes lexical-semantic processing and is a reflection of expectations built up by previous discourse, the immediate situation of utterance, and most crucially for present purposes, lexical, encyclopedic and contextual knowledge. The amplitude of the N400 decreases with increasing predictability (cf. e.g., Kutas and Federmeier 2000) and is inversely related to a word's frequency of occurrence and familiarity (for an overview see Kutas, van Petten and Kluender 2006). Second, a later positive deflection with an onset latency around 600 ms after stimulus-onset (which we here refer to as Late Positivity) reflects discourse updating costs, arising from either the introduction or the reorganization of information units in discourse representation (cf. Burkhardt
knowledge, because the addressee assumes correct and cooperative behavior on the part of the speaker; such an automatic process should result in no Late Positivity differences as a function of expertise. Following this line of reasoning, the generation of an inference should be closely tied to the associated evaluative processes, and previous research has reported modulations of positive potentials as a function of task demands. For instance, when required to judge an antonymy relation (“black − white”) or to carry out a sensicality task on a sentence expressing an antonymy (“The opposite of black is white”), highly expected continuations (“white” vs. “yellow”) registered a positive deflection (P300), reflecting the ease of evaluating the critical antonym (Roehm et al. 2007; Kretschmar 2010). Task-dependent positivities have further been observed for expected completions of idioms and collocations in figurative expressions (Vespignani et al. 2010; Molinaro and Carreira 2010). It thus appears possible that the evaluation of an inferential relation has consequences for discourse-internal operations.

More generally, expertise has been shown to influence cognitive mechanisms, for instance in studies on various aspects of music processing and rhythmic performance comparing trained musicians to laypersons (e.g., Krampe et al. 2000; Koelsch, Schmidt and Kanso 2002; Regnault, Bigand and Besson 2001) or object recognition by experts for birds or dogs (Tanaka and Curran 2001). The general tenor has been that training and experience enhance expert-specific skills and modulate electrophysiological responses. Expertise can further be considered a function of the amount of training and experience, which may account for certain developmental changes as well (cf. e.g., Carey 1992 for findings of face recognition improving with age). This corroborates the idea that the source of the divergent performance observed between children and adults may be more general in nature and represent lack of experience, which makes children more cautious in certain decisions.

Expertise may thus also influence inferential processing, and in particular both the processes underlying the Discourse Linking and the Discourse Updating stage. Expectations built up on the basis of encyclopedic knowledge should be guided by individual knowledge, i.e. expert-based relations should be more familiar and hence easier to retrieve than non-expert relations, resulting in a more pronounced N400-amplitude for non-expert over expert relations. In addition, if the construction and maintenance of discourse representation structure is driven by task-specific instructions, assessing an expertise-supported inference might be easier and hence engender a more pronounced Late Positivity.

2006, 2007). The underlying processes are labeled Discourse Linking and Discourse Updating respectively (as for instance sketched out in Schumacher 2009).

Research on inferential processing has compared the comprehension of an inference-based anaphor – such as “the professor” in “Tim attended a lecture in Berlin. He said that the professor was very nice.” – to the processing of a coreferential anaphor – as in “Tim visited a professor in Berlin. He said that the professor was very nice” (e.g. Burkhart 2006). The ERP data revealed a more enhanced N400 for inferred relations compared to repeated coreference relations, reflecting differences in the lexical fit between the critical expression and prior context, and a more pronounced Late Positivity for inferred expressions, reflecting costs from the introduction of a corresponding new discourse referent and/or the encoding of the corresponding discourse relation (e.g., elaboration, narration — as outlined in Asher and Lascarides 1983). These findings and follow-up investigations indicated that lexical and contextual knowledge facilitates the early Discourse Linking stage (i.e., coreference links are easier than inferential links). Subsequently, discourse-internal maintenance demands for the introduction of an independent discourse representation corresponding to “the professor” in the inferential case, resulting in more enhanced Discourse Updating costs.

While there is ample evidence in the literature for co-textually induced N400-modulations (e.g., van Berkum, Hagoort and Brown 1999; Burkhart 2006), it has also been shown that encyclopedic knowledge (what speakers know about the world) can modulate the N400. Hagoort et al. (2004) reported N400-effects for world knowledge violations (“Dutch trains are white.”); van Berkum and colleagues (2008, 2009) found N400-effects to manipulations of stereotypical knowledge (a child / adult uttering “Every evening I drink some wine before I go to bed.”) and hearer-specific values (a conservative person reading “I think euthanasia is an acceptable / unacceptable course of action”). We therefore predicted that individual expertise, as part of hearer-specific knowledge, should affect the processes underlying the N400, since it facilitates lexical retrieval required for linking an incoming word to the wider discourse.

As far as processes of pragmatic inferencing are concerned that are required for discourse-internal operations, two hypotheses must be tested. On the one hand, expert knowledge may not only license linking processes, but also pragmatic inferencing, and successful inferencing would thus be reflected in a Late Positivity effect. On the other hand, inference drawing may be a rather automatic process that occurs independent of encyclopedic
2. The Current Study

2.1. Methods

2.1.1. Participants

Twenty-eight students from the University of Mainz (15 women) participated in this investigation. Half of them majored in linguistics and half of them in chemistry. All participants were monolingual native speakers of German. Their ages ranged from 19-29 years (mean age: 23.1), they were right-handed (assessed by a German version of the Edinburgh handedness test) and reported normal or corrected-to-normal visual acuity. Five participants had to be discarded from the analysis because of poor performance in the expert knowledge test assessed after EEG recording; an additional two participants were excluded from the analysis because they showed excessive ocular artifacts.

2.1.2. Materials

One hundred and sixty pairs of stimuli were constructed with the factors expert knowledge (linguistics/chemistry) and validity of the inferential relation (correct/false). Eighty sets each were constructed on the basis of expert knowledge from linguistics and chemistry respectively. Examples are provided in (1) and (2) below. Context sentences introduced a particular subfield (e.g., phonology, stereoisomers) and target sentences referred to a specific aspect of this area, allowing for a correct, expertise-based inferential link in the (a)-cases below (e.g., phonology - sonority, stereoisomers - chirality). For the false relation (b)-cases), an illegitimate pairing was constructed (e.g., semantics - sonority, sodium chloride - chirality).

The materials were constructed and checked for accuracy by graduate students in the two disciplines. The stimuli were distributed across lists, so that each participant saw 80 stimuli from the linguistics set and 80 from the chemistry set, representing 40 correct and 40 false relations each. Additional 160 filler trials were constructed and interspersed with the 160 critical trials, including coherent and incoherent mini-discourses, as well as some mini-discourses requiring inferences based on general world knowledge (e.g., accident-casualty).

(1a) Linguistics knowledge – Correct inferential relation:
Flora hat gestern mit ihrer Freundin über Phonologie geredet. Lange Zeit haben sie über die Sonorität diskutiert.
‘Yesterday, Flora talked with a friend about phonology. For a long time, they discussed the sonority.’

(1b) Linguistics knowledge – False inferential relation:
Flora hat gestern mit ihrer Freundin über Semantik geredet. Lange Zeit haben sie über die Sonorität diskutiert.
‘Yesterday, Flora talked with a friend about semantics. For a long time, they discussed the sonority.’

(2a) Chemistry knowledge – Correct inferential relation:
Tim will sein Wissen über Stereoisomerie vertiefen. Am Freitag will er über Chiralität recherchieren.
‘Tim wants to deepen his knowledge on stereoisomers. On Friday, he wants to explore chirality.’

(2b) Chemistry knowledge – False inferential relation:
Tim will sein Wissen über Kochsalz vertiefen. Am Freitag will er über Chiralität recherchieren.
‘Tim wants to deepen his knowledge on sodium chloride. On Friday, he wants to explore chirality.’

The participants’ task was a sensicality judgment. They first had to indicate as quickly and accurately as possible whether the two sentences fitted, by pressing a yes- or no-button on a gamepad. Next, they were asked to indicate their certainty on a four-point scale by pressing one of four buttons.

Following the ERP recording, each participant’s knowledge in their field of studies was assessed in a questionnaire, in which all correct and false pairs of words from the ERP study were presented and participants were asked to rate the connection between the two critical words (e.g., phonology – sonority) on a 7-point-scale.

2.1.3. Procedure

Participants were instructed to read the material for comprehension and to respond to a sensicality and a certainty judgment task at the end of each trial. Stimuli were presented visually on a computer screen in yellow letters against a blue background. Each trial began with a fixation star displayed for 500 ms and followed by a blank screen of 150 ms. Each stimulus was
then displayed in segments (single words for 400 ms and two word segments for 450 ms), with an intersegment blank screen of 150 ms. Following a blank screen of 150 ms, questions for the two tasks were presented and participants were required to respond as quickly and accurately as possible by pressing a button on a gamepad. Following a 500 ms blank screen, the next trial started.

After participants were prepared for the experiment, they completed a short training session to get acquainted with the experimental procedure and were then presented with eight blocks of 40 trials each. After recording, participants filled out a questionnaire that assessed their knowledge of linguistic and chemistry-based relations on a 7-point-scale.

2.1.4. Electrophysiological Measures and Preprocessing

The electroencephalogram (EEG) was recorded from 27 Ag/AgCl scalp electrodes mounted in an elastic cap (ground: AFz; reference: left mastoid). Signals were referenced offline to linked mastoids. Electrooculograms (EOGs) were recorded by means of two sets of electrode pairs placed at the outer canthus of each eye and above and below the participant’s right eye. Electrode impedances were kept below 5 kΩ. All EEG and EOG channels were amplified and digitized with a rate of 500 Hz. The EEG data were bandpass-filtered offline (0.3-20 Hz) to exclude slow drifts. Furthermore, EOG rejections were performed automatically (±40 mV) and manually. In addition to trials with excessive artifacts, trials that registered an incorrect or timed-out response to the sensicality task were also excluded from the ERP data analysis.

2.1.5. Data Analysis

The ERP analyses are based on trials that registered correct responses to the sensicality task. Grand average ERPs were time-locked to the onset of the critical NP (the sonority/the chirality in (1) and (2)). Repeated-measures analyses of variance (ANOVAs) were performed with the factors Expertise (2 levels: expert vs. non-expert) and Validity (2 levels: correct vs. false inferential relation) for lateral and midline electrodes separately. The lateral analyses included the factor region of interest (ROI) (with 4 levels: left anterior (F3/FC1/FC5/C3), right anterior (F4/FC2/FC6/C4), left posterior (CP1/CPS/P3/P7), right posterior (CP2/CP6/P4/P8)) and the midline analysis included the factor ELECTrode (with 2 levels: anterior (Fz/FCz/Cz), posterior (CPz/Pz/POz)). All analyses were carried out in a hierarchical manner on the mean amplitude value per condition, i.e., only significant interactions (p < .05) were resolved. To avoid excessive type I errors from violations of sphericity, we applied Huynh and Feldt correction when the analysis involved factors with more than one degree of freedom in the numerator. Analyses were carried out for the time windows between 300 and 450 ms (N400) and between 500 and 650 ms (Late Positivity).

2.2. Results

2.2.1. ERP Data

Figure 1 illustrates main effects of EXPERTISE, reflected in a more pronounced negativity between 300 and 450 ms for non-expert knowledge and a Late Positivity starting around 500 ms for expert knowledge. These patterns were confirmed by statistical analyses. The ANOVA in the time range from 300-450 ms post-onset revealed a main effect of EXPERTISE (lateral: F(1,20)=8.19, p<.01; midline: F(1,20)=12.44, p<.01). Analyses between 500-650 ms after the onset of the critical NP also registered a main effect of EXPERTISE (lateral: F(1,20)=7.08, p<.02; midline: F(1,20)=8.47, p<.01), and a marginal effects of VALIDITY over midline sites (midline: F(1,20)=3.58, p<.08; lateral: p<.15). All effects were broadly distributed, reflected by the absence of an interaction with the topographical factors.

2.2.2. Expert Knowledge Assessment

Following ERP recordings, each participant was asked to rate the relation of word pairs in a questionnaire on a 7-point-scale (‘1’ indicating no relation between words, ‘7’ indicating a very strong relation). As mentioned in 2.1.1, five participants performed exceptionally poorly in this task and were not capable of correctly identifying the correct and false expert-based pairs. They were discarded from further analysis. The remaining experts evaluated the correct expert-based relations as highly related (mean=6.20, mode=7) and the false expert-based relations as unrelated (mean=2.89, mode=1). Non-expert-based relations received a medium rate irrespective of the validity of the relation (mean=3.21, mode=4), due to the fact that participants were forced to indicate a rating and did not have the possibility of opting out in the absence of sufficient encyclopedic knowledge. The
data therefore clearly demonstrate a good basis of expertise in our participants.

![Expert knowledge and inferencing](image)

**Figure 1.** Grand average ERPs recorded to the onset of the critical anaphor (onset at vertical bar) at 9 selected electrode positions. Time-course is plotted on horizontal axis from -200 ms until 900 ms relative to the onset of the anaphor; voltage fluctuations are plotted on vertical axis with negative voltage going upwards.

3. Discussion

The current findings indicate a strong influence of expertise on both Discourse Linking (N400) and Discourse Updating (Late Positivity). The data demonstrate a main effect of expertise in the N400 signature, with a more pronounced negativity for non-expert-based items over expert-based items. The former are less familiar and not part of the lexical representation of the non-experts. The enhanced N400 thus indicates that the respective knowledge cannot be easily retrieved. This converges with prior findings from the N400 as a marker of lexical and encyclopedic knowledge (Hagoort et al. 2004; Kutas, van Petten, and Kluender 2006; van Berkum et al. 2008, 2009). Crucially, it demonstrates that expert-specific knowledge influences the earliest stages of language processing and that individual encyclopedic and/or lexical knowledge is activated at this point.

Another remarkable finding with respect to the N400 is the absence of an effect of validity, which suggests that the inferential relation is not yet evaluated. This substantiates earlier findings from the processing of negation that have shown that lexical relations overrule truth value judgments (Fischler et al. 1983). In this research, Fischler and colleagues compared utterances like “A robin is (not) a bird.” and “A robin is (not) a tree.” They found no influence of the negation particle at the sentence-final word, but instead a strong lexical N400 effect (“tree” > “bird”). Interestingly, when a negation operator is contextually supported (as in “With proper equipment, scuba-diving is (not) very safe / dangerous”), truth values affect early language processes as reflected by N400 modulations (cf. Nieuwland and Kuperberg 2008).

The data also show a more enhanced Late Positivity for expert-based inferences over non-expert based inferences, suggesting that the former are more computationally demanding than the latter. In addition, the Late Positivity also reveals a marginal effect of validity, which appears to be mainly driven by the difference between correct and incorrect expert-based inferences, reflected by the most pronounced positive deflection for the correct expert-based inferential relation. This indicates that valid expert-based inferences exert the highest discourse updating costs, which may imply that when the inferential relation cannot be evaluated (in the non-expertise cases) or when an inferential relation cannot be established (in the invalid expert-based case), the discourse representation is no longer constructed.

This also converges with the idea sketched out in the Introduction that the Late Positivity may be modulated by task demands and be reflective of the evaluation of the inferential relation. Along these lines the data show...
that the evaluation of an inferential relation - on the basis of expertise - has processing consequences, and we would like to suggest that the validity of an inferential relation determines whether a discourse representation structure is built. Overall, the results thus reveal that expert knowledge is relevant for the derivation of pragmatic inferences.

Rough-and-ready conclusions with regard to some of the issues mentioned in the Introduction are not easily obtainable. We have focused upon knowledge associated with expert lexical items and the pragmatic processes of anaphor resolution triggered by them. But while it is widely acknowledged that there is an intrinsic relationship between lexical meaning and world knowledge (cf. Hobbs 2011), it is far from clear whether the anaphoric relation is triggered on the basis of the lexical meaning of the anaphor (e.g., in the pair phonology - sonority), establishing a sense relation of hyponymy, or through the creation of a context by the hyperonym (phonology), which facilitates the encyclopedic subsumption of the anaphor under this context. Hence it seems necessary to control for the context-creating forces of lexemes, too.

Finally, it is tempting to link our findings to ongoing debates in experimental pragmatics on how to account for so-called GCLs (generalized conversational implicatures) versus PCLs (particularized conversational implicatures). As pointed out in the Introduction, the basic questions are (i) whether such a distinction makes sense at all, and (ii) whether differences in their computation help to decide which approach (Default or Relevance) is the correct one. While the type of pragmatic inferring instantiated by anaphoric resolution may be distinct from the one used for calculating implicatures (the latter always involving conversational maxims such as the Maxim of quantity), a comparison is in order, in as far as in both realms, the lexical meaning of items matters.

It appears then that the ability to derive pragmatic inferences in the expert knowledge scenarios might also have to do with experience or frequency of coming across certain contexts in which hyponymy relations play a role. First of all, recall that seven students had to be excluded because of their incapability to perform the task correctly. Second, there were experts who failed and non-experts who answered correctly. This can only be explained when not only lexical knowledge matters but also the attempt at constructing a fitting context in which the correct result will follow quasi automatically. The idea is then that pragmatic inferring has to do with lexical knowledge (some, sonority, ...) (cf. Guasti et al. 2005), pragmatic inferring (quantity implicature, anaphor resolution) as well as constructing a fitting context, be it a context supporting the implicature of some or not all, be it a context of constructing sonority as a hyponym of phonology. We therefore propose to control more meticulously for the construction of context in the course of pragmatic inferring (cf. Meibauer t.a.; Finkelbeiner, Meibauer and Schumacher 2012).

In sum, we assume that the anaphoric relation in the present experiment is a type of pragmatic inference based on lexical knowledge (of the experts) linked to world knowledge. A semantic relation between the hyponym and the hyperonym supports the derivation of the correct inference. Expert knowledge, this is the overall result of the current study, is indeed relevant for anaphor resolution. This finding possibly has consequences for our understanding of developmental processes, since the acquisition of expert knowledge has to do with lexical learning. In the course of learning, true experts are supposed to come across many contexts, in which a certain anaphoric relation is valid. Here, we may see a parallel to the assumption that adults react differently from children, because they have more experience with certain types of scalar implicature (cf. Katsos 2009). We conclude, then, that studies that take the experience of participants into account would be worthwhile to sharpen our understanding of the language architecture.

Acknowledgements

We would like to thank Anika Jôdicke, Manuel Dangl and Oliver Henrich for their assistance at various stages in the preparation, collection and analysis of the data.

References

Bott, Lewis, and Ira A. Noveck

Burkhardt, Petra

Carey, Susan

Finkbeiner, Rita, Jörg Meibauer, and Petra Schumacher (eds.)

Fischler, Ira, Paul A. Bloom, Donald G. Childers, Salim E. Roucos, and Nathan W. Perry

Geurts, Bart

Guasti, Maria Teresa, Gennaro Chierchia, Stephen Crain, Francesca Foppolo, Andrea Gualmini, and Luigi Meroni

Hagoort, Peter, Lea Hald, Marcel Bastiaansen, and Karl Magnus Petersson

Hobbs, Jerry R.

Huang, Yan

Katsos, Napoleon

Koelsch, Stefan, Björn-Helmer Schmidt, and Julia Kansok

Krampe, Ralf T., Reinhold Klieg, Ulrich Mayr, Ralf Engbert, and Dirk Vorberg

Kretzschmar, Franziska
2010 The electrophysiological reality of parafascial processing: On the validity of language-related ERPs in natural reading. Ph.d. diss., Department of German Linguistics, Philipps-Universität Marburg.

Kutas, Marta, and Kara D. Federmeier

Kutas, Marta, Cyma van Petten, and Robert Kluender

Levinson, Stephen C.

Meibauer, Jörg
2010 Pragmatic evidence, context, and story design. *Language Sciences* (Special Issue on “Converging Data Sources in Cognitive Linguistics”)

Molinaro, Nicola, and Manuel Carreiras
2010 Electrophysiological evidence of interaction between contextual expectation and semantic integration during the processing of collocations. *Biological Psychology* 83 (3): 176–190.

Nieuwland, Mante, and Gina R. Kuperberg

Noveck, Ira A.
2001 When children are more logical than adults: experimental investigations of scalar implicature. *Cognition* 78: 165–188.

Noveck, Ira A., and Andres Posada
Noveck, Ira A., and Dan Sperber

Regnault, Pascaleine, Emmanuel Bigand, and Mireille Besson

Roehm, Dietmar, Ina Bornkessel-Schlesewsky, Frank Rösler, and Matthias Schlesewsky

Schumacher, Petra B.

Tanaka, James W., and Tim Curran

van Berkum, Jos J. A.

van Berkum, Jos J. A., Peter Hagoort, and Colin M. Brown

van Berkum, Jos J. A., Danielle van den Brink, Cathelijne M. J. Y. Tesink, Miriam Kos, and Peter Hagoort

van Berkum, Jos J. A., Bregje Holleman, Mante Nieuwland, Marte Otten, and Jaap Murre

Vespignani, Francesco, Paolo Canal, Nicola Molinaro, Sergio Fonda, and Cristina Cacciari