Presupposition Processing - The Case of German *wieder*⋆

Florian Schwarz[1] and Sonja Tiemann[2]

[1] Department of Linguistics and IRCS, University of Pennsylvania
florians@ling.upenn.edu
[2] Eberhard-Karls Universität Tübingen
sonja.tiemann@uni-tuebingen.de

Abstract. Presuppositions are vital for language comprehension, but little remains known about how they are processed. Using eye tracking in reading, we investigated two issues based on German *wieder* ('again'). First, we looked at the time course of presupposition processing by testing for processing costs of unsupported presuppositions. Secondly, we tested whether embedding *wieder* under negation affected a potential mismatch effect. Presupposition-induced effects showed up immediately after *wieder*, but only in the unembedded context, suggesting that embedding interferes with the detection of the mismatch. However, judgments in a follow-up rating study indicate that a mismatch is perceived in both the embedded and unembedded conditions when the PSP is not supported by the context. Taken together, these results suggest that detection of the mismatch under embedding is delayed in processing.

Keywords: Presuppositions, Psycholinguistics, Presupposition Processing, Presupposition Projection, Eye Tracking

1 Introduction

While the recent literature has seen a renewed peak in theoretical discussion of presuppositions, together with consideration of ever more intricate data (see Schlenker 2010 and Beaver and Geurts, to appear, for recent surveys), experimental approaches to presuppositional phenomena are still in their beginnings. Based on the general notion that presuppositions require some form of contextual support, previous experimental studies have found that lack of such support is reflected in various processing effects, e.g. regarding the choice of interpretation of a syntactically ambiguous structure and increase in reading times (Schwarz 2007) or the need for accommodation (Tiemann et al. 2011; see also Chemla 2009, and Chemla and Bott 2010 for other recent experimental studies). In a reading study using eye tracking, we investigated two issues concerning the processing of presuppositions. First, we test what form effects of presupposition failure, which have been found previously using Self-Paced Reading, have in

⋆ This work was supported by EURO-XPRAG (ESF) and SFB 833 (DFG).

this methodology. Eye tracking is more naturalistic and faster than Self-Paced Reading, and thus provides a more precise perspective on the time course of cognitive processes during reading. Secondly, our design compares these effects for unembedded occurrences with cases where a presupposition is introduced in the (syntactic) scope of negation (but standardly interpreted globally).

2 Background

Since different expressions that introduce presuppositions seem to vary in terms of ease of accommodation and other general properties (see, for example, the distinction between soft and hard triggers in Abusch 2009), it seems most prudent to focus experimental investigations on one presupposition trigger at a time. The present experiment focuses on the German trigger *wieder* ("again").

In the theories of Stalnaker 1973 and Heim 1990, presuppositions are restrictions on appropriate contexts. This means that a sentence like (1) is only felicitous if the context entails that Sue had danced before.

(1) Sue danced again.

An unmet presupposition results in presupposition failure and thus in uninterpretability of the sentence (cf. Heim and Kratzer 1998). Within this tradition, it is generally assumed that presuppositions are lexically encoded in the meaning of the presupposition trigger. The lexical entry for *wieder* then looks as in (2):

(2) $[\![\text{wieder}]\!] = \lambda P.\lambda x.\lambda t.\lambda w{:}\ \exists t'[t'<t\ \&\ P(x)(t')(w)].\ P(x)(t)(w)$

(3) captures formally that a sentence like (1) can only update a context if the context entails that there is a time t' before t at which Sue has danced (c is Stalnaker's context set).

(3) $\lambda c{:}\ c \subseteq \{w{:}\ \exists t'<t\ \&\ \text{Sue danced at } t' \text{ in } w\}.\ c \cap \{w{:}\ \text{Sue dances at } t \text{ in } w\}$

In contrast with semantic theories of presupposition along these or similar lines, e.g. in the frameworks of dynamic semantics (Heim 1982) and Discourse Representation Theory (DRT, Kamp 1981, van der Sandt and Geurts 1991), several issues have recently given rise to a revived debate that includes various proposals for deriving presuppostions (of at least some presupposition triggers) pragmatically (Schlenker 2009, Simons 2001). In addition to the theoretical arguments that have been brought fourth in order to distinguish between these theories, experimental investigations can contribute to the debate in that the two types of accounts suggest different time courses for computing presuppositional content in online processing. The reasoning here is very much parallel to that presented in the experimental literature on scalar implicature processing. Increases in processing time have been taken to argue in favor of accounts of implicature generation where Gricean reasoning is carried out after the core literal content is computed (Bott and Noveck 2004, and much subsequent work). To the extent that pragmatic accounts of presuppositions also appeal to Gricean reasoning, we then might expect similar delays in effects related to the interpretation of presuppositions. If there is no delay in processing presupposed content, on the other hand, that would seem to fit more squarely with a view where presuppositions are encoded conventionally as part of the lexical content of the triggers (though it may not necessarily rule out certain versions of pragmatic accounts). Our experiments vary whether the context that *again* appears in supports its presupposition or not, which allows us to evaluate the time course of presupposition processing using the high temporal resolution of eye tracking during reading and thus contributes new empirical evidence to this debate.

A second experimental manipulation relates to one of the key properties of presupposed content, namely the fact that it is not affected by various embedding operators (including negation and various attitude verbs). For example, in (4),

(4) Sue [did NOT [dance again]].

even when assuming the syntactic structure indicated by the bracketing, the presupposition escapes the scope of negation, as it were, so that the entire sentence still presupposes that Sue danced before (rather than that she didn't dance before), just like the original version without negation. From the perspective of processing, keeping apart asserted and presupposed content and taking care to interpret these distinct aspects of meaning appropriately with respect to operators like negation constitutes a fairly delicate and complex task. *Again* appears syntactically in the scope of negation and has to combine with the verb in order to derive the appropriate presupposition, but the result of this then has to be interpreted globally, rather than in the scope of negation (unlike the asserted content contributed by, e.g., the verb). Investigating the online processing of presupposition triggers in the scope of operators like negation thus has the potential to provide important insights to our understanding of the underlying processes by which the global interpretation of presupposed content is derived. Our experiment is a first attempt to shed light on this issue by directly comparing processing effects based on presuppositional content both in configurations where it appears in the scope of negation as well as in global ones.

3 Experiment

3.1 Methods and Material

Design & Stimuli. Our design makes use of a feature of German syntax, where *wieder* (again) and *nicht* (not) can appear in adjacent positions in either order. This makes it possible to construct target sentences which are minimally different with respect to whether *again* appears inside or outside the scope of negation. We presented such sentences in two different contexts, each of which supported the presupposition of one of the orders of *wieder* and *nicht* and contradicted the other. In the sample item from our materials below, the context sentence (5) supports the presupposition of (7a) (that Tina went ice-skating before), while (6) contradicts it (if not strictly speaking logically, then at least pragmatically).

Conversely, (6) supports the presupposition of (7b) (that there was a preceding occasion where Tina did not go ice-skating), while (5) is inconsistent with it.[1]

(5) *Tina* ***went ice*** *skating for the first time last week with Karl. The weather was beautiful, and they* ***had a great time****.*

(6) *Tina* ***wanted*** *to go ice skating for the first time with Karl last week. But the weather was miserable and they* ***gave up on their plan****.*

(7) Dieses Wochenende war Tina {(a) **nicht wieder** / (b) **wieder nicht**}
This weekend, was Tina {(a) not again / (b) again not}
Schlittschuhlaufen, weil das Wetter so schlecht war.
ice skating because the weather so bad was

The pairing of sentences and contexts yielded a fully counterbalanced 2×2 interaction design with two factors: **Firstword** (whether *wieder* or *nicht* appeared first) and **Felicity** (whether the context supports the presupposition or not).

Procedure & Participants. 24 sentences with versions for each of the four conditions were created. In addition to the experimental items, there were 48 unrelated filler items. Subjects read the sentences on a computer screen while there eyes were being tracked by an EyeLink 1000 eye tracker from SR Research. For half of the items (of both the fillers and experimental sets), participants had to answer yes/no questions, which followed directly after the sentence, to ensure full comprehension of the materials. 32 native speakers of German from the University of Tübingen community participated in the experiment. Subjects were split into 4 groups, where each subject saw 6 of the sentences per condition.

3.2 Results

The primary focus in our analysis were the reading times on the verb following the {*wieder nicht*} sequence. Since the presupposition of *wieder* crucially relies on the verb of its clause, it is only at the point of the verb that it becomes recoverable from explicitly given materials. Reading times were also examined for {*wieder nicht*} itself. Standard reading measures were calculated for purposes of analysis. Based on prior self-paced reading experiments using the same general approach (Schwarz 2007, Tiemann et al. 2011), we expect increases in reading time when sentences are presented in contexts that are inconsistent with the presupposition. The time point at which such increases arise is indicative of the relevant presupposition having been computed at this point.

Data Analysis. All analyses used mixed-effect models with subjects and items as random effects, using the *lmer* function of the *lme4* package in *R* (Bates 2005), together with MCMC estimates for significance (Baayen et al. 2008). All effects significant at the $p < .05$ threshold are reported. Five reading time measures were

[1] At least on a global interpretation. See below for local interpretations.

computed (Rayner 1998): first fixation duration, which measures the length of the very first fixation on the region of interest (here the verb); go-past time, which here is taken to measure the sum of all fixations on the region of interest prior to any fixations to the right of this region (but not including the time of regressive fixations); first pass time, which includes all fixations on the region when it is looked at the first time, up until leaving the region (to either the left or right); total duration, which sums all the fixations on the region of interest, no matter when they occur; regression path duration, which measures all fixations from first entering the region to first leaving it to the right (including all potential regressive fixations; this is sometimes also referred to as go past time); first pass regression proportion, which is the proportion of regressive eye movements following the first time of entering the region.

Means for the reading time measures on the verb are presented in table 1.[2] The primary result is an interaction betweeen **Firstword** and **Felicity**: when *wieder* was first (i.e., not embedded under negation), reading times on the verb were significantly higher in the infelicitous condition. When *nicht* was first, on the other hand (resulting in *wieder* being embedded under negation), there was no such slow-down (and except in total reading time and first pass time, no significant difference between the felicitous and infelicitous context conditions).

Table 1. Reading time measures (in ms) and First Pass Regression Proportion (in %) on the verb

	wieder nicht		*nicht wieder*	
Reading Measure	felicitous	infelicitous	felicitous	infelicitous
First Fixation	194	210	199	192
Go-Past	292	359	324	285
First Pass	270	281	275	247
Total	309	405	370	307
Reg. Duration	395	619	438	479
Reg. Proportion	17.0%	33.5%	17.4%	20.6%

There was a significant interaction for first fixation duration ($p < .05$), go-past ($p < .01$), and total ($p < .001$) reading times, as well as a marginal interaction for first pass duration ($p = 0.067$), regression path duration ($p = 0.056$), and and for first pass regression proportion ($p = .058$). There was a main effect of **Firstword** on all measures ($p < .05$) with faster reading times (and lower regression proportions) on the verb in the *nicht wieder* conditions, which is not generally interpretable on its own given the cross-over interaction. The interaction was primarily driven by a significant simple effect of **Felicity** for the *wieder nicht* conditions, with increases in reading measures for the infelicitous context

[2] The numbers and statistics below differ slightly from the pre-proceedings version, due to the discovery of minor errors in data treatment and reading time computation.

(go-past time: $p < .05$, total time: $p < .01$, first fixation: $p < .05$, regression path: $p < .01$, regression proportion: $p < .001$). For the *nicht wieder* conditions, the only simple effect of **Felicity** appeared in the total reading time and first pass time (both $p < .05$), and it was in the opposite direction, with a decreased reading time in the infelicitous condition. Regarding simple effects of **Firstword**, the only effect for the felicitous conditions was for total reading time ($p < .05$), where reading times were faster in the *wieder nicht* condition than in the *nicht wieder* condition. In the infelicitous conditions, the *wieder nicht* conditions displayed faster reading times on the verb than the *nicht wieder* conditions (total: $p < .01$, first fixation: $p < .05$, first pass: $p < .01$, go-past: $p < .0.1$, regression path duration: $p < .1$, regression proportion: $p < .01$). With respect to reading times on {*wieder nicht*} (taken as a unit, in either order), a parallel interaction effect showed up in the total reading times, with corresponding simple effects of **Felicity** for the *wieder nicht* condition and of **Firstword** for the infelicitous condition. There were no other significant effects in this region.

Given the lack of an effect of **Felicity** in the *nicht wieder* condition on the verb, follow-up analyses on later and larger regions were carried out. No increases in reading times were found for regions consisting of the verb plus 2 following words, the 3 words following the verb, the entire section of the sentence from *nicht wieder* to the end, or, for that matter, for the entire trial duration (i.e., total reading time for the entire paragraph).

3.3 Discussion

There are two main points to discuss with respect to the experimental results. The interaction shows that the effect of encountering a presupposition in a context that is inconsistent with it differs based on whether we are dealing with an embedded or an unembedded trigger. Furthermore, the presupposition of unembedded *wieder* gives rise to fairly immediate effects of inconsistency that are reflected throughout a variety of reading time measures. Of particular interest with regards to the latter point are the simple effects for first fixation duration and first pass regression proportion. Already during the first fixation of the verb (which last less than 200 ms), the beginning of which arguably is the logically earliest point possible to fully compute the presupposition of *wieder* based on what has been explicitly provided, a 12 ms effect emerges. Based on the experimental design, the delay can be attributed to the inconsistency between the expressed presupposition and the provided context. But for such an inconsistency to arise, the relevant presupposition must of course have been computed. Similarly, the increase in first pass regression proportions indicates that upon first looking at the verb, there is an increased likelihood of returning to look at the preceding context, which is presumably triggered by noticing the same inconsistency. The experiment thus provides evidence that the presupposition of *again* is computed rapidly online. As mentioned above, this seems most consistent with theoretical proposals that assume it to be conventionally encoded, rather than derived by some type of pragmatic reasoning, which - based on what we know about scalar implicature processing - would seem to require some extra processing time.[3]

Returning to the first point, the picture is rather different for cases where *wieder* is embedded under negation. Given the standard global interpretation, the two contexts vary in precisely the same way as was the case for the unembedded occurrence of *wieder* (albeit their roles are reversed), with one context supporting the presupposition, while the other is inconsistent with it. If the global interpretation of the presupposition were available while reading the verb, we would expect to see an effect on reading times similar to the unembedded condition. However, on none of the reading measures was there a significant increase for the infelicitous condition. In fact, the only significant simple effect (for total reading times) went in the opposite direction (which is something that we do not yet have a clear explanation for). The lack of such an increase thus can be taken as an indication that the global interpretation is not available while the verb is being read.

In principle, there are two possible explanations for why this might be the case. First, it could be that more is involved in deriving a global interpretation in the context of an embedding operator like negation, compared to simply recognizing the presupposition of an unembedded trigger. Thus, the lack of an effect might be due to a lag in generating the appropriate presupposition in this more complex sentential context. An alternative exists, however, based on the possibility (ignored in our discussion so far) of local interpretations of presuppositions in the scope of negation. Perhaps the most well-known case of this concerns the existence presupposition of the definite article, as in (8), where a global interpretation of the presuppostion is inconsistent with the continuation.

(8) The King of France is not bald - because there is no King of France!

Similarly, in (9) it seems possible to negate the presupposition that Tina had been ice-skating before, rather than the asserted content.

(9) Tina didn't go ice-skating again last weekend - this was the first time!

A simple way of modeling this local interpretation is to simply assume that both the presupposed and the asserted content remain in the scope of negation, so that the overall interpretation of the sentence can be paraphrased as follows:

(10) NOT [Tina went ice-skating before AND went ice-skating this weekend]

While in principle, the falsity of either conjunct in the scope of negation would suffice to make this true, the fact that one could express the negation of the second conjunct more straightforwardly (by simply leaving out the presupposition trigger altogether) might bias this towards an interpretation where it is indeed the falsity of the conjunct contributed by the presuppostion trigger that is conveyed by an utterance of this sentence.

[3] The extent to which this generalizes to other presupposition triggers remains to be explored. Triggers very well may vary precisely in this respect (cf. Simons 2001).

In any case, given a paraphrase along the lines of (10), if a local interpretation were available for the target sentence in the experimental materials, the *nicht wieder* sentences have interpretations that are perfectly consistent with either context. If the context states that Tina had been ice-skating some time recently, then the regular global presupposition of course remains consistent with that (and the local interpretation is not strictly speaking inconsistent with this either, if the paraphrase above is correct). And if the context states that she did not go ice-skating (and had never done it before, either), then the local interpretation (which is generally taken to convey that she had not been ice-skating before) is perfectly consistent with that. Thus, if both global and local interpretations for the presupposition of *wieder* in the scope of negation are available, we would not expect to see any reflexes of inconsistency in the reading times, since at least one of the readings always is consistent with the given contexts. In order to test whether local interpretations are indeed available for the experimental materials, a follow-up rating experiment was carried out.

3.4 Follow-up Rating Experiment

If local interpretations are indeed available for the presupposition of *wieder* when it appears in the scope of negation, we would expect this to affect speakers' acceptability judgments of these sentences in the two contexts. In particular, the type of interaction that we saw in the reading times should also be present in the judgments. If the local interpretation is not available (or only to a very limited extent), on the other hand, the *nicht wieder* sentences in what we have labeled as the infelicitous context above should be judged to be less acceptable than in the felicitous context. A rating questionnaire was conducted via the web using the WebExp2 software (http://www.webexp.info). The materials were exactly the same as those used in the eye tracking experiment, including all the fillers. Subjects were asked to rate the appropriateness of a given discourse on a scale from 1 (least appropriate) to 5 (most appropriate). The results are summarized in table 2.

Table 2. Results of the rating experiment

	wieder nicht		*nicht wieder*	
	felicitous	infelicitous	felicitous	infelicitous
Mean Rating	3.94	2.63	3.23	2.34

While there was a marginally significant interaction between **Firstword** and **Felicity** ($p = .059$), as well as a marginally significant main effect of **Firstword** ($p = .059$), more importantly there was a clearly significant main effect of **Felicity** ($p < 0.001$), with items containing felicitous contexts getting higher (= better) ratings than those containing infelicitous contexts. While this effect was slightly more pronounced in the *wieder nicht* items (giving rise to the marginal interaction), there nonetheless is a significant simple effect for *nicht wieder* in the same direction ($p < .001$), just as there is for *wieder nicht* ($p < .001$). Thus the rating results clearly show that for both embedded and unembedded *wieder*, the **Felicity** manipulation had a clear effect and resulted in decreased acceptability when the context sentence was inconsistent with the (global) presupposition of *wieder*. This would be unexpected if the local interpretation of *wieder* under negation were readily available. The explanation of the reading time results in terms of the availability of such an interpretation thus is undermined by the rating results.

4 Conclusion

The results from the eye tracking experiment showed that reading times on the verb following {*wieder nicht*} were affected differently based on the order (and corresponding scope) of negation and *wieder*, with clear effects of infelicity in the unembedded *wieder* condition and no (or opposite) effects in the embedded *wieder* condition. The immediate presence of presupposition-based effects arguably is more consistent with *semantic* accounts of presupposition, which assume that the presupposed content is conventionally encoded in the lexical entries for the triggers.

With respect to the absence of reading time effects of **Felicity** in the embedded *wieder* condition, the results from the rating study show that this cannot be attributed to the general availability of a local interpretation of the presupposition of *wieder*, as this would predict the same interaction to show up in the ratings. An alternative explanation for the lack of reading time effects in this condition is that computing the global interpretation in the syntactic context of negation is more complex in terms of processing, and that this interpretation therefore is not immediately available. What remains somewhat mysterious at this point is that no slow-downs in reading are to be found on any of the subsequent and larger regions that we analyzed. Characterizing the result as involving a delayed computation of the presupposition might lead us to expect to find the same type of increase in reading times, but on a later region. Nonetheless, the rating study clearly shows that the target sentences are perceived to be infelicitous in the infelicitous context, and it's hard to explain the absence of a reading time effect for the *nicht wieder* conditions if we assume that this infelicity becomes apparent immediately. Furthermore, there was a suggestive numerical increase in response times for the ratings in the infelicitous *nicht wieder* condition. While this did not give rise to a statistically significant interaction, there was a potential hint of a marginal simple effect of **Felicity** for this order ($p = .12$). If it were possible to substantiate such an increase in a study that is more directly targeted at capturing the time course of the acceptability judgment, that could lend further and even more direct support to the hypothesis that computing global presupposition interpretations in the context of negation is more costly.

While limitations of space as well as the experimental focus of the present research have kept us from evaluating the impact on theoretical discussions of the interaction in detail, it would be of high theoretical significance if embedded presuppositions indeed involve more processing effort. In particular, this would seem very much consistent with theories that posit explicit and complex operations on levels of representation in the computation of global interpretations, such as the DRT analysis by van der Sandt and Geurts 1991 and van der Sandt 1992 (though this is certainly not the only possible account consistent with the data). But a more thorough exploration of such theoretical implications must await future occasion.

References

Abusch, D.: Presupposition Triggering from Alternatives. Journal of Semantics 27, 37–80 (2009)

Bates, D. M. and Sarkar, D.: lme4: Linear mixed-effects models using S4 classes, R package version 0.99875-6. (2007)

Bates, D. M.: Fitting linear mixed models in R. R News 5, 27–30 (2005)

Baayen, R. H., Davidson, D. J., and Bates, D. M.: Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language 59, 390–412 (2008)

Beaver, D. and Geurts, B.: Presupposition. The Stanford Encyclopedia of Philosophy (to appear)

Bott, L. and Noveck, I. A. Some utterances are underinformative: The onset and time course of scalar inferences. Journal of Memory and Language 51, 437–457 (2004)

Chemla, E., Bott, L.: Processing Presuppositions: Dynamic Semantics vs. Pragmatic Enrichment.Manuscript, Cardiff University, IJN, LSCP (2010)

Chemla, E.: Presuppositions of Quantified Sentences: Experimental Data. Natural Language Semantics 17, 299–34 (2009)

Heim, I. and Kratzer, A.: Semantics in Generative Grammar (1998)

Heim, I.: Presupposition Projection. Reader for the Nijmegen Workshop on Presupposition, Lexical Meaning, and Discourse Processes (1990)

Heim, I.: The Semantics of Definite and Indefinite Noun Phrases, Doctoral Dissertation, University of Massachusetts at Amherst (1982)

Kamp, H.: A Theory of Truth and Semantic Representation. Formal Methods in the Study of Language, 277-322 (1981)

Rayner, K.: Eye Movements in Reading and Information Processing: 20 Years of Research. Psychological 124, 372–422 (1998)

van der Sandt, R.: Presupposition Projection as Anaphora Resolution. Journal of Semantics 9, 333–377 (1992)

van der Sandt and R., Geurts, B.: Presupposition, Anaphora, and Lexical Content. Text Understanding in LILOG, 259–296 (1991)

Schwarz, F.: Processing Presupposed Content. Journal of Semantics 24, 373–416 (2007)

Schlenker, P.: Presuppositions and Local Contexts. Mind, 119, 474: 377–391 (2010)

Schlenker, P.: Local Contexts. Semantics & Pragmatics 2, 1–78 (2009)

Simons, M.: On the Conversational Basis of Some Presuppositions. Proceedings of Semantics and Linguistics Theory 11 (2001)

Stalnaker, R.: Presuppositions. Journal of Phillosophical Logic 2, 447–457 (1973)

Tiemann, S., Schmid, M., Bade, N., Rolke, B., Hertrich, I., Ackermann, H., Knapp, J. and Beck, S.: Psycholinguistic Evidence for Presuppositions: On-line and Off-line Data. Sinn und Bedeutung 15. Proceedings of the Annual Conference of the Gesellschaft für Semantik, 581–597 (2011)