Judith Köhne-Fuetterer, Heiner Drenhaus, Francesca Delogu
and Vera Demberg*

The online processing of causal and concessive discourse connectives

Abstract: While there is a substantial amount of evidence for language processing being a highly incremental and predictive process, we still know relatively little about how top-down discourse based expectations are combined with bottom-up information such as discourse connectives. The present article reports on three experiments investigating this question using different methodologies (visual world paradigm and ERPs) in two languages (German and English). We find support for highly incremental processing of causal and concessive discourse connectives, causing anticipation of upcoming material. Our visual world study shows that anticipatory looks depend on the discourse connective; furthermore, the German ERP study revealed an N400 effect on a gender-marked adjective preceding the target noun, when the target noun was inconsistent with the expectations elicited by the combination of context and discourse connective. Moreover, our experiments reveal that the facilitation of downstream material based on earlier connectives comes at the cost of reversing original expectations, as evidenced by a P600 effect on the concessive relative to the causal connective.

Keywords: discourse connectives; event-related potentials (ERPs); eye-tracking; N400; P600; prediction

*Corresponding author: Vera Demberg**, Department of Mathematics and Computer Science, Department of Language Science and Technology, Saarland University, Campus C7.2, 66123 Saarbrücken, Germany, E-mail: vera@coli.uni-saarland.de
Judith Köhne-Fuetterer, TestDaF-Institute, University of Bochum, Bochum, Germany, E-mail: koehne-fuetterer@testdaf.de
Heiner Drenhaus and Francesca Delogu, Department of Language Science and Technology, Saarland University, Saarbrücken, Germany, E-mail: drenhaus@coli.uni-saarland.de (H. Drenhaus), delogu@coli.uni-saarland.de (F. Delogu)

1 Introduction

An increasing body of research has provided evidence that language comprehension is generally incremental and predictive (see e.g., Kuperberg and Jaeger 2015; Kutas et al. 2011; Marslen-Wilson 1973; Tanenhaus and Trueswell 1995). However, the majority of studies on predictive processing has been concerned primarily with processing at the level of syntax or semantics, leaving us with a less clear picture regarding the interplay of top-down predictions based on information given in earlier parts of the discourse and new bottom-up information which may be discourse-relevant and hence alter predictions. A particularly interesting device for studying the integration of top-down discourse expectations with a discourse-relevant bottom-up signal is the discourse connective.

Long-standing experimental evidence suggests that discourse connectors such as *therefore* and *however* facilitate coherence building and hence comprehension: Millis and Just (1994) for instance found that when sentences were connected by discourse markers (*because* and *although*), people were able to more successfully answer comprehension questions and to more quickly read the second sentence.

Existing work on the processing of discourse relations has already provided some evidence that discourse connectors may be rapidly and incrementally integrated with earlier parts of the discourse (e.g., Traxler et al. 1997a; Xiang and Kuperberg 2015), and that comprehenders are sensitive to fine-grained discourse structure (Delogu et al. 2018; Scholman et al. 2017). It is, however, still unclear how the bottom-up information from the connective is integrated with top-down predictions based on the previous discourse, and whether the observed facilitation at later content words is a prediction effect or reveals facilitated integration.

In this article, we present a series of three studies (one visual world study and two ERP studies), comparing the time-course of integrating causal connectors (e.g., *therefore*) versus concessive connectors (e.g., *however*). Since concessives have sometimes been referred to as "negative causals" (König and Siemund 2000), it is particularly interesting to examine whether processing them (compared to causals) resembles processing negations. In more detail, our experiments aim to answer the following questions:

1. Can we replicate previous findings revealing that connectives facilitate processing downstream material, and are these findings stable across paradigms (visual world, ERPs) and languages (English, German)?
2. Can we unambiguously show that the facilitative effect of connectives on later content words is related to prediction as opposed to facilitated integration?

3. Regarding global interpretation, are concessive discourse relations integrated as smoothly as causal discourse relations or do they cause processing difficulties (Carpenter and Just 1975)?
4. Do concessives elicit a search for alternatives (as has been shown for negation, Kaup et al. 2006) or an update of the situation model?

1.1 Processing discourse connectives

Different questions regarding the processing of discourse connectors have been debated in the literature, one of which is whether they facilitate comprehension. A number of scholars found facilitation effects due to discourse connectors: faster reading times (e.g., Haberlandt 1982; Sanders and Noordman 2000), better recall (Sanders and Noordman 2000), a more accurate comprehension (Millis and Just 1994; Sanders and Noordman 2000), as well as a more immediate inference of the causal relation (Cozijn et al. 2011) for marked rather than unmarked discourse relations.

A less well-studied question is how quickly connectors are integrated. Early evidence for a slow processing of discourse markers comes from Millis and Just (1994). In short discourses of two clauses, they observed longer wrap-up times at the end of the second clause when a (causal or concessive) discourse connector was present, as compared to the same sentences without a discourse connector. As a result, Millis and Just (1994) hypothesized that a representation of the second clause was constructed without taking into account the first clause, and only later integrated with the first clause.

Millis and Just's *Connective Integration Model* of late integration of discourse connectors and late integration with earlier parts of the discourse was however challenged. Traxler et al. (1997b), for instance, found evidence for an early integration of *because* with the preceding discourse: When comparing processing of causal and diagnostic sentences, the greater difficulty in diagnostics occurred well before the end of the second clause. This indicates that processing of the second clause was affected early on by its relation to the preceding context. A related reading time study by Canestrelli et al. (2013) compared the Dutch connectives *want* and *omdat*, which signal subjective (claim-argument) versus objective (consequence-cause) causality. They find that these connectives lead to different time courses in reading the sentence in the regions immediately following the connective, which also indicates that these connectives are processed incrementally and may affect discourse expectations.

The exact time course of the integration of different types of discourse connectors however still is a matter of interest: the above studies provide evidence for

a processing difficulty when discourse expectations elicited by causal connectives is inconsistent with later content of the discourse relational argument. This may however be very different for different types of connectives, in particular with connectives that are potentially more difficult to process. Previous research has consistently shown that comprehenders process causal relations more quickly than other types of relations, and that they are able to recall content from causally related sentences more accurately than from additive relations (Brehm-Jurish 2005; Murray 1995; Sanders 2005; Sanders and Noordman 2000; see also, Townsend 1983). This effect, which is unusual as fast encoding is typically associated with worse memory recall, while increased processing difficulty and slower encoding usually leads to better recollection, is also referred to as the *causality by default* hypothesis (Sanders 2005: 39). In a more recent ERP study, Kuperberg et al. (2011) found further evidence supporting the causality by default hypothesis: the processing of adjacent sentences induced a larger N400 for causally unrelated scenarios than intermediately related scenarios which, in turn, elicited a larger N400 than causally related scenarios.

The view that causal and continuous discourse relations are generally expected by comprehenders is also consistent with recent analyses of discourse relation marking in production data: Asr and Demberg (2012) found that causal and continuous discourse relations are less likely to be marked using explicit discourse connectors than other (presumably less expected) discourse relations. This observation is in accordance with the uniform information density hypothesis (Levy and Jaeger 2007), which holds that optional linguistic elements (like discourse connectors) may be omitted when they do not convey much new information, i.e., when they mark an expected discourse relation. Taken together, these findings lead to the prediction that discourse connectors that signal discontinuity should be more difficult to process because they require a re-interpretation of the discourse relation that was assumed by default. Studies that confirm this hypothesis by directly comparing the processing of marked causal and for instance concessive relations are however still missing.

Interesting in this context is a series of EEG experiments by Xiang and Kuperberg (2015) that reveals that concessive connectors can have a facilitating effect: While comprehenders experience processing difficulty (inducing an increased N400) on a word that is unexpected given their world knowledge, this expectation effect can be reversed when a concessive connective (*even so*) is present (e.g., "*Elizabeth had a history exam on Monday. She took the test and aced/ failed it. (Even so), she went home and celebrated wildly*", Xiang and Kuperberg 2015: 648). This provides evidence for the hypothesis that a concessive connective can be integrated with world knowledge to reverse expectations. The authors argue that this reversal of expectations comes at the price of later sustained negativity

effects at the end of the sentence. It would be interesting to see whether this effect is also already visible on the connective; Xiang and Kuperberg however do not compare these regions, as there is no corresponding positive causal connective that *even so* can be compared to in their study. Moreover, the study does not provide direct evidence of prediction at the critical word, as the observed N400 effect is located on the target word itself and can alternatively be interpreted as a facilitated integration effect. To more precisely differentiate prediction from integration, the position where the match/mismatch effect is tested has to precede the target word itself.

1.2 Concessives as negative causals

Studies revealing that concessive relations are more effortful to process than, for instance, causal relations (e.g., Kuperberg et al. 2011; Millis and Just 1994) support an idea elaborated on by König and Siemund (2000), who consider concessives as "negative causals". This notion is also supported by experimental studies, indicating that causals and concessives establish the same type of relation, but differ in polarity (Louwerse 2001; Sanders et al. 1992).

An increased processing effort with concessives on the one hand and defining concessives as negative causals on the other hand seems to be in line with results revealing that processing negation is causing a delay in processing (e.g., Carpenter and Just 1975). Kaup et al. (2006), for instance, compared negated sentences with contradictory predicates (e.g., *The door is not open/closed*) in a self-paced reading study combined with a picture-naming task, which was timed to happen 750 versus 1,500 ms after the critical region (not open). Their results reveal that at the early stage, people are mentally simulating the positive state (open door); only at the late state, the positive state is negated, which means that a search for alternatives happened and they have mentally closed the door (see also Lüdtke et al. 2008). In line with this, Ferguson et al. (2008) found that counterfactual negated discourse information was not used incrementally but had a delayed effect on comprehension.

Other studies, on the contrary, reveal that a delay can be attenuated or removed when the negation is expected or pragmatically licensed (Dale and Duran 2011; Nieuwland and Kuperberg 2008). Staab (2007), for instance, reports a series of ERP studies showing that negation in discourse context was processed fast. Interestingly, if readers were forced to process slowly and deeply, negation was even used as a cue to rapidly anticipate how the sentence continues. Nordmeyer and Frank (2014) find that also a visual context can affect the processing of negation: Following a matching visual scene, negated sentences were verified as fast as non-negated ones.

These very different results suggest that the time course of processing negation may be influenced by a number of factors, in particular the kind of negated information and the discourse context. It is an interesting question, how processing negative causals, that is, concessive discourse markers, may enrich this picture with and without a visual context.

1.3 The present experiments

In the following, we present results of one visual-world study and two ERP studies to evaluate the exact time course, potential predictive effects, processing facilitation, and processing difficulties occurring in comprehending marked causal versus concessive discourses.

The design of our experiments is motivated by studies by Wicha and colleagues (e.g., van Berkum et al. 2005; Wicha et al. 2004). In both ERP studies, the prediction of linguistic material is evaluated exploiting gender marking preceding a target noun (see also Koornneef 2021): If readers anticipate the target noun, they should not only show a reaction when this target noun does not match their expectations but already when the preceding article or adjective does grammatically not match the expected target noun. The advantage of measuring at the gender marked pretarget region rather than the target region is that an effect could only be due to prediction but not facilitated integration (see Section 1.2).

Our ERP studies use the same logic (using grammatical gender for German and the difference between the articles *a* and *an* in English). Our visual world study also exploits grammatical gender but does not contain a mismatch condition and is rather based on analyzing the redirection of gaze due to changes in discourse expectations.

2 Experiment 1

Our first experiment investigates the time course of anticipating specific lexical items based on integrating top-down and bottom-up cues in a visual world paradigm (i.e., spoken sentences and static visual scenes). While earlier discourse content and world knowledge are integrated and provide a top-down cue for forming expectations on upcoming content, the causal versus concessive discourse connective, as well as grammatical gender information encoded in the determiner and adjective preceding the target word provide bottom-up information that could help comprehenders to further specify discourse expectations.

2.1 Method

2.1.1 Participants

We tested 36 participants, four of which had to be excluded due to eye-tracking problems. Data of 32 participants (eight male, average age 26) was analyzed. Participants were paid 5 € for taking part.

2.1.2 Materials and design

We constructed 20 items, each consisting of three spoken sentences in German, and a static scene (see (1a) for a glossed example, (1b) for an illustration of the analyzed regions, and Figure 1).

(1) a. *Mark   denkt   über   einen   kleinen   Snack   nach.   Er*
         Marc   thinks   about   a   small   snack   VPART   he
         hat   gerade   Lust,   etwas   Süßes   zu   essen.
         has   right now   desire   something   sweet   to   eat
         Daher   holt   er   sich   aus   der   Küche   die   appetitliche
         therefore   gets   he   REFL   from   the   kitchen   the.F   delicious.F
         Waffel.
         waffle
         'Marc fancies a small snack. He feels like having something sweet.
         Therefore, he gets the delicious waffle from the kitchen.'

Figure 1: Stimulus for visual world experiment.

b.   *Marc denkt über einen kleinen [Snack nach. Er hat gerade Lust, etwas]*topic *[Süßes/Salziges zu essen]*category*. [Daher/Dennoch holt er sich]*connector *[aus der Küche]*extended *[die appetitliche/den appetitlichen]*pretarget *[Waffel/Kuchen/Brezel/Käse]*target*.*
'Marc fancies a [snack. He feels like having something]topic [sweet]category. [Therefore/Nevertheless, he gets]connector [from the kitchen]extended [the[fem]/[masc] delicious[fem]/[masc]]pretarget [waffle/cake/pretzel/cheese]target.'

The first sentence introduces a situation or topic, such as food (*Marc denkt über einen kleinen Snack nach*, 'Marc fancies a Snack'). The second always identifies a category (e.g., sweet things), matching two of the depicted objects (waffle and cake; *Er hat gerade Lust, etwas Süßes zu essen*, 'He feels like having something sweet'). Two other objects in the scene belong to another category (the *counter category*, here salty things: cheese and pretzel).[1] Sentence 3 begins with a causal or a concessive connector (*Daher/Dennoch*: causal condition, *Deswegen/Trotzdem*: concessive condition; within-participant factor *Connector Type*), followed by subject and verb (*holt er sich*, 'he gets'; connector region). This region precedes another phrase such as a prepositional phrase (*aus der Küche*, 'from the kitchen'; extended connector region), the gender-marked pretarget noun region (*die appetitliche*, 'the delicious'), and the target noun (causal: *Waffel*, 'waffle', concessive: *Brezel*, 'pretzel').

Target nouns are always congruent with the preceding discourse. Visuals worlds were designed to include not only the four objects belonging to the category and the counter category, but to embed them in a little scene (here, kitchen furniture), also including two distractor objects (here, cup and egg whip).

The experimental design thus includes four conditions (two connector type causal/concessive × two gender of target noun). Due to full counter-balancing of target objects, which required alternating the category given in sentence 2, we split up the items into eight lists. Every participant was assigned to one of the lists and saw each of the 20 items in one version only (each participant thus saw five instances of each of the four experimental conditions).

Fourty filler discourse-scene pairs were included, following the same general pattern but using a range of discourse relations and markers (e.g., *später*, 'later'), making the target noun unpredictable.

---

1 We would like to thank the CUNY audience for bringing to our attention the fact that pretzels can be sweet in the US. For our German subjects, pretzels cannot be conceived of as a sweet food though.

All items and half of the fillers were followed by a comprehension question about the target noun but referring to it by category rather than name (*Holt Marc sich etwas Süßes?*, 'Does Marc get something sweet?'), which participants answered by button press (YES/NO). Half of the questions' correct answer was "yes", the other half "no". Order of presentation was pseudo-randomized, with at least one filler in between two items.

2.1.3 Procedure

Participants were tested individually, seated in front of a computer screen. They were given a button box to respond, their eye-movements were tracked using an Eyelink II head-mounted eye-tracker (sampling rate 500 Hz, spatial resolution about 0.1°). Before the start of the experiment, participants went through a calibration procedure. Viewing was binocular but only the participant's dominant eye was tracked.

Trials started with a fixation cross at the center of the screen that participants were asked to focus such that drift correction could be performed. Then they were presented a scene and a discourse of three spoken sentences. After the third sentence either the comprehension question or the fixation cross preceding the next trial appeared on the screen automatically.

Participants' task was to look and listen carefully enough to reply to comprehension questions. The experiment lasted about 30 min.

2.2 Predictions

When the category (e.g., sweet) is mentioned, fast and incremental processing predicts participants to look more often at the two objects matching this category (waffle and cake) in both conditions (e.g., Altman and Kamide 1999).

For Sentence 3, predictions for the two connector types differ: In the causal condition, people are predicted to keep looking at the category objects until the case-marked pretarget region. During the pretarget region then, fast and incremental processing predicts more looks towards the gender-congruent object, and finally, when the target is mentioned, more looks to the target. In the concessive condition, hypothesizing that the concessive connector is processed eagerly and incrementally predicts participants to direct their looks to the counter-category objects (pretzel and cheese), as soon as the scope of the concessive connector is clear.

In particular, the scope could be inferred and a search for alternatives could be initiated after the subject and verb following the connector (connector region). The hypothesis that the concessive connector is processed as fast and incrementally as

the causal one, also predicts participants to start looking more at the final target object during the gender-marked pretarget region. Note that a late integration account, or a simple lexical priming account, would not predict this pattern but that participants keep looking at the category objects (sweet things) until they hear the target word.

Processing difficulties in the concessive compared to the causal case moreover may manifest themselves in lower accuracy and higher reaction times for answering the comprehension questions in the concessive case.

2.3 Data analysis and results

For eye-movement analyses, we compared inspections to the four areas of interest (AOIs): target (e.g., waffle), category competitor (sharing category with target, cake), gender competitor (sharing gender with target, pretzel), and unrelated competitor (sharing neither category nor gender with target, cheese). Four time regions were of interest: category region, connector region, extended connector region, and pretarget region. Eye-movements were analyzed using logistic regressions, entering the data into linear mixed effect models with logit-link function (from the lme4 package in R; Bates 2005). AOI and Connector Type (causal/concessive) were used as a Fixed Factors and Participant and Item as random factors. Main effects were tested based on model comparison using a χ^2-test (Baayen et al. 2008). In models with several factors, models are built incrementally; we report the contribution of second factor with respect to a model already containing the first factor. Interactions are analyzed through model split: if we find a significant interaction with connector type, the data is split by connector type. We included the most complete random-effect structure (intercepts and slopes) that allowed the models to converge (see Barr et al. 2013). Models that did not converge were simplified using a step-by-step backwards elimination approach until the model converged. For contrasts between levels (AOIs), we report Wald-z values and p-values as well as coefficients (b) and standard errors (SE).

Analyses reveal that in the category region, participants inspected the objects matching this category (causal: target + category competitor; concessive: gender competitor + unrelated competitor) significantly more frequently than the counter-category objects, independent of the connector type (effect AOI: $\chi^2(1) = 25.64$, $p < 0.001$; no effect Connector Type: $\chi^2(1) = 1.98$, $p = 0.16$; no interaction: $\chi^2(1) = 1.96$, $p = 0.16$).

In the connector region, there was no effect of AOI ($\chi^2(1) = 1.14$, $p = 0.29$), no effect of Connector Type ($\chi^2(1) = 0.31$, $p = 0.58$) but, importantly, an interaction ($\chi^2(1) = 16.05$, $p < 0.001$). In the causal condition (see Figure 2), the category objects were still looked at significantly more often than the counter-category objects ($\chi^2(1) = 12.69$, $p < 0.001$);

in the concessive condition, however, participants inspected the two counter-category objects just as much as the category objects ($\chi^2(1) = 1.59$, $p = 0.21$). As illustrated in Figure 3, this is due to them first looking more at the category objects, but gradually starting to look more at the counter category objects, as the scope of the concessive becomes clear.

In the extended connector region then, we find significantly more looks to the objects of the target category (i.e., category in causal and counter-category in concessives), independent of the connector type (effect AOI: $\chi^2(1) = 34.03$, $p < 0.001$, no effect Connector Type: $\chi^2(1) = 0.27$, $p = 0.60$, no interaction: $\chi^2(1) = 2.04$, $p = 0.15$). This reveals that the concessive marker was immediately interpreted and people engaged in an active search for alternatives.

In the pretarget region (when shifted 200 ms),[2] the target was looked at more frequently than all other objects for both connector types (effect AOI: $\chi^2(3) = 110.26$, $p < 0.001$, no effect Connector Type: $\chi^2(1) = 0.03$, $p = 0.87$, no interaction: $\chi^2(3) = 0.76$, $p = 0.86$). The difference between target and category competitor was smallest ($b = -0.26$, $SE = 0.12$, $z = -2.14$, $p = 0.03$), while there was a large difference between target and gender competitor ($b = -1.00$, $SE = 0.13$, $z = -7.59$, $p < 0.001$) and target and unrelated object ($b = -1.17$, $SE = 0.14$, $z = -8.57$, $p < 0.001$).
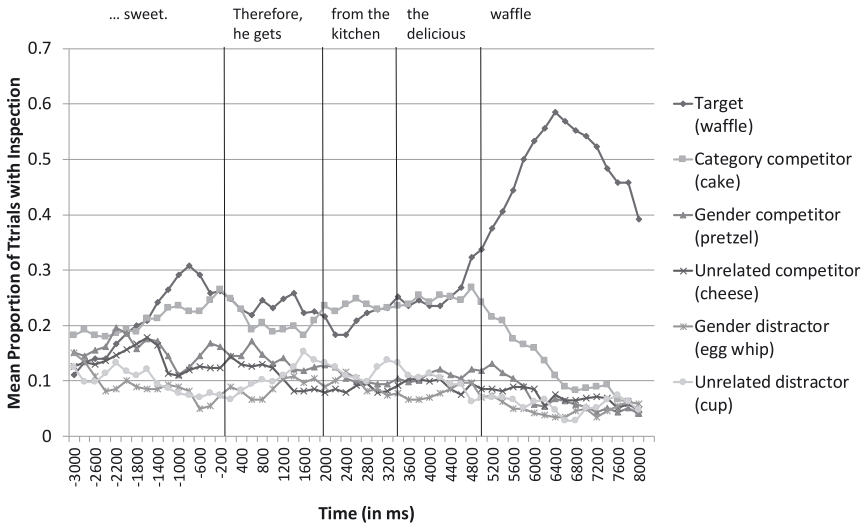

Figure 2: Results for causal condition.

---

2 This is frequently done for short eye-tracking regions in visual world studies because 200 ms is known as the amount of time needed to program an eye-movement.
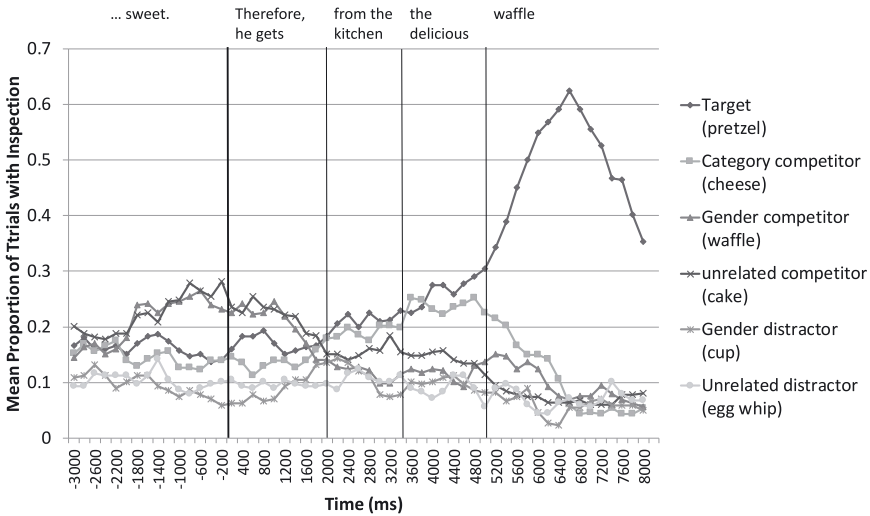
Figure 3: Results for concessive condition.

    Accuracies and Reaction Times for comprehension questions were analyzed the same way as eye-movements, except that we used linear regressions rather than logistic regressions for response times. Accuracy was significantly lower in the concessive condition (78%) than in the causal condition (84%; $\chi^2(5) = 6.48$, $p = 0.01$). More detailed analyses reveal that this difference was driven by the lower answer accuracy for those questions in the concessive condition for which the correct answer was "yes" (concessive: 70% vs. causal: 83%, $\chi^2(1) = 13.76$, $p < 0.001$): There were significant main effects of Correct Answer (yes vs. no; $\chi^2(1) = 17.74$, $p < 0.001$), Connector Type (causal vs. concessive; $\chi^2(1) = 6.76$, $p = 0.01$), and an interaction between both factors ($\chi^2(1) = 6.66$, $p = 0.01$). For questions to which the correct answer was "no", answer accuracies were high in both conditions (concessive: 85%, causal 86%, $\chi^2(1) = 1.39$, $p = 0.24$). The analysis of the response times showed that participants took a longer time to answer the questions to which the correct answer was "no" ($\chi^2(1) = 5.40$, $p = 0.02$). There was however no effect of Connector Type ($\chi^2(1) = 0.47$, $p = 0.50$), and no interaction ($\chi^2(1) = 0.30$, $p = 0.58$).

2.4 Discussion

These results reveal that both causal and concessive discourse markers were integrated rapidly into online comprehension: while attention was already on the causally congruent target in the causal condition, the concessive condition

showed a reversal of expectations in the time windows following the connective, as visual attention shifted from the causally congruent objects to the objects that are consistent with the reversed expectations elicited by the concessive connective. As in Staab (2007), listeners used negation as a cue to anticipate upcoming information, searching for alternatives. In our experiment, however, they did so fast and without explicit instruction to process deeply. Possibly the visual context facilitated prediction compared to processing written language as in Staab (2007).

Eye-movements furthermore indicate that processing was rapid and stable enough to combine with another bottom-up cue, the grammatical gender information, to predictively identify the exact target referent. Since this was found in the pretarget region, it cannot be explained by integration.

The finding that accuracy of question answering was worse in the concessive than the causal condition (when the correct answer was "yes", see Example 2) might suggest that processing in the concessive condition was shallower, causing a late cognitive burden for global interpretation.

(2)　　*Marc fancies a snack. He feels like having something sweet. Nevertheless, he gets from the kitchen the delicious pretzel. – Does Marc get something salty?*

More precisely, it is possible that suppressing the category directly mentioned in the second sentence (*sweet* in Example 2) in combination with having to categorize the named target (e.g., *pretzel* = salty) might be difficult: Firstly, it might require suppression to answer "yes" to the question whether something salty was chosen when "something sweet" was just said. Secondly, the cue to the correct answer to the question that it is correct that something salty was chosen must be inferred since "something salty" has never been explicitly mentioned but only an instance of this category (a pretzel). Both of these potential challenges may have caused participants to give the wrong answer more often than in other cases. The answers to questions where the correct answer is "no" might be easier, because in the causal case, nothing salty has been mentioned, making it easy to answer "no", while in the concessive case, the negation to the sweet category from the second sentence was already made explicit through the concessive connective.

3 Experiment 2

In order to explore whether our finding that discourse markers can be integrated rapidly, shaping predictions about upcoming words, can be replicated even when the potential referents are less clearly defined by a visual scene, we conducted two ERP experiments. The main goal was to examine: a) whether readers also predict upcoming linguistic content without the support of a visual scene, both in causal

and concessive discourses and b) whether processing concessives is more difficult than processing causals.

Two ERP components are particularly relevant for the present investigation: the N400 and a late positive component (P600). The N400, a broadly distributed negative deflection peaking around 400 ms post-stimulus onset, was initially observed in response to sentence-final semantically incongruent words, but was soon discovered to be part of the normal response to words and other potentially meaningful stimuli (Kutas and Federmeier 2011). Importantly, the N400 amplitude is sensitive to the predictability of a word in its sentential (e.g., Kutas and Hillyard 1984) or discourse (e.g., Federmeier and Kutas 1999; van Berkum et al. 2005) context.

The P600, a positive shift with a latency varying between 600 and 900 ms, has been traditionally associated with syntactic reanalysis and repair processes (e.g., Osterhout and Holcomb 1995). More recently, however, late positivities – rather than N400 effects – have been reported to semantic/pragmatic violations (e.g., Drenhaus et al. 2011; Kuperberg et al. 2003; van Herten et al. 2005). Interestingly, the "semantic P600" has been discussed as reflecting the reorganization or updating of the mental representation of the unfolding discourse (Brouwer et al. 2012).

3.1 Method

3.1.1 Participants

Sixteen undergraduate students (mean age 23, six male) from Saarland University took part in the experiment. All participants were native speakers of German, had normal or corrected-to-normal vision, and were paid for their participation.

3.1.2 Material and design

We constructed 24 experimental items, each in four different conditions, crossing the type of discourse connector (Connector Type: causal/concessive) and the congruency of the target (congruent/incongruent), as shown in (3a) and (3b):

(3)　a.　*Tim und Kim überlegen, ob　　sie　lieber tanzen gehen oder*
　　　　　Tim and Kim wonder　　whether they rather dance go　　or
　　　　　den　neuen　Film　sehen　wollen. Kim findet tanzen toll.
　　　　　the　new　　movie watch want　Kim finds　dance　great
　　　　　Daher　gehen sie　in　die　frisch　renovierte　Disko,
　　　　　therefore go　　they to the.F freshly　renovated.F disco
　　　　　um　sich　　richtig　gut　zu　amüsieren.
　　　　　to　themselves really　well　to　amuse

'Tim and Kim are wondering whether they prefer to go dancing or to watch the new movie. Kim likes dancing a lot. Therefore, they go to the freshly renovated night club to enjoy themselves really well.'

b. *Tim und Kim überlegen, ob sie lieber tanzen gehen oder den neuen Film sehen wollen. Kim findet tanzen toll. [Daher/Dennoch]*connector *gehen sie in [die frisch renovierte/das frisch renovierte]*pretarget *[Disko/Kino]*target, *um sich richtig gut zu amüsieren.*

'Tim and Kim are wondering whether they prefer to go dancing or to watch the new movie. Kim likes dancing a lot. [Therefore/Nevertheless]connector, they go to [the[fem] newly renovated/the[neut] newly renovated]pretarget [night club/cinema]target, to enjoy themselves really well.'

Each item consisted of three-sentence discourses. The first sentence always introduced two alternatives (e.g., going dancing vs. watching a film) and the second sentence then identified a preference for one of them (going dancing). The third sentence (i.e., the target sentence) began either with a causal (*Deswegen/Daher* 'therefore') or a concessive connector (*Trotzdem/Dennoch* 'however') and included a gender-marked prenominal region consisting of a determiner and an adjective preceding the target noun. The target noun could either be congruent with the expectations generated by the connector together with the context (e.g., *Kim likes dancing + Therefore → night club*; or *Kim likes dancing + Nevertheless → cinema*) or incongruent (e.g., *Kim likes dancing + Therefore → cinema*; or *Kim likes dancing + Nevertheless → night club*).

We conducted a cloze test to ensure that the congruent target words were highly predictable in both causal and concessive conditions. The 30 passages in the two conditions were truncated before the pretarget region and presented to 21 independent participants to be completed. The mean cloze probability for the congruent target words in the causal condition was 0.58 (SD = 0.26), while the mean cloze probability was 0.49 (SD = 0.25) for the concessive condition. The incongruent target words were almost never produced as completions: the mean cloze probability for the causal condition was 0.02 (SD = 0.04), while for the concessive condition was 0.03 (SD = 0.06).

The 96 experimental passages (24 items in four conditions each) were intermixed with 72 unrelated filler discourses and arranged in four lists.

3.1.3 Procedure

During the experiment, participants were seated in a sound-proof, electromagnetically shielded chamber. Discourses were presented in black fonts

(28-point Times) on a white background using E-Prime (Psychology Software Tools). After a short training session (six sentences), the 96 experimental trials and the 72 fillers were presented in pseudo-randomized order, in four blocks with intervening breaks. Each trial started with the presentation of the first two context sentences as a whole for 2,500 ms, followed by a blank screen (150 ms), and a fixation star in the center of the screen (500 ms). The target sentence was then presented word-by-word in the center of the screen, for 350 + 100 ms inter-stimulus interval (RSVP). In 25% of the cases, the target sentence was followed by a plausibility judgment task: participants were asked to press one of two buttons on a response pad within a maximal interval of 2,500 ms.

3.2 Predictions

We hypothesized that processing the connector together with the context would elicit predictions about the target noun. In particular, we expect a congruency effect (i.e., a larger N400) in both the causal and the concessive conditions when the gender of the determiner in the prenominal region does not match the gender of the predicted noun (i.e., in the incongruent condition). Finding this effect on the pre-target region would provide strong evidence that comprehenders predict upcoming lexical content based on the context and the type of discourse relation. Furthermore, under the assumption that comprehenders by default expect causal relations, we predict a larger positivity on the concessive connector compared to the causal one, reflecting a discourse model updating process (e.g., Brouwer et al. 2012).

3.3 EEG recording

The EEG was recorded by means of 26 Ag/AgCl scalp electrodes. Electrodes were placed according to the 10–20 system (Sharbrough et al. 1995). Impedance was kept below 5 kOhm. The signal was referenced and digitized at a sampling rate of 500 Hz. The EEG data were re-referenced to an average of both mastoid electrodes offline. The horizontal electro-oculogram (EOG) was monitored with two electrodes placed at the outer canthus of each eye and the vertical EOG with two electrodes above and below the right eye. During recording, no online filters were used.

3.4 Data analysis and results

The EEG data were band-pass filtered offline with 0.01–40 Hz. Single-participant averages were computed in a 1,000 ms time-window per condition relative to the

onset of the critical item and aligned to a 200 ms pre-stimulus baseline. Averaged ERPs were semi-automatically screened for electrode drifts, amplifier blocking, eye movements, and muscle artifacts (4% of data points were excluded). Only artifact-free ERP averages time-locked to the onset of the critical regions entered the statistical analyses. Critical regions were the connector (Causal vs. Concessive) and each word in the target NP: Determiner (*das/die*), adverb (*frisch*), adjective (*renovierte*), noun (*Kino/Disko*).

We fitted linear mixed models (LMM) (Bates and Sarkar 2007) with ERP values averaged over critical items for each participant as dependent measure. Fixed factors were Connector Type (Causal vs. Concessive) for the analysis of the connector region, and Connector Type and Congruency (Congruent vs. Incongruent) for the analysis of the NP region (contrast coding: −0.5 for the concessive condition and +0.5 for the causal condition, −0.5 for the incongruent condition, and +0.5 for the congruent condition). We included by-subject random intercepts,[3] as well as the maximal by-subject random slope structure (see Barr et al. 2013). We report only significant effects of (maximal) models that converged. Models that did not converge were simplified using a step-by-step backwards elimination approach until the model converged.[4] Following Baayen et al. (2008), a given coefficient was judged to be significant at $\alpha = 0.05$ if the absolute value of t exceeded 2. We report only significant effects.

Nine representative electrode sites from frontal, central, and parietal regions were included in the statistical analyses (F3, Fz, F4, C3, Cz, C4, P3, Pz, P4).

For all critical regions, we performed baseline analyses on the −200 to 0 ms time window to ensure that there were no systematic effects present prior to the presentation of the critical word. Additionally, we analyzed the eye electrodes to exclude any interference with experimental effects. None of these analyses revealed significant effects ($t < 1$).

3.4.1 Connector region

The ERP patterns to the connectors are displayed in Figure 4. The concessive connector elicited a larger fronto-central positivity than the causal connector between 300 and 600 ms. The LMM analysis collapsing over the nine electrodes revealed a significant effect *(b = 1.2924, SE = 0.3146, t = 4.109)*. Separate analyses of frontal

---

3 We did not include items in the random structure of the LMM because item data with satisfactory signal-to-noise ratio are generally not available with ERPs (see Laszlo and Federmeier 2011). Our LMEMs were based on mean amplitudes averaged across subjects, time-windows, and electrodes.
4 However, the simplified converging models did not show significant effects regarding the regions of interests.
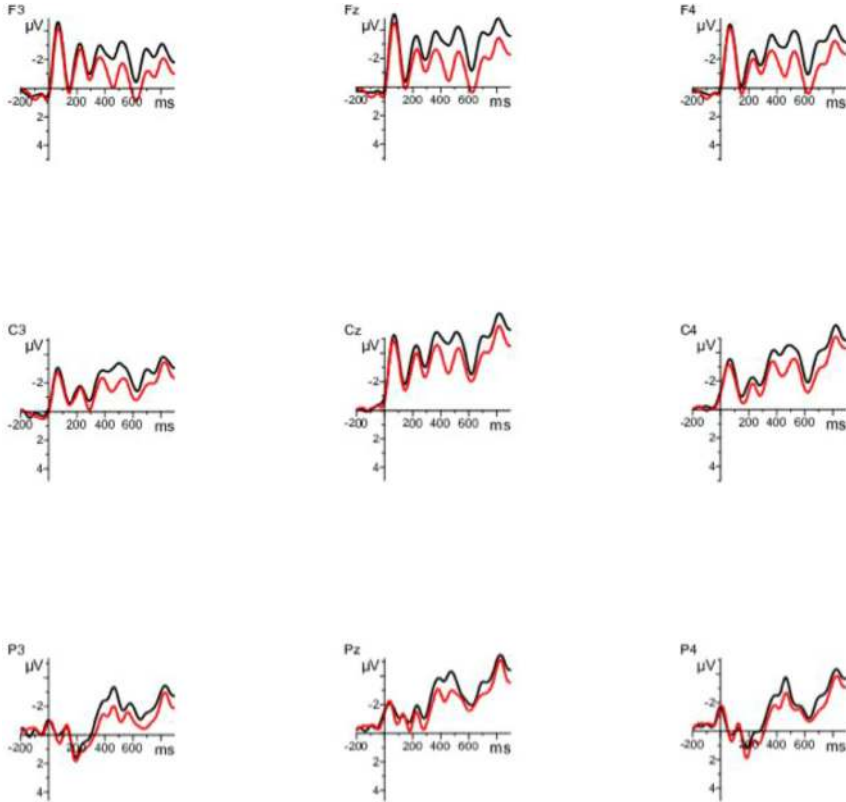
Figure 4: ERP effects on the discourse connective, on a subset of nine electrodes. Negativity is plotted upwards. Black line: causal connector condition; red line: concessive connector condition. For presentation purpose only, ERPs were filtered off-line with 10 Hz low pass.

(F3, FZ, F4), central (C3, CZ, C4), and posterior (P3, PZ, P4) electrode regions of interest (ROIs) revealed a significant effect for the frontal ROI ($b = 1.5645$, $SE = 0.4123$, $t = 3.795$) and the central ROI ($b = 1.1817$, $SE = 0.3510$, $t = 3.367$) but not at the posterior ROI ($t < 1$).

3.4.2 NP region

Neither visual inspection nor statistical analyses revealed significant effects at the determiner, the adverb or the noun. However, between 250 and 400 ms in the adjective region (*renovierte*), the incongruent condition elicited a larger negativity than the congruent condition following a causal as well as a concessive connector (Figures 5 and 6). The LMM analysis, collapsing over the nine channels, showed a significant effect of Congruency ($b = -0.4690$, $SE = 0.2023$, $t = -2.318$) but no effect
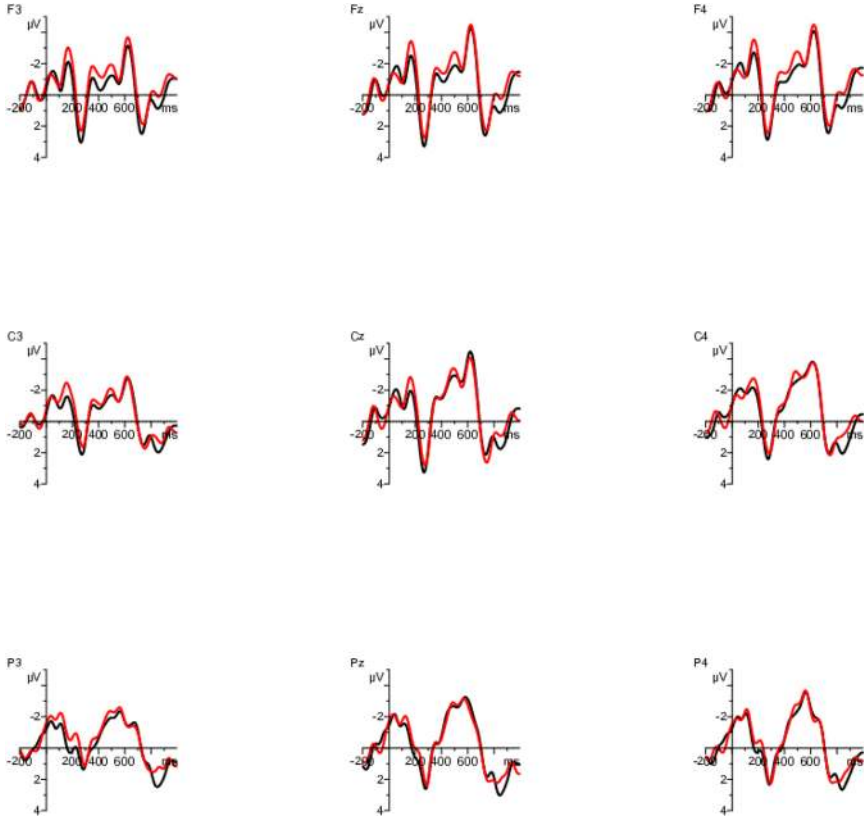
Figure 5: ERP effects on the adjective region (*renovierte*) following a causal connector at a subset of nine electrodes. Black line: congruent condition; red line: incongruent condition. Negativity is plotted upwards. For presentation purpose only, ERPs were filtered off-line with 10 Hz low pass.

of Connector Type ($t < 1$). Separate analyses of frontal, central, and posterior electrode ROIs revealed a significant effect of Congruency at the frontal ROI ($b = -0.9222$, $SE = 0.3709$, $t = -2.486$), but no effect of Connector Type ($t < 1$). No effects for Connector Type or Congruency were found at the central ($t < 1$) or posterior ($t < 1$) ROIs.

3.5 Discussion

The ERP analyses revealed a fronto-central positivity for the concessive connector compared to the causal connector. Possibly comprehenders need to update or
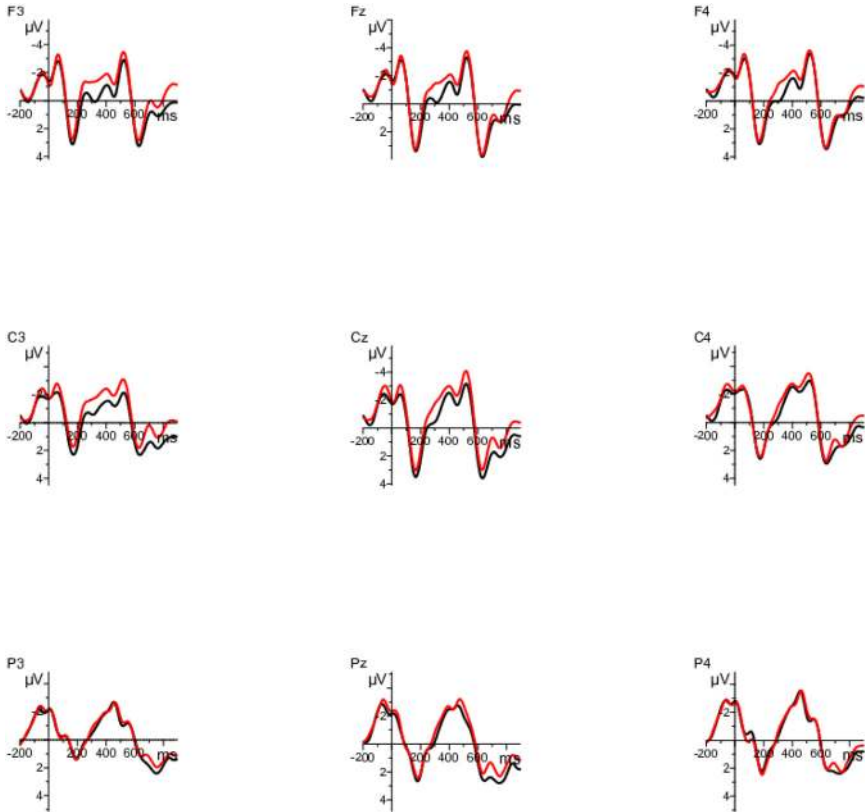
Figure 6: ERP effects on the adjective region (*renovierte*) following the concessive connector at a subset of nine electrodes. Black line: congruent condition; red line: incongruent condition. Negativity is plotted upwards. For presentation purpose only, ERPs were filtered off-line with 10 Hz low pass.

revise their representation of the discourse (e.g., Brouwer et al. 2012). One possible explanation is that their expectations for a causal discourse relation are violated (Kuperberg et al. 2011). This finding supports the claim that processing concessive connectors comes at the cost of a mental effort.

The frontal N400-like effect on the adjective region when the gender of the preceding determiner does not match with the gender of the predicted target noun is in line with van Berkum et al. (2005) and Wicha et al. (2004) (but see Koornneef 2021, for differing results when reading is not self-paced). The effect is also consistent with the anticipatory looks to the target referents during the pretarget region observed in Experiment 1. Our findings thus further support the view that

comprehenders predict a very specific lexical item based on the connector type, even when the setting is less constrained than in the visual-world experiment. Again, integration cannot explain this effect since the adjective preceded the target noun. Note however, that we did *not* find any effect on the determiner. Given that the determiner is also gender marked, one might have expected to already observe an effect in this region. This could be due to slow integration (it just takes a little bit of time to combine the bottom-up gender marking information with the anticipated target and show a mismatch effect), while at the adjective, more time has passed to show a mismatch effect.

4 Experiment 3

The goal of Experiment 3 was to replicate the results from Experiment 2 in English. Since English has no grammatical gender, we tested for predictive processing by exploiting the a/an indefinite article phonological alternation, as done by DeLong et al. (2005; see also van Berkum et al. 2005; Wicha et al. 2004). Given that the pretarget region was therefore also much shorter than in the German case, Experiment 3 was also a second evaluation of the time course of the facilitating effect of causal versus concessive connectors.

4.1 Method

4.1.1 Participants

Fourteen students (mean age 26, eight male) from Saarland University took part in the experiment. All participants were native speakers of English, had normal or corrected-to-normal vision, and were paid for their participation.

4.1.2 Materials and design

We created 96 experimental passages in English like in Example (4).

(4)   *Mr. Brown was planning to look for new glasses and shoes today. The glasses really are more urgent. [Therefore/However]$_{connector}$, he now heads towards [an/a]$_{pretarget}$ [optician/shoe shop]$_{target}$ that a friend recommended.*

The target region contained a noun beginning with a vowel versus a consonant and the pretarget region contained the phonological appropriate indefinite determiner (*an* vs. *a*). Similar to the gender-marking manipulation in Experiment 2, the

prenominal determiner was used to test whether comprehenders rapidly use the connector together with the context to predict the upcoming target noun.

As in Experiment 2, we collected cloze probabilities of the target words in the causal and concessive conditions. The mean cloze probability for the congruent target words in the causal condition was 0.56 (SD = 0.27), while for the concessive condition the mean cloze probability was 0.52 (SD = 0.33). The incongruent target words were almost never produced as completions: the mean cloze probability for the causal condition was 0.04 (SD = 0.08), while for the concessive condition was 0.10 (SD = 0.13).

4.1.3 Procedure

The procedure was the same as in Experiment 2.

4.2 Predictions

Analogous to the German ERP experiment, we expect incongruent targets to elicit a larger negativity than congruent targets. If comprehenders can rapidly use information from the context and the connector to predict the upcoming noun, the congruency effect should be observed already at the determiner, when it does not match the phonological representation of the predicted noun. Furthermore, we expect to replicate the positivity for concessive connectors observed in Experiment 3.

4.3 EEG recording, data analysis and results

EEG recording and data analyses were the same as in Experiment 2.

4.3.1 Connector region

The ERP patterns to the connectors are displayed in Figure 7. Concessive connectors elicited a globally distributed positivity compared to causal connectors. The positivity starts at around 300 ms and lasts for approximately 300 ms.

The LMM analysis, collapsing over the nine channels revealed a significant effect (b = 1.1327, SE = 0.2565, t = 4.416) for the concessive connector compared to the causal connector.

A separate analysis of frontal (F3, FZ, F4), central (C3, CZ, C4), and posterior (P3, PZ, P4) electrode ROIs revealed a significant effect for the frontal (b = 1.6009,
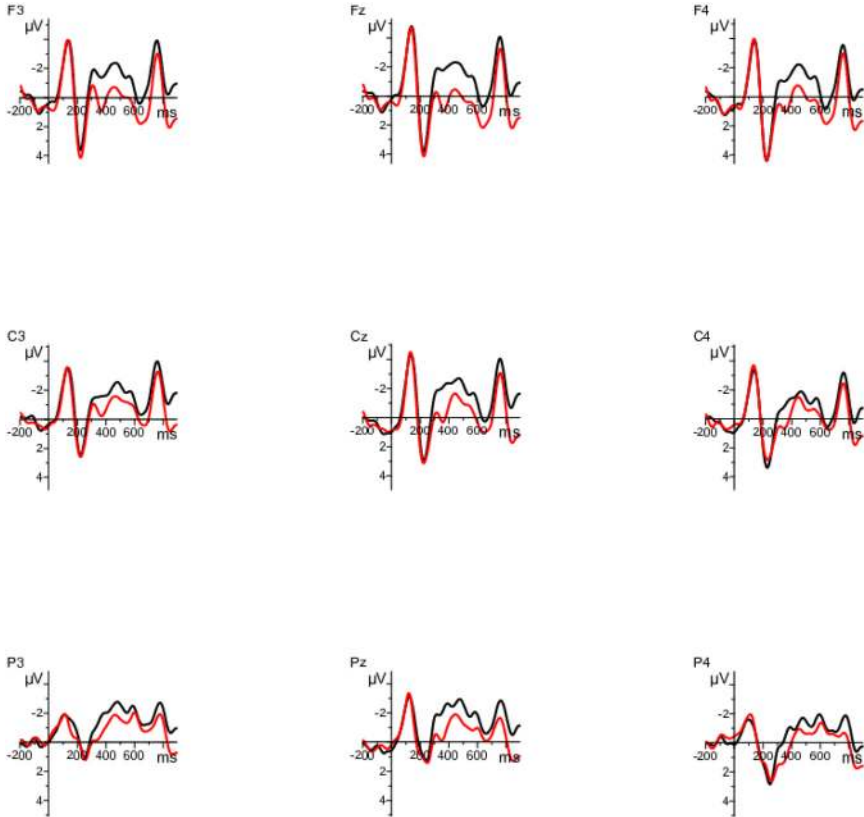
Figure 7: ERP effects on connectors. Black line: causal connector condition; red line: concessive connector condition. Negativity is plotted upwards. For presentation purpose only, ERPs were filtered off-line with 10 Hz low pass.

$SE = 0.4867$, $t = 3.29$), central ($b = 0.9133$, $SE = 0.4299$, $t = 2.125$) and the posterior ROI ($b = 0.8839$, $SE = 0.3995$, $t = 2.212$).

4.3.2 NP region

Visual inspection as well as statistical analyses revealed no significant effects on the determiner. On the target noun, however, the incongruent condition elicited a larger negativity between 400 and 600 ms compared to the congruent condition (Figures 8 and 9).

The LMM analysis, collapsing over the nine channels showed a significant effect of Congruency ($b = -1.0118$, $SE = 0.4120$, $t = -2.456$) but no effect of Connector
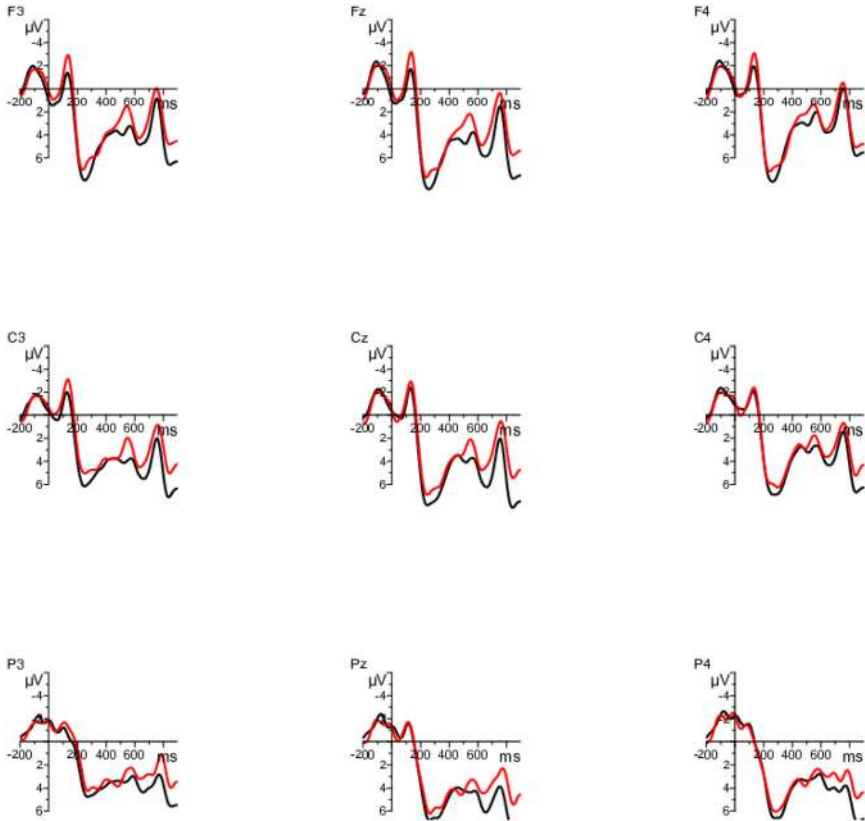
Figure 8: ERP effects on the noun region following the causal connector at a subset of nine electrodes. Black line: congruent condition; red line: incongruent condition. Negativity is plotted upwards. For presentation purpose only, ERPs were filtered off-line with 10 Hz low pass.

Type ($t < 1$). A separate analysis of frontal, central, and posterior electrode ROIs showed significant effects of Congruency at frontal electrodes ($b = -1.0118$, $SE = 0.4120$, $t = -2.456$), central electrodes ($b = -1.1921$, $SE = 0.4288$, $t = -2.780$) and posterior electrodes ($b = -1.2292$, $SE = 0.5440$, $t = -2.260$). No effects of Connector Type were found in any of the ROIs ($t < 1$).

4.4 Discussion

The ERP analyses revealed a significantly larger positivity for the concessive connector compared to the causal one, replicating the model updating effect on the
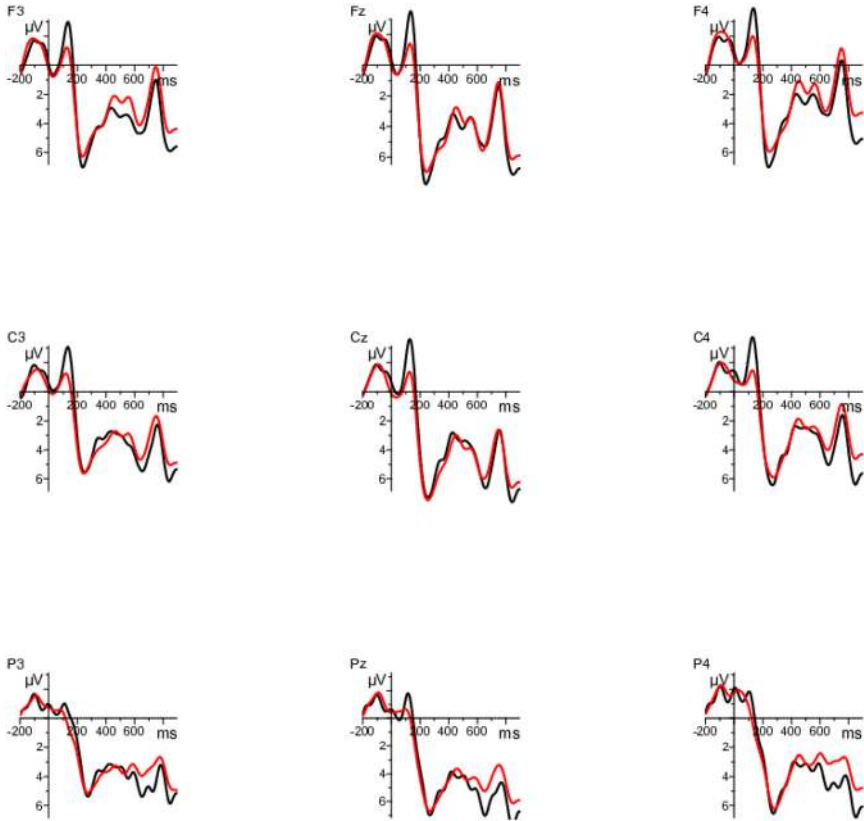
Figure 9: ERP effects on noun region following a concessive connector at a subset of nine electrodes. Black line: congruent condition; red line: incongruent condition. Negativity is plotted upwards. For presentation purpose only, ERPs were filtered off-line with 10 Hz low pass.

connector observed in Experiment 2 on German. This provides additional support for the immediate reconstruction of representation and reversal of expectations following the concessive connective.

We furthermore found an N400-like effect on the noun region for the incongruent compared to the congruent condition, following both causal and concessive connectors, showing that participants incrementally integrated discourse context and connector with upcoming content. Consistent with results in Experiment 2, we did not find any significant effect on the prenominal determiner. This suggests that predictive processing in our experiments needs some time to manifest – since in the English items there was no adjectival region between the determiner and the noun, the effect appeared on the noun itself.

While overall, results from the two ERP experiments reveal that discourse connectives are processed incrementally and help anticipating upcoming content, the immediacy of the facilitating predictive effect of both causal and concessive discourse did not affect processing of the determiner, even if that determiner was at odds with expectations. One might argue that in the English case, congruency with the determiner only holds if one assumes that the noun follows the determiner immediately and that this congruency effect should be less reliable because other material like adjectives is often encountered in this position. This argument does however not hold for the German determiner that is incongruent independent of other intervening material.

5 Conclusions and general discussion

We investigated the time course of processing in causal and concessive discourse relations within three experiments. Taken together, the presented experiments contribute to the four research questions presented in the introduction in the following way:

1. The first question was whether we could replicate previous findings in that connectives facilitate processing of downstream material. Our studies showed clear evidence corroborating this idea: In the visual-world study (Experiment 1), both causal and concessive connectors caused listeners to focus their attention to objects that were consistent with the discourse connective. Exp. 2 supports the idea that both causal and concessive connectives can be integrated rapidly enough to enable the reader to predict the target already in the pretarget region also in the absence of a constraining visual context (N400 effect on the gender-marked adjective preceding the head noun). This result is consistent with previous findings on predictive processing (e.g., van Berkum et al. 2005).

2. Our second research question was whether any such facilitating effect could be attributed unambiguously to prediction instead of integration. Experiments 1 and 2 (for both causal and concessive) are only consistent with prediction, but not with an integration view of facilitation: visual attention in the first experiment showed evidence of reversed expectations substantially before the occurrence of the target region, and the N400 effect in the second experiment was found on the gender-marked adjective, which precedes the mismatching noun.

3. Our third question was whether concessives cause higher processing difficulty than causals. Our studies showed different indications for concessive markers causing a processing difficulty. In Experiment 1, comprehension questions

were answered less accurately in the concessive case when the correct answer was "yes". This indicates that a reactivation of the content of the sentence containing the concessive can be difficult. A potential reason for the difficulty could be the reversal of discourse expectations, which may potentially cause both representations to still be active in memory (cf. good enough processing effects, Ferreira and Patson 2007; Slattery et al. 2013). This might be a short-lived effect in the case of our stimuli: comprehension questions were raised immediately after processing the sentence. It would be an interesting question how memory after a longer period of time would be affected.

4. In the EEG experiments, we found a late positivity on the concessive connector compared to the causal connector in both German (Experiment 2) and English (Experiment 3). We interpret this late positivity to reflect processes related to the updating, revision, or verification of a discourse model (e.g., Brouwer et al. 2012; Donchin and Coles 1988; van Herten et al. 2005; see also Arbel et al. 2011 for an interpretation of the P600 as an instantiation of the P300 component), or a more general process of pragmatic reanalysis (Drenhaus et al. 2006, 2011; van Herten et al. 2005; Xiang et al. 2009, see also, DeLong et al. 2014; Wlotko and Federmeier 2012).

5. The EEG results were remarkably consistent across languages: even though the materials were different, we found very consistent effects on the concessive connectors in English and German.

6. The results that processing concessives was demanding for comprehenders is also in line with the idea that concessive discourse markers are a type of negation (König and Siemund 2000). We did not find any indication of a processing delay, however.

7. Our final question concerned the issue whether connectives, similar to negation, give rise to a search for alternatives (i.e., negation was used as a direct cue to anticipation, see Staab 2007). Experiment 1 supports this hypothesis: In the extended connector region in concessive sentences, listeners started redirecting their looks (e.g., from the two sweet objects to the two salty ones). This supports the idea that concessives can be considered negative causals: As soon as the comprehender assumes that the so-far expected state of affairs needs to be reversed, he or she tries to find alternatives. This of course is a rather easy task in a visual world with few objects and it is still when different alternatives have been mentioned linguistically, as in Experiments 2 and 3. Further research is needed to explore what this search for alternatives looks like in scenarios that are more complex.

A further interesting finding in our study, when putting the experiments in relation to the results reported in Xiang and Kuperberg (2015), is that the N400 effect was

similarly large for the causal and concessive conditions in our experiments, while the N400 effect was actually *larger* in the concessive condition than in the unmarked condition in Xiang and Kuperberg's experiment. Xiang and Kuperberg interpret this as meaning that constrained contextual *content* is not the only factor to trigger active prediction. Strong communicative cues – in the current case *even so* – can put people in a "predictive mode", even when the specific content itself is not highly constraining. and remark that it will be important in future research to determine whether the effect generalizes beyond *even so* to other concessive connectives.

We see several aspects that differ between the studies reported here versus those in Xiang and Kuperberg that can potentially explain this discrepancy in findings: firstly, our causal condition is marked with the connective *because*, whereas the causal condition in Xiang and Kuperberg (2015) is not marked with a connective. The connective *because* may hence strengthen predictions. Furthermore, the two studies used different concessive connectives. While Xiang and Kuperberg evaluated *even so*, our Experiment 3 concentrated on the connective *however*. These connectives differ in their scopes, in that *even so* takes a narrower scope regarding what it may be contrasting with, and may hence lead to stronger predictions than *however*. Knott, in his 1996 thesis (pp. 185–186), observes that *however* cannot always be substituted by *even so*, but that *even so* can always be substituted by *however*. The example he gives is *John was starving. However/But/#Nevertheless/#Even so, there was no food in the house.* A related study by Köhne and Demberg (2013) reports an eye-tracking-while-reading study using the experimental materials from the second experiment reported here. This experiment uses the connectives *dennoch* and *trotzdem* in the concessive condition (for half of the items each). The connective *dennoch* is less constraining and behaves similar to *however*, while *trotzdem* is similar to the English *even so*. Köhne and Demberg (2013) speculate that their failure to find a significant difference between the match and mismatch conditions in concessive connectives may be due to the lower constraint of connectives like *however* compared to *even so*. The findings regarding the difference in strength of the N400 effect reported here versus in Xiang and Kuperberg (2015) is thus consistent with this interpretation.

We therefore conclude here that concessives are not *generally* more constraining or more likely to elicit predictions than other discourse markers, but that connectives inside the same class of connectives can exhibit different degrees of constraints. These may then be evident in the size of the observed reduction in the N400 as a function of connective based anticipation. It is a common phenomenon that, in general, discourse particles of the same class are not always interchangeable because they subtly differ in meaning (e.g., regarding communicative knowledge, see van Bergen and Hoogeweg 2021), which means that they should also elicit differing expectations.

To summarize, the data of Experiments 1–3 presented here provide us with a consistent picture about how top-down and bottom-up factors together shape discourse expectations: top-down predictions from previous discourse context were updated in the light of bottom-up information from a discourse connective.

List of abbreviations

F	feminine
INF	infinitive
M	masculine
VPART	verb particle
REFL	reflexive

References

Altman, Gerry T. M. & Yuki Kamide. 1999. Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition* 73. 247–267.

Arbel, Yael, Kevin M. Spencer & Emanuel Donchin. 2011. The N400 and the P300 are not all that independent. *Psychophysiology* 48(6). 861–875.

Asr, Fatemeh T. & Vera Demberg. 2012. Implicitness of discourse relations. In Martin Kay & Christian Boitet (eds.), *COLING 2012: Technical papers*, 2669–2684. Mumbai: Indian Institute of Technology Bombay.

Baayen, R. Harald, Douglas J. Davidson & Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59. 390–412.

Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278.

Bates, Douglas M. 2005. Fitting linear mixed models in R. *R News* 5. 27–30.

Bates, Douglas M. & Deepayan Sarkar. 2007. lme4: Linear mixed-effects models using S4 classes. R package version 0.99875-6.

Brehm-Jurish, Eva U. 2005. *Connective ties in discourse: Three ERP studies on causal, temporal, and concessive connective ties and their influence on language processing*. Potsdam: Potsdam University Dissertation.

Brouwer, Harm, Hartmut Fitz & John C. J. Hoeks. 2012. Getting real about semantic illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research* 1446. 127–143.

Canestrelli, Anneloes R., Willem M. Mak & Ted J. M. Sanders. 2013. Causal connectives in discourse processing: How differences in subjectivity are reflected in eye movements. *Language and Cognitive Processes* 28(9). 1394–1413.

Carpenter, Patricia A. & Marcel Just. 1975. Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review* 82(1). 45–73.

Cozijn, Reinier, Leo G. M. Noordman & Wietske Vonk. 2011. Propositional integration and world-knowledge inference: Processes in understanding because sentences. *Discourse Processes* 48(7). 474–500.

Dale, Rick & Nicholas D. Duran. 2011. The cognitive dynamics of negated sentence verification. *Cognitive Science* 35(5). 983–996.

Delogu, Francesca, Heiner Drenhaus & Matthew W. Crocker. 2018. On the predictability of event boundaries in discourse: An ERP investigation. *Memory and Cognition* 46(2). 315–325.

DeLong, Katherine A., Laura Quante & Marta Kutas. 2014. Predictability, plausibility, and two late ERP positivities during written sentence comprehension. *Neuropsychologia* 61(1). 150–162.

DeLong, Katherine A., Thomas P. Urbach & Marta Kutas. 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience* 8(8). 1117–1121.

Donchin, Emanuel & Michael G. Coles. 1988. Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences* 11(3). 357–427.

Drenhaus, Heiner, Malte Zimmermann & Shravan Vasishth. 2011. Exhaustiveness effects in clefts are not truth-functional. *Journal of Neurolinguistics* 24(3). 320–337.

Drenhaus, Heiner, Peter Beim Graben, Saddy Douglas & Stefan Frisch. 2006. Diagnosis and repair of negative polarity constructions in the light of symbolic resonance analysis. *Brain and Language* 96. 255–268.

Federmeier, Kara D. & Marta Kutas. 1999. A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language* 41(4). 469–495.

Ferguson, Heather J., Anthony J. Sanford & Hartmut Leuthold. 2008. Eye-movements and ERPs reveal the time course of processing negation and remitting counterfactual worlds. *Brain Research* 1236. 113–125.

Ferreira, Fernanda & Nikole D. Patson. 2007. The 'good enough' approach to language comprehension. *Language and Linguistics Compass* 1(1/2). 71–83.

Haberlandt, Karl. 1982. Reader expectations in text comprehension. In Jean F. Le Ny & Kintsch Walter (eds.), *Language and language comprehension*, 239–249. Amsterdam: North-Holland.

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

Knott, Alistair. 1996. *A data-driven methodology for motivating a set of coherence relations*. Edinburgh: The University of Edinburgh Dissertation.

Köhne, Judith & Vera Demberg. 2013. The time-course of processing discourse connectives. In Markus Knauff, Markus Pauen, Natalie Sebanz, & Ipke Wachsmuth (eds.), *Proceedings of the 35th annual conference of the Cognitive Science Society (CogSci 2013, Berlin)*, 2760–2765. Austin, TX: Cognitive Science Society.

König, Ekkehard & Peter Siemund. 2000. Causal and concessive clauses: Formal and semantic relations. *Topics in English Linguistics* 33. 341–360.

Koornneef, Arnout. 2021. The processing signature of anticipatory reading: an eye-tracking study on lexical predictions. *Linguistics* 59(2). 449–479.

Kuperberg, Gina R. & T. Florian Jaeger. 2015. What do we mean by prediction in language comprehension? *Language Cognition & Neuroscience* 3798. 1–70.

Kuperberg, Gina R., Tatiana Sitnikova, David Caplan & Phillip J. Holcomb. 2003. Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research* 17(1). 117–129.

Kuperberg, Gina, Paczynski Martin & Tali Ditman. 2011. Establishing causal coherence across sentences: An ERP study. *Journal of Cognitive Neuroscience* 23(5). 1230–1246.

Kutas, Marta & Kara D. Federmeier. 2011. Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology* 62. 621–647.

Kutas, Marta & Steven A. Hillyard. 1984. Brain potentials during reading reflect word expectancy and semantic association. *Nature* 307(5947). 161–163.

Kutas, Marta, Katherine A. DeLong & Nathaniel J. Smith. 2011. A look around at what lies ahead : Prediction and predictability in language processing. In Moshe Bar (ed.), *Predictions in the brain: Using our past to generate a future*, 190–207. Oxford: Oxford University Press.

Laszlo, Sarah & Kara D. Federmeier. 2011. The N400 as a snapshot of interactive processing: Evidence from regression analyses of orthographic neighbor and lexical associate effects. *Psychophysiology* 48(2). 176–186.

Levy, Roger & T. Florian Jaeger. 2007. Speakers optimize information density through syntactic reduction. In Bernhard Schölkopf, John Platt & Thomas Hoffman (eds.), *Twentieth annual conference on neural information processing systems*, 849–856. Cambridge, MA: MIT Press.

Louwerse, Max. 2001. An analytic and cognitive parametrization of coherence relations. *Cognitive Linguistics* 12(3). 291–316.

Lüdtke, Jana, Claudia K. Friedrich, Monica De Filippis & Barbara Kaup. 2008. Event-related potential correlates of negation in a sentence-picture verification paradigm. *Journal of Cognitive Neuroscience* 20(8). 1355–1370.

Marslen-Wilson, William. 1973. Linguistic structure and speech shadowing at very short latencies. *Nature* 244. 522–523.

Millis, Keith K. & Marcel A. Just. 1994. The influence of connectives on sentence comprehension. *Journal of Memory and Language* 33. 128–147.

Murray, John D. 1995. Logical connectives and local coherence. In Robert F. Lorch Jr. & Edward J. O'Brien (eds.), *Sources of coherence in reading*, 107–125. Hillsdale, NJ: Lawrence Erlbaum.

Nieuwland, Mante & Gina R. Kuperberg. 2008. When the truth is not too hard to handle an event-related potential study on the pragmatics of negation. *Psychological Science* 19(12). 1213–1218.

Nordmeyer, Ann E. & Michael C. Frank. 2014. A pragmatic account of the processing of negated sentences. In Paul Bello, Marcello Guarini, Marjorie McShane & Brian Scassellati (eds.), *Proceedings of the 36th annual conference of the Cognitive Science Society (CogSci 2014, Quebec City)*, 2699–2704. Austin, TX: Cognitive Science Society.

Osterhout, Lee & Phillip J. Holcomb. 1995. Event-related potentials and language comprehension. In Michael D. Rugg & Michael G. H. Coles (eds.), *Electrophysiology of mind (Oxford psychology series 25)*, 171–215. Oxford: Oxford University Press.

Sanders, Ted J. M. 2005. Coherence, causality and cognitive complexity in discourse. In Michael Aumague, Myriam Bas, Anne Le Draoulec & Laure Vieu (eds.), *Proceedings of the first international symposium on the exploration and modelling of meaning (SEM-05)*, 105–114. Toulouse: Université de Toulouse le Mirail.

Sanders, Ted J. M. & Leo G. M. Noordman. 2000. The role of coherence relations and their linguistic markers in text processing. *Discourse Processes* 29(1). 37–60.

Sanders, Ted J. M., Wilbert P. M. Spooren & Leo G. M. Noordman. 1992. Toward a taxonomy of coherence relations. *Discourse Processes* 15(1). 1–35.

Scholman, Merel C. J., Hannah Rohde & Demberg Vera. 2017. 'On the one hand' as a cue to anticipate upcoming discourse structure. *Journal of Memory and Language* 97. 47–60.

Sharbrough, Frank, Gian E. Chartrian, Ronald P. Lesser, Hans O. Lüders, Mark R. Nuwer & Terence W. Picton. 1995. American Electroencephalographic Society guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology* 8. 200–202.

Slattery, Timothy J., Patrick Sturt, Kiel Christianson, Masaya Yoshida & Fernanda Ferreira. 2013. Lingering misinterpretations of garden path sentences arise from competing syntactic representations. *Journal of Memory and Language* 69(2). 104–120.

Staab, Jenny. 2007. *Negation in context: Electrophysiological and behavioral investigations of negation effects in discourse processing*. San Diego/San Diego State University: University of California Dissertation.

Tanenhaus, Michael K. & John C. Trueswell. 1995. Sentence comprehension. In Joanne Miller & Peter D. Eimas (eds.), *Handbook of perception and cognition, vol. 11: Speech, language, and communication*, 2nd edn., 217–262. San Diego, CA: Academic Press.

Townsend, David J. 1983. Thematic processing in sentences and texts. *Cognition* 13(2). 223–261.

Traxler, Matthew J., Anthoony J. Sanford, Joy P. Aked & Linda M. Moxey. 1997b. Processing causal and diagnostic statements in discourse. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23(1). 88–101.

Traxler, Matthew J., Michael D. Bybee & Martin J. Pickering. 1997a. Influence of connectives on language comprehension: Eye tracking evidence for incremental interpretation. *The Quarterly Journal of Experimental Psychology: Section A* 50(3). 481–497.

van Bergen, Geertje & Lotte Hoogeweg. 2021. Managing interpersonal discourse expectations: a comparative analysis of contrastive discourse particles in Dutch. *Linguistics* 59(2). 333–360.

van Berkum, Jos J. A., Colin M. Brown, Pienie Zwitserlood, Valesca Kooijman & Hagoort Peter. 2005. Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31(3). 443–467.

van Herten, Marieke, Herman H. J. Kolk & Dorothee J. Chwilla. 2005. An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research* 22(2). 241–255.

Wicha, Nicole Y. Y., Eva M. Moreno & Marta Kutas. 2004. Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience* 16. 1272–1288.

Wlotko, Edward W. & Kara D. Federmeier. 2012. So that's what you meant! Event-related potentials reveal multiple aspects of context use during construction of message-level meaning. *NeuroImage* 62(1). 356–366.

Xiang, Ming & Gina R. Kuperberg. 2015. Reversing expectations during discourse comprehension. *Language, Cognition and Neuroscience* 30(6). 37–41.

Xiang, Ming, Brian Dillon & Colin Phillips. 2009. Illusory licensing effects across dependency types: ERP evidence. *Brain and Language* 108(1). 40–55.