To appear in: A. Goel, C. Seifert, & C. Freska (2019, eds.), *Proceedings of the 41st Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Processing of affirmation and negation in contexts with unique and multiple alternatives: Evidence from event-related potentials

Maria Spychalska[1,2] **(mspychal@uni-koeln.de)** & **Viviana Haase**[2] **(viviana.haase@rub.de)**

Jarmo Kontinen[2] **(kontinen.jarmo@gmail.com)**

Markus Werning[2] **(markus.werning@rub.de)**

1. Department of German Linguistics, University of Cologne, Germany
2. Institute of Philosophy II, Ruhr University Bochum, Germany

Abstract

We employ a scenario-sentence-verification paradigm to investigate the role of scenario-given alternatives for the processing of affirmative and negative sentences. We show that for both affirmative and negative sentences the N400 amplitude is larger if the context model provides multiple alternatives for a true sentence continuation relative to the case when it provides only a unique referent. Additionally, we observe a late positivity effect for negative relative to affirmative sentences, independent of the context model.

Keywords: Negation; alternatives; N400; P600

Introduction

Negation occurs in every human language and is essential for communication. Nevertheless, negative sentences seem to require elevated processing resources when compared to affirmative ones and negation has posed a challenge to psycholinguistic theories. Language comprehension is generally considered to happen incrementally, i.e. as the linguistic input unfolds in real time; yet, empirical evidence suggests that negation is not compatible with this view, and it is argued that negation may be integrated into sentence meaning only at a later stage of the comprehension process. Furthermore, language processing is considered to be predictive, meaning that we not only process incoming linguistic information but also anticipate upcoming content (cf. DeLong et al., 2005; Pickering & Garrod, 2007; Van Petten & Luka, 2012; Kuperberg & Jaeger, 2015). However, despite of the large amount of experimental literature on prediction, it remains an open question to what extent and at which processing stage the number of contextually available alternatives for an upcoming word modulates the prediction of that word during sentence comprehension. Furthermore, as it is unknown at which stage negation becomes a part of compositional meaning, it is also not clear whether and how negation modulates predictive processing.

In the psycholinguistic literature emphasis has been put on providing a model of negation processing and experimental work has focused on investigating the cognitive costs related to the processing of sentences with negation. Since negative sentences are structurally more complex, they are expected to involve more cognitive resources than their affirmative counterparts. From a semantic point of view, the potential need to suppress positive information is also likely to result in an increase of processing costs. Early on it has been shown that negative sentences are associated with higher error rates as well as longer response and reading times than affirmative sentences (Just & Carpenter, 1971; Clark & Chase, 1972; Lüdtke & Kaup, 2006; Dale & Duran, 2011).

Early electroencephalography (EEG) studies on negation processing further suggested that the integration of negation into the compositional sentence meaning is delayed. For instance, many studies (Fischler et al., 1983; Dudschig et al., 2019) found no effect of negation on the N400 event-related potential (ERP) component. The N400 is a negative shift in the ERP waveform, of a latency between 200 and 600 ms post-stimulus onset and maximal over centro-parietal scalp sites. Its amplitude tends to be larger for words that are (semantically) less expected given the background context, or world-knowledge, as well as for words of lower corpus-based frequency (see Kutas & Federmeier (2011) for an overview). The N400 has furthermore been shown to be inversely correlated with the triggering word's cloze probability, i.e. the percentage of individuals who would continue a given sentence fragment with that word (Federmeier et al., 2007). Various theories interpret the N400 as a marker of (i) lexical retrieval (Brouwer & Hoeks, 2013), (ii) integration into the prior context (Hagoort et al., 2004), (iii) predictive preactivation (DeLong et al., 2005), or even (iv) meaning-related probabilistic prediction (Lau et al., 2013; Kuperberg & Jaeger, 2015; Rabovsky et al., 2018). According to a recent account (Rabovsky et al., 2018), the N400 reflects meaning-related prediction error, where prediction is understood in a non-intentional sense, as an implicit state of the system that is tuned to anticipate the upcoming input in a graded manner.

In the first ERP study on negation by Fischler et al. (1983), the N400 was only modulated by the lexical-semantic relation of the predicate to the main noun and it was larger for the sentence-final predicates when this relation was weak, as for instance in *A robin is/is not a truck* relative to the case when the relation was strong, as in *A robin is/is not a bird*, independently of the presence of negation in a sentence, which reversed the sentence truth-value. This result is usually interpreted as evidence that negation is not processed incrementally and thus is not immediately integrated into the compositional sentence meaning. This interpretation was further supported by an EEG-study employing a sentence-picture verification paradigm that revealed a delayed integration of the negation in the sentence meaning (Lüdtke et al., 2008). In their experiment, affirmative and negative sentences, such as *In front of the tower there is a/no ghost*, were followed by matching or mismatching pictures, e.g. a tower with a lion or a tower with a ghost, after a short (250*ms*) or a long (1500*ms*) delay. Note that in the case of the affirmative sentences, the matching pictures are explicitly mentioned and thus primed by the sentences, but the negative sentences primed the mismatching pictures. In the short-delay condition, the N400 ERPs reflected a priming effect, namely, for the affirmative sentences the mismatching pictures were associated with a larger N400 than the matching pictures, whereas for the negative sentences the effect was opposite. An effect of negation was only reflected by a late positivity effect that was identified as the P600 effect. In contrast, for the long-delay condition, main effects of truth-value and negation in addition to the priming effect were already observed in the N400 time-window. This result was taken as evidence that integrating negation into the sentence meaning required additional time after the sentence was read. The P600 effect observed in response to the use of negation is an especially noteworthy result. The P600 is a slow, late (around 500-800 ms post-onset) positive shift in the ERP waveform that is maximal over posterior scalp sites. It is often observed for structural violations, grammatical errors or syntactically more complex sentences (Hagoort et al., 1993) but also for some pragmatic and semantic anomalies (Kuperberg et al., 2003). It has been argued to reflect combinatorial aspects of linguistic processing (Kuperberg, 2007) or even semantic integration mechanisms (Brouwer et al., 2012). In the case of negative sentences it may be taken as a marker of the increased processing demands related to integrating the negation into sentence meaning.

However, the comprehension of negated sentences is facilitated if they are embedded into context. Nieuwland & Kuperberg (2008) showed that pragmatically licensed negative sentences such as for example *With proper equipment, scuba diving isn't very safe/dangerous...* did not lead to elevated N400-components for true compared to false sentences. Instead, without pragmatic licensing, e.g. *Bulletproof vests aren't very safe/dangerous...* the true negated sentences led to higher N400s than the false sentences, in line with Fischler et al. (1983). Tian et al. (2016) furthermore showed that for cleft-structures which narrow the scope of the negation and therefore the potential alternatives such as in, e.g. *It is John who hasn't ironed his shirt*, incremental comprehension of negated sentences is facilitated as well.

The role of alternatives for the processing of both negative and affirmative sentences was directly studied in an eye-tracking experiment by Orenes et al. (2014). They investigated whether the presence of multiple alternatives in the context has an effect on the processing of negative and affirmative sentences. An auditory context sentence was introduced that either indicated that all objects are possible choices (*multary* condition: *The figure could be red, green, blue, or yellow*), or restricted the choice to only two objects (*binary* condition: *The figure could be red or green*). Then, the visual context appeared, which always included four different objects, e.g. circles of different colors. The target sentence *The figure is red/not red* was presented auditorily while the four figures were shown on the screen and eye-movements were monitored. For affirmative target sentences subjects fixated the target object (*red circle*) in both context conditions. In contrast, for negative sentences, subjects fixated the target object (*green circle*) only in the *binary* condition, whereas in the *multary* condition they fixated the object that had the mentioned, negated feature (*red circle*). The interpretation of these results is problematic, since the affirmative and negative conditions were not logically comparable. Whereas affirmative sentences directly mentioned the target object, for negative ones the identification of the target was only possible in the *binary* condition, where one could infer the color of the target object from the pair of the context and the target sentence. In the *multary* condition the target object was not identifiable: The intended figure that was described as *not red*, could be blue, yellow, or green.

It is not surprising that in the case when the referent cannot be identified, the processing of a sentence differs relative to the case when the referent can be uniquely established in the context. However, the question arises whether this effect has anything specifically to do with negation. In our current experiment we aimed at directly comparing the processing of affirmative and negative sentences in situations when the context scenario provides multiple or unique potential true sentence continuations. Suppose that Julia is dealt three cards (*ace*, *king*, *queen*) and the game is to choose some of these cards, while rejecting the rest of them at the same time. If Julia selects one card (e.g. *ace*), one can describe her choice by saying that *Julia selected the ace*. There is only one card that can be mentioned in a true affirmative sentence. Additionally, about each of the remaining two cards one can say that it was not selected, e.g. *Julia did not select the king/queen*. In this case there are two potential true sentence continuations. The situation is exactly opposite if Julia had selected two cards, while rejecting only one. By manipulating a game situation of this type we can create unique and multiple contexts equivalent for negative and affirmative descriptions. It has previously been suggested that the N400 amplitude elicited by a given word seems to be directly dependent on the number of contextually given alternatives to this word. Spychalska et al. (2016) showed that, for sentences such as *Some pictures contain X*, if the context scenario provided additional objects *Y* that could be mentioned instead of *X*, the N400 was larger when *Y* would complete a true sentence (thus was a true alternative to *X*) relative to the case when *Y* would complete a false sentence (and was not a true alternative to *X*). Since the N400 is known to inversely correlate with the cloze probability of the triggering word, one can hypothesize that if the context scenario provides alternative referents for a true sentence condition, the N400 should be larger relative to the case when the context allows to uniquely predict the referent. By contrasting negative and affirmative sentences in logically equivalent conditions, the design allows us to directly measure the effect of negation on the processing of a referent in the context with multiple alternatives.

Method

The experimental design used a scenario-sentence verification paradigm. Participants were informed that they follow a player's moves in a game. In each target trial the player (introduced in the form of a clipart-like image) was dealt three cards, each depicting a different object, which were presented on the screen. Then, the player selected or rejected one or two of the cards. Subjects were informed that this action leads to an exhaustive divide of the set of cards. Selection was marked by framing the selected cards green, which implied that the unframed cards are rejected. Rejection was marked by framing the rejected cards red, which implied that the unframed cards are selected. After the cards were marked, the scene disappeared and a sentence (in German) was presented phrase-by-phrase. At the end of the trial, participants were asked whether the sentence is a true description of the action taken by the player. The target sentences were either of the form 1a (*affirmative* conditions) or of form 1b (*negative* conditions), where *X* is a proper name referring to the player and *Y* denotes the critical noun.

(1) a. *X hat den/die/das Y ausgewählt.*
X has chosen Y.

b. *X hat nicht den/die/das Y ausgewählt.*
X has not chosen Y.

In each target scenario there was only one object of the given type, thus a definite article was used in the sentence. All objects presented in a given trial were of the same grammatical gender to rule out that the noun in the sentence could be predicted based on the article. We ran two experiments: In **Experiment I**, all target sentences referred to one of the **unframed** objects. Thus, negative sentences followed scenarios with green frames, whereas affirmative sentences followed scenarios with red frames. In this way, both the affirmative and negative conditions required the participant to make an inference about the unmarked cards from the information provided visually (i.e. the marked set of cards). In the **unique** conditions, two out of three cards were framed, which left only one possible and therefore unique referent (unframed picture) to be named in the target sentence, whereas in the **multiple** conditions only one card was framed leaving two and hence multiple possible referents (unframed pictures) to be potentially named in the target sentence. In **Experiment II**, all target sentences referred to one of the **framed** cards. Thus, negative sentences followed scenarios with red frames, whereas affirmative sentences followed scenarios with green frames. In this experiment both negative and affirmative target sentences directly described the player's action and did not involve any inferential step. Both experiments used a

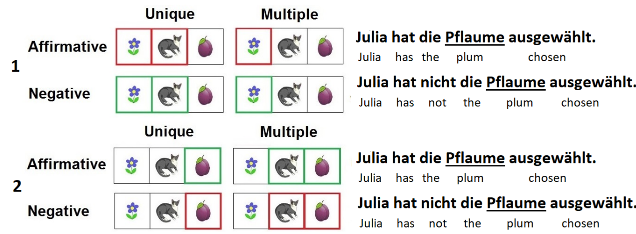

Table 1: Example for the conditions in Experiment 1 (top) and Experiment 2 (bottom)

2x2 design with the factors (i) *Alternatives* (unique/multiple); and (ii) *Polarity* (affirmative/negative), resulting in four conditions as shown in Table 1. [1]

Our main hypothesis was that the N400 recorded for the nouns referring to the target objects should be larger for the *multiple* relative to the *unique* context. This effect was in principle expected both for negative and affirmative sentences; however, we also hypothesized a possible interaction effect between *Polarity* and *Alternatives*, indicating that the processing of negative and affirmative sentences involves non-overlapping processes especially in the context where multiple alternatives are available. Furthermore, based on prior studies (Lüdtke et al., 2008), we hypothesized that the main effect of negation may occur in the late (P600) time-window.

Materials We created a list of 240 German nouns that are depictable and concrete. They were all mono- or bi-syllabic, moderately frequent[2], had a length between three to nine characters and were used in their singular form. The words were combined into 240 unique triples $\langle N_1, N_2, N_3 \rangle$, i.e. each word was used with each other word only once. These triples were used to generate scenarios presenting three different objects, assigned pseudo-randomly to experimental conditions, in such a way that each word was a critical noun only once. This resulted in 60 trials per condition, 240 target trials in total. To balance out the overall truth-value ratio and to make the material more variable we added 240 filler trials: 200 false and 40 true[3], based on a list of 84 new nouns and with affirmative and negative sentences evenly distributed. To rule out the possibility of creating expectations for negative or affirmative sentences based on the color of the frames, the framing of the filler's pictures exploited all possible options cross-balanced, so that affirmative/negative filler sentences followed red or green frames the same number of times.

Procedure Upon arrival participants signed an informed consent of participation. They were given a written instruction including few examples and completed an exercise ses-

[1]The framing was realized as a between-subject factor in order to have a sufficient number of trials per condition (60 in the current design) without inflating the length of the experiment. In addition, false filler sentences were needed to balance the materials (see below).

[2]The logarithmic frequency of all stimuli words was controlled with the use of Leipzig Wortschatz http://wortschatz.uni-leipzig.de/

[3]Overall, 41.6% of the trials in the experiment were false.

sion consisting of eight trials. Feedback was provided for the exercise to make sure that participants understood the task, especially the meaning of the framing and the exhaustive divide of the set of cards. The experiment comprised of eight blocks lasting approximately 10 minutes, with optional breaks in between.[4] The whole experiment lasted approximately 90 minutes. Each trial started with the presentation of three pictures in the center of the screen. The pictures were first shown without any frame. Subsequently, one or two of the objects were framed green or red, followed by the sentence that was displayed phrase-by-phrase on the screen (see Figure 1). The main ERP trigger was the noun referring to (one of) the selected or rejected target object(s). After each sentence, subjects had to respond to a truth-value judgment task by pressing a left- or a right-hand button on a response pad. The buttons were assigned pseudo-randomly by displaying ”TRUE” and ”FALSE” on the screen sides.

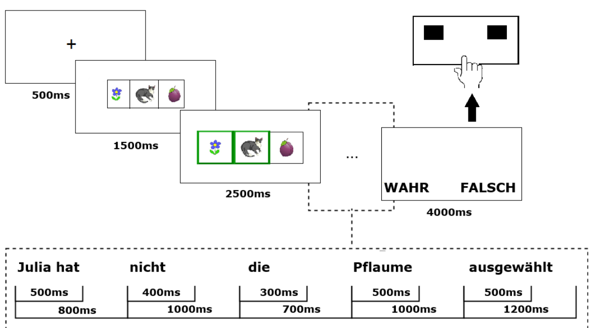

Figure 1: The time course of an example trial with presentation times. The main ERP trigger is the word *Pflaume* (*plum*).

EEG recording and data preprocessing The EEG was recorded with a BrainAmp acticap 64 channel recording system. Electrode position AFz was used as ground and FCz as physical reference. The electro-oculogram (EOG) was measured with four electrodes (PO09, PO10, FT9, FT10), which were reprogrammed and placed above and below the right eye and on both temples. Electrode impedances were kept below 5kΩ. The EEG was recorded with a sampling rate of 500Hz, a 10 sec low cut-off filter and an anti-aliasing hardware filter. The EEG-data was analyzed in the BrainVision Analyzer 2.0 software. We applied a 0.1-30 Hz off-line bandpass filter. All trials with an absolute amplitude difference over $200\mu V/200ms$ or with an activity lower than $0.5\mu V$ in intervals of at least $100ms$ were automatically rejected. The maximal voltage step that was allowed was $50\mu V/200ms$. Eyeblinks were corrected by means of an independent component analysis. The data was re-referenced to the averaged mastoids (TP9, TP10). The baseline-correction was done based on the $200ms$ pre-onset interval of the stimulus. Segments with any remaining physical artifacts (lower than -90μV or higher than 90μV) were removed before averaging. At least 50% of segments in each condition were preserved.

[4]For each block a player was introduced. In total, four different players (two female) appeared during one experiment, each assigned to two of the eight blocks.

Statistical analysis of the ERPs following the onset of the critical noun For the analysis of the ERPs we used a repeated measures ANOVA with the factors *Polarity* (negative/affirmative), *Alternatives* (multiple/unique) and *Region* (anterior/posterior).[5] The anterior ROI covered frontal, anterior frontal, frontal parietal, frontal temporal as well as frontal central regions of both hemispheres. The posterior ROI reached across temporal posterior, central posterior, posterior, posterior occipital and occipital regions of both hemispheres. The electrodes from the horizontal midline (central electrodes) were analyzed separately. We analyzed the averaged subjects' ERPs in two time-windows: 250-550 ms post-stimulus onset for the N400 effect and 550-850 ms post-onset for the P600. The assumptions of parametric data (e.g. normal distribution) were met.

Results

Twenty-five volunteers participated in Experiment 1 (nine male; mean age 25.2 ($SD = 4.42$, range $18-33$). In Experiment II we measured twenty-five new (not participating in Experiment I) volunteers (nine male; mean age 23.88 years (SD 3.94), age range 18-33). We excluded one subject per experiment from the analyses due to excessive artifacts in the EEG-data.

Behavioral results. Accuracy was at ceiling level in all conditions, indicating that the task was not too difficult for the subjects (see Table 2). Although minor differences are observed across conditions, due to space limitations, the statistical analysis of the behavioral data is omitted in the paper.

		Unique	**Multiple**
Experiment I	**Affirmative**	97.22(3.25)	94.17(7.17)
	Negative	95.63(5.83)	95.42(4.12)
Experiment II	**Affirmative**	97.01(2.82)	95.26(5.00)
	Negative	95.34(4.14)	95.27(4.74)

Table 2: Mean accuracy in Experiment I and II, in percentages, and the standard deviation ($\mu(\sigma)$) for all conditions

Polarity independent modulation of the N400 by alternatives. The visual inspection of the grand averages revealed that critical nouns in the multiple conditions elicited clearly larger N400 ERPs than critical nouns in the unique conditions for both sentence polarities and in both experiments. The ANOVA for the time-window **250-550 ms** for **Experiment I**, revealed a main effect of *Alternatives* ($F(1,23) = 30.040$, $p<.001$, $\eta^2 = .566$), with the mean difference between the multiple and unique conditions of $-2.157\mu V$ ($M_{mult} = -.59\mu V$, $M_{unq} = 1.567\mu V$). There was also a main effect of *Polarity* ($F(1,23) = 10.854$, $p = .003$, $\eta^2 = .321$), namely, the negative sentences showed more positive ERPs compared to the affirmative sentences ($\Delta_{Neg,Aff} = .919\mu V$),

[5]The regions were chosen based on the visual inspection of the effect's topography that suggested clear anterior-posterior differences, but no clear lateralization differences. Since we had no specific hypotheses regarding potential lateralization effects, we decided to include only AP as a factor in order to keep the analysis more transparent.

as well as an effect of *Region* ($F(1,23) = 67.535$, $p<.001$, $\eta^2 = .746$, $\Delta_{Post,Front} = 3.411\mu V$). The interaction *Alternatives×Region* was significant ($F(1,23) = 11.401$, $p = .003$. $\eta^2 = .331$), which can be attributed to a larger *multiple vs. unique* N400 effect in the posterior ($\Delta_{Mult,Unq} = -2.558\mu V$) relative to the frontal regions ($\Delta_{Mult,Unq} = -1.756\mu V$). Given that the three-way *Polarity×Alternatives×Region* interaction was also significant ($F(1,23) = 9.712$, $p = .005$, $\eta^2 = .297$), we broke down this interaction by *Region*. For the frontal region, there was a significant effect of *Alternatives* ($F(1,23) = 18.357$, $p < .001$, $\eta^2 = .444$, $\Delta_{Mult,Unq} = -1.756\mu V$), *Polarity* ($F(1,23) = 16.947$, $p < .001$, $\eta^2 = .424$, $\Delta_{Neg,Aff} = 1.232\mu V$) and the *Polarity×Alternatives* interaction ($F(1,23) = 5.288$, $p = .031$, $\eta^2 = .187$): The *unique vs. multiple* effect was larger for the affirmative sentences $\Delta_{Mult,Unq} = -2.4\mu V$ than for the negative sentences $\Delta_{Mult,Unq} = -1.112\mu V$. However, for the posterior region only the main effect of *Alternatives* was significant ($F(1,23) = 38.511$, $p < .001$, $\eta^2 = .626$, $\Delta_{Mult,Unq} = -2.557\mu V$). For the midline electrodes the effects were similar, i.e. there was a significant effect of *Alternatives* ($F(1,23) = 33.788$, $p < .001$, $\eta^2 = .595$, $\Delta_{Mult,Unq} = -2.473\mu V$), as well as of *Polarity* ($F(1,23) = 8.58$, $p = .008$, $\eta^2 = .272$, $\Delta_{Neg,Aff} = .943\mu V$), but no interaction.

The results of **Experiment II** were in line with the first experiment. There was a main effect of *Alternatives* ($F(1,23) = 21.045$, $p<.001$, $\eta^2 = .478$), i.e. the multiple conditions showed larger negativity than the unique conditions ($\Delta_{Mult,Unq} = -.972\mu V$), as well as *Region* ($F(1,23) = 36.042$, $p<.001$, $\eta^2 = .610$, $\Delta_{Post,Front} = 3.109\mu V$). No main effect of *Polarity* was observed. Unlike in the first experiment, there were no significant interactions. For the midline electrodes only a main effect of *Alternatives* was found ($F(1,23) = 24.920$, $p < .001$, $\eta^2 = .520$, $\Delta_{Mult,Unq} = -1.160\mu V$).

Alternatives independent Late Positivity for negated sentences. The visual inspection of grand averages revealed a late positivity effect for the negative relative to affirmative conditions, that was apparent for both alternatives conditions and both experiments. The analysis in the late time-window **550-850 ms** for **Experiment I** revealed a main effect of *Polarity* ($F(1,23) = 25.714$, $p<.001$, $\eta^2 = .528$), driven by the negative sentences showing more positive average amplitudes than affirmative sentences ($\Delta_{Neg,Aff} = 1.177\mu V$), as well as a main effect of *Region* ($F(1,23) = 108.986$, $p<.001$, $\eta^2 = .826$, $\Delta_{Post,Front} = 2.58\mu V$). No effect of *Alternatives* was observed; however, there was a significant *Alternatives×Region* interaction ($F(1,23) = 19.308$, $p < .001$, $\eta^2 = .456$): The mean amplitude difference between multiple and unique conditions was more negative in the frontal ($\Delta_{Mult,Unq} = -1.028\mu V$) than in the posterior region ($\Delta_{Mult,Unq} = .172\mu V$). There was also a significant three-way interaction *Polarity×Alternatives×Region* ($F(1,23) = 9.430$, $p = .005$, $\eta^2 = .291$), which we broke down by *Region*. In the frontal region, we found a significant effect of *Polarity* ($F(1,23) = 16.387$, $p < .001$, $\eta^2 = .416$, $\Delta_{Neg,Aff} = 1.192\mu V$), *Alternatives* ($F(1,23) = 11.088$, $p = .003$, $\eta^2 = .325$, $\Delta_{Mult,Unq} = -1.028\mu V$), as well as significant *Polarity×Alternatives* interaction ($F(1,23) = 6.867$, $p = .015$, $\eta^2 = .230$): the mean amplitude difference between the negative and affirmative sentences was larger for the multiple ($\Delta_{Neg,Aff} = 1.894\mu V$) than for the unique ($\Delta_{Neg,Aff} = 0.49\mu V$) condition. In the posterior region only the effect of *Polarity* was significant ($F(1,23) = 23.064$, $p < .001$, $\eta^2 = .501$, $\Delta_{Neg,Aff} = 1.161\mu V$). For the midline electrodes, we found an effect of *Polarity* ($F(1,23) = 29.341$, $p < .001$, $\eta^2 = .561$, $\Delta_{Neg,Aff} = 1.32\mu V$), but no effect of *Alternatives*, and no *Polarity×Alternatives* interaction.

The results for **Experiment II** were again generally consistent with **Experiment I**. There was a main effect of *Polarity* ($F(1,23) = 15.269, p = .001$, $\eta^2 = .399$) driven by the negative sentences showing more positive ERPs than the affirmative sentences ($\Delta_{Neg,Aff} = .883\mu V$), and a main effect of *Region* ($F(1,23) = 45.121$, $p<.001$, $\eta^2 = .662$, $\Delta_{Post,Front} = 2.234\mu V$), but there was no main effect of *Alternatives*. The interaction *Polarity×Region* was significant ($F(1,23) = 9.915$, $p = .004$, $\eta^2 = .301$): The difference between negative and affirmative conditions was larger in the posterior ($\Delta_{Neg,Aff} = 1.373\mu V$) than in the frontal regions ($\Delta_{Post,Front} = .393\mu V$). Additionally, and similar to the first experiment, the interaction *Alternatives×Region* was significant ($F(1,23) = 10.739$, $p = .003$, $\eta^2 = .318$, frontal $\Delta_{Mult,Unq} = -.214\mu V$ and posterior $\Delta_{Mult,Unq} = .688\mu V$). The interaction *Polarity×Alternatives* was not significant and neither was the three-way interaction. For the midline electrodes, again we only found an effect of *Polarity* ($F(1,23) = 17.516$, $p < .001$, $\eta^2 = .432$, $\Delta_{Neg,Aff} = 1.066\mu V$).

Comparison across experiments. Although both experiments showed similar main effects, some differences between the two framing variants were observed, in particular, some interactions showed significant in one experiment and not in the other. As the experiments were otherwise identical, as a meta-analysis we conducted a full-factorial ANOVA with *Polarity*, *Alternatives* and *Region* as within-subject factors and *Experiment* as a between-subject factor.

This analysis in the early time window **250-550 ms**, showed no main effect of *Experiment*; however, the interaction *Alternatives × Experiment* was significant ($F(1,46) = 7.029$, $p = .011$, $\eta^2 = .133$): the multiple vs. unique N400 effect was larger in Experiment I than in Experiment II. There was also a significant *Polarity × Region × Experiment* interaction ($F(1,46) = 5.311$, $p = .026$, $\eta^2 = .104$).[6] For the midline electrodes there was again no effect of *Experiment*, but the *Alternatives×Experiment* interaction was significant ($F(1,46) = 7.338$, $p = .009$, $\eta^2 = .138$).

No main effect of Experiment was found for the time-window of **550-850 ms**, but the interaction between *Alternatives* and *Experiment* was significant both in the main analysis

[6] See the experiment-specific analyses reported above for the relevant mean differences.

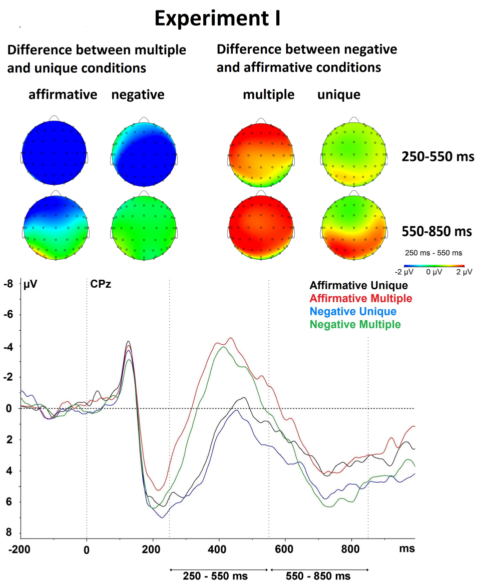

Figure 2: The comparison of grand averages at the critical word in all four conditions at the electrode CPz and the topographical distribution of the effects in Experiment I.

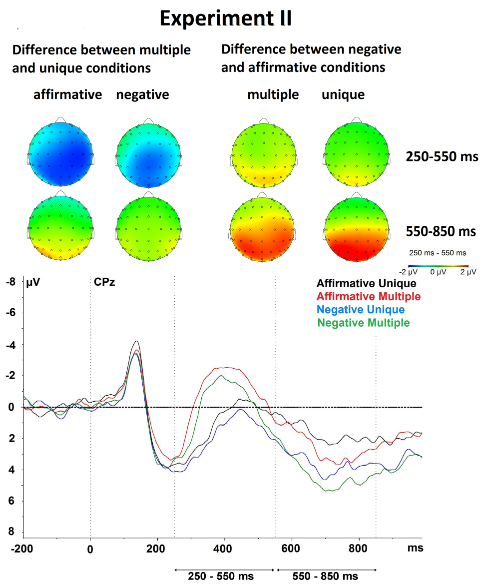

Figure 3: The comparison of grand averages at the critical word in all four conditions at the electrode CPz and topographical distribution of the effects in Experiment II.

($F(1,46) = 4.808$, $p = .033$, $\eta^2 = .095$, $\Delta_{Mult,Unq} = -.428\mu V$ in Experiment 1, and $\Delta_{Mult,Unq} = .237\mu V$ in Experiment 2), as well as for the midline electrodes ($F(1,23) = 4.603$, $p = 0.037$, $\eta^2 = .091$, $\Delta_{Mult,Unq} = -.443\mu V$ in Experiment 1 and $\Delta_{Mult,Unq} = .305\mu V$ in Experiment 2).

Discussion

In two experiments we compared the processing of affirmative and negative sentences in two contexts: (i) a unique context, that allows to make a specific prediction of the critical noun to be mentioned in a true sentence, (ii) a multiple context, where two alternatives can potentially be mentioned in a true sentence. In our design the affirmative and negative conditions were fully comparable: In both cases, sentence verification either required inferring the status of the unframed cards from the status of the marked set of cards (Experiment 1), or no inference was involved, since the sentence directly referred to the marked cards (Experiment 2).

It is generally accepted that the N400 reflects meaning-related expectancy of the stimulus. What this means precisely remains debated and the theories vary between taking the N400 to be a marker of lexical retrieval, lexical predictive preactivation or even meaning-related probabilistic prediction. We hypothesized that the presence of multiple alternatives, where the processor cannot uniquely predict the referent, should lead to a higher N400 relative to the case of a unique referent. This hypothesis was supported, as we observed a larger N400 effect for multiple vs. unique conditions independent of the sentence polarity. Although it is well-established that the N400 is correlated with expectancy and cloze probability, no prior experiments focused directly on the relation between the N400 and the availability of equally predictable alternatives in the scenario. Furthermore, our study is the first that directly compares how the scenario-based cloze probability of the upcoming word affects the processing of that word in affirmative and negative sentences. As the main result we show that the effect of alternatives is similar for both sentence polarities, thus, the possibility of predicting a unique true sentence continuation facilitates the processing also under the scope of negation.

In addition we also showed that, for both alternatives conditions, the presence of the negative particle led to a late positivity effect at the sentence critical noun occurring after the negation, i.e. when the noun's expectancy was related to the prior use of negation in the sentence. This effect forms a clear and large P600 effect in both experiments, although in the first experiment some modulation is already observed in the early time-window. Under the assumption that the P600 amplitude reflects integration of the lexical information into the semantic representation of the sentence (see Brouwer & Hoeks, 2013), this effect may be taken to indicate that in the case of negative sentences the construction of the semantic representation was more effortful.

Although the main pattern of effects was the same in both experiments, the type of framing made a significant contribution to the size of the effects, namely, the N400 effect for multiple vs. unique alternatives turned out to be significantly larger in Experiment 1 than in Experiment 2, as shown by the *Alternatives*Experiment* interaction. Furthermore, only in Experiment I we observe an interaction between *Polarity* and *Alternatives* by *Region* in both time-windows, specifically, in the posterior region the effect of alternatives was similar for both sentence polarities, but in the frontal region it was larger for affirmative sentences. This interaction result indicates that, in Experiment 1, affirmative and negative sentences possibly engaged slightly different processes in the two alternatives conditions. These between-experiment differences may be explained in terms of different task demands. Although from a logical perspective the two tasks were equivalent, cognitively they differ in an important manner. In Experiment 1, all target sentences referred to the unframed objects, whose status (chosen vs. unchosen) could only be inferentially determined based on the status of the framed objects and the assumed exhaustivity of the set divide. Thus, to determine the status of the unframed objects one had to include the so-called closed world assumption, which basically means that what is not known to be true is false. Given this assumption, in the negative condition, one could infer that if A was chosen (framed green), then B & C were not (or if A & B were chosen, then C was not). Similarly, in the affirmative condition, one could infer that if A was not chosen (framed red), then B & C were chosen (or if A & B were not chosen, hence C was). This inferential process is slightly different in the two conditions as it either goes from a negative premise to a positive conclusion, or the other way round. In contrast, in Experiment 2, all target sentences mentioned the framed, highlighted objects and hence there was no need to reason about the status of the remaining cards.

In sum, we showed that if the scenario allows to uniquely predict the upcoming word in a true sentence continuation, the processing of that word is significantly facilitated both in the case of affirmative and negative sentences, which is observed in the form of a reduced N400 component for the uniquely predicted words. The (higher) cognitive cost of processing the negative particle is observed in the form of a late positivity effect. Finally, the task demands, i.e. whether the status of the referent is directly marked or has to be inferentially determined, make a significant contribution to the size of the effects and appear to differently affect the negative and affirmative sentences. Further research should explore how the N400 is modulated by a larger number of alternatives provided and how the effect depends on the sentence truth-value.

Acknowledgments

This collaborative research was funded by *Stiftung Mercator* within the project *Structure of Memory* as well as by the German Research Foundation, DFG, within the priority program XPrag.de.

References

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: Rethinking the functional role of the p600 in language comprehension. *Brain Res. 2012 Mar 29;1446:127-43. doi: 10.1016/j.brainres.2012.01.055.*, *29*(1446), 127-43.

Brouwer, H., & Hoeks, C. (2013). A time and place for language comprehension: mapping the N400 and the P600 to a minimal cortical network. *Frontiers of Human Neuroscience*, *7*(758).

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, *3*(3), 472-517.

Dale, R., & Duran, N. (2011). The cognitive dynamics of negated sentence verification. *Cognitive Science*, *35*(5), 983-996.

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

Dudschig, C., Mackenzie, I. G., Maienborn, C., Kaup, B., & Leuthold, H. (2019). Negation and the n400: investigating temporal aspects of negation integration using semantic and world-knowledge violations. *Language, Cognition and Neuroscience*, *34*(3), 309-319.

Federmeier, K., Wlotko, E., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, *18*(1146), 75-84.

Fischler, I., Bloom, P., Childers, D., Roucos, S., & Perry, N. (1983). Brain potentials related to stages of sentence verification. *Psychophysiology*, *20*(4), 400-409.

Hagoort, P., Brown, C. M., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, *8*(4), 439-483. (Doi:10.1080/01690969308407585) doi: 10.1080/01690969308407585

Hagoort, P., Hald, L. A., Bastiaansen, M. C. M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304*(5669), 438-441.

Just, M. A., & Carpenter, P. A. (1971). Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior*, *10*(3), 244-253.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23-49.

Kuperberg, G. R., & Jaeger, T. F. (2015). What do we mean by prediction in language comprehension? *Language, Cognition, and Neuroscience*, *31*(1), 32-59.

Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophisiological distinction in processing conceptual relationship within simple sentences. *Cognitive Brain Research*, *17*(1), 117-129.

Kutas, M., & Federmeier, K. D. (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.

Lau, E., Holcomb, P., & Kuperberg, G. (2013). Dissociating N400 effects of prediction from association in single word contexts. *Journal of Cognitive Neuroscience*, *25*(3), 484-502.

Lüdtke, J., Friedrich, C. K., De Filippi, M., & Kaup, B. (2008). Event-related potential correlates of negation in a sentence-picture verification paradigm. *Journal of Cognitive Neuroscience*, *20*(8), 1355-1370.

Lüdtke, J., & Kaup, B. (2006). Context effects when reading negative and affirmative sentences. In R. Sun (Ed.), *Proceedings of the 28th annual conference of the cognitive science society* (p. 1735-1740). Lawrence Erlbaum.

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

Orenes, I., Beltran, D., & Santamaria, C. (2014). How negation is understood: Evidence from the visual world paradigm. *Journal of Memory and Language*, *74*, 36-45.

Pickering, M., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*, 105110.

Rabovsky, M., Hansen, S. S., & McClelland, J. L. (2018). Modelling the N400 brain potential as change in a probabilistic representation of meaning. *Nature Human Behaviour*, *2*(693-705).

Spychalska, M., Kontinen, J., & Werning, M. (2016). Investigating scalar implicatures in a truth-value judgment task: Evidence from event-related brain potentials. *Language, Cognition and Neuroscience*, *31*(6), 817840.

Tian, Y., Ferguson, H., & Breheny, R. (2016). Processing negation without context - why and when we represent the positive argument. *Language, Cognition and Neuroscience*, *31*(5), 683-698.

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, *83*(2), 176-190.