frontiers | Frontiers in Language Sciences

TYPE Hypothesis and Theory
PUBLISHED 01 March 2024
DOI 10.3389/flang.2024.1304613

OPEN ACCESS

EDITED BY
Lucia Colombo,
University of Padua, Italy

REVIEWED BY
Ashley Glen Lewis,
Max Planck Institute for Psycholinguistics, Netherlands
Hassan Banaruee,
University of Education Weingarten, Germany

*CORRESPONDENCE
Viviana Haase
viviana.haase@rub.de

RECEIVED 29 September 2023
ACCEPTED 02 February 2024
PUBLISHED 01 March 2024

CITATION
Haase V (2024) Disentangling inhibition and prediction in negation processing.
Front. Lang. Sci. 3:1304613.
doi: 10.3389/flang.2024.1304613

Disentangling inhibition and prediction in negation processing

Viviana Haase*

Institute for Philosophy II, Ruhr University Bochum, Bochum, Germany

Negation as a universal feature of human language is used effortlessly in everyday communication. However, experimental research has shown that the comprehension of negated sentences seems to require additional cognitive resources compared to affirmative sentences. Many studies investigating the processing of negation report longer reading and reaction times for negative compared to affirmative sentences and many studies report a Polarity by Truth interaction: false affirmative sentences lead to longer response times and larger N400 event-related potentials (ERPs) than true affirmative sentences, whereas the pattern is reversed for negative sentences where it is the true sentence that elicits longer reaction times and higher N400 ERPs compared to false negative sentences. These interactions have been discussed in the light of lexical associations, predictability, and the need to construct two subsequent mental representations. Furthermore, recent studies have shown that the comprehension of negated sentences seems to make use of neural resources that are typically involved in cognitive control and inhibitory mechanisms. As both processes have been associated with two different and temporally overlapping ERP components (the N400 and the P300), we focus on studies with high temporal resolution. We discuss linguistic aspects of negation, such as semantic similarity and contextual invariance of negation. We furthermore discuss the role of the verb as well as the position of the negative marker with respect to the verb, and their respective relevance for predictive and inhibitory mechanisms in negated sentences.

KEYWORDS

negation, inhibition, prediction, N400, P300

1 Introduction

While negation is an essential part of every human language and is used effortlessly in everyday language, psycho- and neurolinguistic studies have repeatedly shown that negative sentences seem to require different, possibly additional, cognitive resources compared to affirmative sentences. Among the recurring results within the past 60 years of research are the findings that (i) negative sentences lead to higher reading and reaction times and higher error rates compared to affirmative sentences, (ii) the presence of a negation in a sentence does not modify the N400 event-related potential (ERP) in true and false sentences and (iii) affirmative and negative sentences elicit a Truth by Polarity interaction with false compared to true affirmative sentences leading to longer reaction times and higher N400 ERPs but the opposite pattern for true and false negative sentences, that is, higher reaction times and N400 amplitudes for true compared to false negative sentences. See Kaup and Dudschig (2020) for a review. In addition to the processing difficulty of negated sentences, the literature on negation addressed further research questions in more recent years. While early on it has been assumed that mental representations are of propositional format, more recent accounts assume that

language is grounded in cognitive systems for action and perception, and therefore, mental representations are assumed to be of simulative nature, influenced by experiential traces that are created over time. Negation functions as a test case for the idea of simulative views on language comprehension as it might be challenging if not impossible to simulate the absence of something. Furthermore, two mechanisms that are part of human cognition have been assumed to be of essential nature for the comprehension of negation and negated sentences: Prediction of upcoming sentence material leading to the activation of representations on the one hand and inhibition of negated representations on the other hand.

1.1 Prediction

Prediction is a fundamental feature of human cognition, that is, we update an internal model of the world to predict further sensory input. In case of a mismatch between prediction and the actual input a prediction error occurs. Prediction is an important part of visual and emotional processing, attention, and motor control (see e.g., Bubic et al., 2010 for an overview). Prediction is furthermore an essential part of language processing (DeLong et al., 2005; Pickering and Garrod, 2007, 2013; Altmann and Mirković, 2009; Kutas et al., 2011; Pickering and Gambi, 2018). The fast and efficient use of language and our ability to form an infinite amount of sentences generated the now widely accepted view of prediction being an essential part of incremental language comprehension in which upcoming linguistic material is predicted based on the information already processed in discourse (see e.g., Bar, 2009; Freunberger and Roehm, 2016). It has furthermore been debated what it means to form a prediction in contexts that offer multiple options for prediction (DeLong et al., 2005; Van Petten and Luka, 2012; Kuperberg and Jaeger, 2016). In this article, prediction is understood in terms of pre-activation, that is, memory retrieval of upcoming linguistic material where the ease of memory retrieval depends on the retrieval cues that are presented by the prior context. This assumption is in line with the Retrieval-Integration Account (Brouwer et al., 2012, 2016; Brouwer and Hoeks, 2013).

1.2 Inhibition

Apart from prediction, inhibition is an essential feature of human cognition, necessary to adapt behavior and inhibit inappropriate or unsafe actions (Chambers et al., 2009). Furthermore, as part of language acquisition, children learn the connection between negation and inhibition: negative imperatives signal the need to end initiated or intended actions. In parts of the literature on negation its effects have been described in terms of reduced availability of representations or as suppression (MacDonald and Just, 1989; Kaup et al., 2006, 2007) or as reduced availability of neural information (Tettamanti et al., 2008; Papeo et al., 2016). This link between negation and reduced availability of conceptual representations and reduced neural activation suggests that inhibitory mechanisms are to some degree involved in the processing of negated sentences. More recent literature even goes further and suggests that neurocognitive mechanisms that are relevant for inhibition are reused for negation (Beltrán et al., 2021).

1.3 Aim of current article

It is still debated how and to what degree mental representations are activated and especially how negation modulates the predictive process as it is more difficult to predict something that is absent or does not hold than something that is present or is the case. Regarding inhibition, recent studies have shown that negation comprehension and inhibition seem to share some neural mechanisms. However, it remains to be clarified how and when the inhibition of negated representations takes place. Overall, the temporal dynamics of both processes taken together need to be clarified, that is, how and when does prediction take place once a negation is encountered in a sentence and what are its neurocognitive correlates and at what point after the occurrence of a negative marker does the inhibition process take place and what are its neurocognitive correlates.

The aim of this article is to shed further light on the interplay and temporal dynamics of prediction and inhibition in the comprehension of negated sentences. After reviewing the role of prediction in language comprehension in general, we present aspects relevant to the prediction in negated sentences, such as the contextual invariance of negation and semantic similarity, among others. We furthermore address the role of the verb and the position of the negation with respect to the verb and their relation to predictive and inhibitory mechanisms in negated sentences. Subsequently, we review inhibition as part of language comprehension with a focus on negation before we discuss the interplay of both cognitive mechanisms in the last part of the article. We include suggestions for upcoming studies in our discussion.

2 Prediction

2.1 Prediction in language comprehension

Context facilitates the processing of upcoming material in the linguistic input (Kuperberg and Jaeger, 2015) and modulates language comprehension (for an overview of context effects during language comprehension see for example Schumacher (2012)), not only on the level of lexical semantics but also regarding world-knowledge, mutual knowledge and interlocutor specific knowledge (Bornkessel-Schlesewsky and Schumacher, 2016). In many cases, there is not only one potential word that would continue a true sentence plausibly and grammatically but several potential words that would be both plausible and grammatical given a specific context. This raises the question of how this multiplicity, that is, the availability of true alternative sentence continuations, affects prediction. One possibility would be the prediction of several potential sentence continuations (here: words) in parallel, with each word being of equal probability. Another possibility would be the prediction of several words but each word is assigned a specific probability, resulting in weaker and stronger predictions.

For example, after receiving the contextual information that Maria likes to eat fruits for breakfast as in example 1, several equally probable sentence completions for the following sentence are possible upon hearing a sentence fragment such as *Today she ate...* Hence, both example 1.1 and 1.2 can be considered equally probable (assuming that both apples and bananas are available throughout the year and that she generally likes both of them).

1. Maria usually eats fruits for breakfast.

1.1. Today she ate two apples.

1.2. Today she ate two bananas.

Therefore, according to the second possibility named above, a reader's brain can gradually pre-activate multiple potential sentence continuations to a degree that can be estimated from the probability that each word is given as a continuation for a sentence fragment (DeLong et al., 2005). During the computation of predictions for upcoming words in the linguistic input, coherence is maximized (Bornkessel-Schlesewsky and Schumacher, 2016), that is, co-reference relations are preferred (Burkhardt, 2006), causal relations facilitate the processing (George et al., 1997; Kuperberg et al., 2011), and events are processed more easily when presented in their natural order (Münte et al., 1998). Empirical evidence for (graded) prediction during language comprehension comes, among others, from event-related potential studies reporting an N400 ERP. The N400 is a negative deflection with a peak typically around 400 ms. Its amplitude varies in response to stimulus expectancy (Kutas and Hillyard, 1980; Kutas and Van Petten, 1994; Kutas and Federmeier, 2000, 2011; Swaab et al., 2012) and inversely correlates with a word's cloze probability (Taylor, 1953), independent of context (Federmeier et al., 2007). The N400 has been interpreted in terms of facilitation of lexical retrieval (Kutas and Federmeier, 2000), ease of retrieval (Brouwer et al., 2012, 2016; Brouwer and Hoeks, 2013), stimulus pre-activation (DeLong et al., 2005; Schumacher, 2014), semantic integration into prior context (Hagoort et al., 2004) and, more recently, also in terms of probabilistic meaning-related prediction (Lau et al., 2013; Kuperberg and Jaeger, 2015; Rabovsky et al., 2018; Delaney-Busch et al., 2019; Rabovsky, 2020).

The (gradual) pre-activation of multiple sentence continuations leads to the question of what happens once the right alternative has been processed. While the literature on the N400 ERP provides information about the degree of pre-activation of a specific word in relation to other words, it does not reveal what happens to these words subsequently. Do the competing alternatives stay active in memory even if they turn out to be wrong in the prediction process, or do they get inhibited? Is there a difference between strong predictions that turn out to be wrong and weak predictions (with multiple alternatives predicted in parallel) with regard to a potential inhibition? These questions will be addressed in Section 3.

While in affirmative sentences prediction still seems to be relatively straightforward, it becomes much harder in negative sentences since it is in general more difficult to predict something absent. This leads to the question of what is predicted once we encounter a negation. The next subsection introduces various aspects that are of relevance for prediction-making in general, and especially concerning negated sentences.

2.2 Prediction in negated sentences

2.2.1 Probability

If context provides multiple true potential referents, the probability for each of them can be assumed to be split between them. If these alternatives, that is, potential and true referents, are of equal probability, it can furthermore be assumed to be split evenly across all potential referents. With an increasing number of alternatives, however, the cognitive load of predicting a range of alternatives increases as well, raising the question of whether not predicting anything would be the more reliable strategy in that case. Depending on the individual memory capacities, the application of this strategy might occur for varying numbers of alternatives.

However, what happens in negative sentences? Provided there is only one potential true referent in the context, the prediction of a specific word in a negative sentence should not be more difficult than the prediction of a specific word in an affirmative sentence. This assumption is confirmed by results from an Eye-Tracking study by Orenes et al. (2014) who show that in a context with a binary choice of two items participants immediately focus on the alternative as soon as one item is negated. Instead when being presented with multiple potential referents, participants stayed focused on the negated item as there was no prediction cue for any of the alternatives, which all had equal probability. In communication, the choice between alternatives often is not binary, leading to the question of how predictions are being made, and which cues in the linguistic input help form adequate predictions.

2.2.2 Negation as a truth-function

As a truth function, negation changes the semantic value from *p* to *not p*. In communication, it thereby changes the addressees attention from one set to the other. Related to that it has been argued that sentential negation opens up two mental spaces and has two epistemic states (Verhagen, 2010). From the perspective of logic, everything in the logical space must be either *p* or *not p*, and hence, *not p* is the complement set of *p* and vice versa. Therefore, if *cat* denotes the set of all cats, *no cat* denotes the complement of that set. Therefore, in an affirmative sentence, predictions must be made within the set of *p* to render the sentence true. Instead, encountering a negation requires the comprehender to predict an entity from the complement set *not p* to render the negative sentence true. Yet, that leaves them with a vast and seemingly unpredictable amount of possibilities, because the complement set *not p* contains everything that would render *p* false, and therefore is much larger than the set of *p*. Consider for example sentence 2.1. Once encountering the verb (underlined) the reader can predict the grammatical status of the upcoming phrase or word, for example, a noun, a noun phrase, or a prepositional phrase, e.g., *in Africa*. They could furthermore predict several meanings such as, for example, *a river* or *a long river*, which would render the sentence true as well. Instead, in example 2.2, it is much harder for the reader to predict what comes after the verb phrase, as the set of possibilities in the complement set is vast, for example, any continent other than Africa could be listed, any adjective that does not describe the Nile adequately, every noun that does not denote that river

correctly. Hence, the form-based predictability seems to be on par with the one in the affirmative sentence, but on the semantic level, a seemingly infinite number of meanings would render the sentence true. Of course, there are more plausible options, e.g., *not in Asia, not in Egypt* and less plausible options, e.g., *not purple, not in Berlin, not a tomato* that would make the sentence true but are not plausible to be uttered (cf. Wason, 1965). Yet, on the logical level, the set of truthmakers for the negated sentence is much larger than the set of truthmakers for the affirmative sentence. Here, the term *truthmaker* refers to the set of words and expressions that function as a true sentence continuation, independent of other factors such as plausibility.

2.1. The Nile is in Africa.

2.2. The Nile is not in Asia.

Furthermore, this example demonstrates that the truthmaker of the affirmative sentence and the truth maker of the negated sentence are similar. This is not a coincidence but a typical feature of negation.[1]

2.2.3 Contextual invariance of negation

Predicates and their negations as well as contrasting words tend to occur in similar contexts. Fauconnier (1999) accordingly describes the function of elementary negation as being used to contrast one situation with another one that is similar. Therefore, entities or scenarios that might be feasible as alternatives (that is, from the set of *not p*) to the negated situation (that is, from the set of *p*) can be expected to occur in similar contexts as well. For example, the assertions "I am hungry", "I am not hungry" and "I am full" are feasibly uttered in similar situations, e.g., in a restaurant, or anywhere else where buying, cooking, or eating a meal is or can be plausibly addressed. Similarly, for "This is not a cat…", "…it is a dog" is a more plausible continuation than "…it is a meeting". In light of this contextual invariance of negation, an overlap of features between the negated entity or proposition and its affirmative alternative(s) can be assumed. However, what makes one alternative more plausible than another one?

2.2.4 Semantic similarity

Categorization research suggests that many human categories are taxonomic, that is, concepts are grouped based on shared perceptual and functional attributes (Kay, 1971; Rosch et al., 1976). Membership within a category is graded, determined by whether and how many attributes a concept shares with other concepts of a category (e.g., Rosch (1973, 1975)). Assuming that we predict potential alternatives during online sentence comprehension, a graded spread of activation in a semantic network can be assumed, in which the level of activation depends on the overlap of attributes and in which the ease of retrieval is determined by the retrieval cues that are present in a word's prior context (Brouwer et al., 2012, 2016; Brouwer and Hoeks, 2013). See Naumann and Petersen (2021) for a recent proposal on the dependency of the N400 ERP on categorization. Given the contextual invariance of negation, the notion of semantic similarity is an important aspect of the comprehension of negated sentences as it influences the prediction and therefore activation and retrieval of alternatives during the comprehension process by suggesting that affirmative and negative sentences share a solid amount of retrieval cues.

For example, if someone suddenly, that is, without further context, says "This is not an apple", the most probable concepts that function as alternatives to the negated concept [NOT APPLE] can be assumed to be of maximal similarity to an apple, e.g., a peach which differs in taste and consistency, but has the same shape and belongs to the category of fruits as well. Instead, an entity from the same category but with a higher amount of distinct features, such as, for example, a banana, seems to be a less likely candidate. Instead, an entity of another category, e.g., a ship, would be a completely odd -yet true- alternative.[2] However, to limit the alternatives evoked by negation, mere category membership is a helpful yet insufficient constraint since it does not allow for a fine-grained distinction between similar concepts, and hence, further cues and information in the linguistic input are necessary to minimize the set of alternatives.

2.2.5 The role of the verb

The verb is the crucial element of each sentence as it determines the number of arguments[3], and the relation between them, that is, the verb is the decisive part of an event's linguistic description. It furthermore imposes selectional restrictions about the arguments, for example, whether they are animate or not or whether they are humans or instruments. So far, its role in the comprehension of negated sentences has been neglected to our knowledge.

In an MEG study, Maess et al. (2016) report differences in the predictability of verbs and nouns. Their participants were presented auditorily with simple German sentences with verbs that were either highly predictive (e.g., *He drives the car*) for the occurrence of a particular noun given a specific context, or were not predictive (e.g., *He gets the car*). They report a reduction of the N400 for nouns in sentences with highly predictive verbs as compared to nouns in sentences with weakly predictive verbs but report an opposite pattern for the preceding verbs: highly predictive (that is, more informative) verbs led to stronger neural magnitude compared to less predictive verbs. Furthermore, in a distributed brain network, a negative correlation between the N400 effect of the verb and the N400 effect of the noun was found, indicating an integral relation between the predictive power of the verb and the processing of the subsequent noun (Maess et al., 2016). According to the authors

1 Furthermore, this example demonstrates that the plausibility and therefore also the probability of the entity describing the subject is similar to the subject entity in terms of their semantics, not only in the affirmative but also in the negative cases. Hence, in copular structures of the type *x is z; x is not y*, the most plausible *z*, *y* seem to be entities that share several features with *x*.

2 Let us remark at this point that for this reason (among others) the experimental material of some studies (e.g., Fischler et al., 1983; Wiswede et al., 2013) is not ideal for the investigation of processes underlying negation comprehension.

3 The valency of the verb controls the number of arguments that are required, e.g., monovalent verbs such as *sleep* only need one argument, as in *He sleeps*, whereas bivalent verbs such as *He eats pizza* require two arguments, in this case, the grammatical subject and direct object. Trivalent verbs, for example *give*, require three arguments, as in *He gave the letter to Sarah*.

of the study, predictions are only made if evidence is sufficiently high and therefore, the processing of the predictive verbs leads to a co-activation of expected nouns that often co-occur with those verbs and therefore have stronger weightings, while less predictive verbs do not cause multiple predictions for all possible sentence continuations as they only have weak co-occurrence weightings with other words.

Negative particles are often preceding the verb (cf. *Neg-First-Principle* (Horn, 1989)) and in various cases are reinforced by a postverbal negative particle, as in the French *ne...pas* (Payne, 1985; Zanuttini, 2001). Yet, in languages as for example German, in which the placement of a verb is constrained to a specific position in the sentence (V2), the negative marker follows upon the verb. In certain languages that contain multiple possibilities to express negation, though, the different types of negation of course can bear various functions.

The resulting question is whether the position of the negation before or after the verb affects prediction to different degrees. Assuming that we continuously make predictions about the upcoming input, preverbal negation allows us to make our predictions concerning a sentence of negative polarity at an earlier stage of sentence comprehension as postverbal negation does. However, early on in the sentence, before encountering the verb, prediction is extremely hard. For example, consider which predictions can be made upon hearing a sentence beginning with *Sarah*.... Theoretically, every verb that is at least monovalent and that has an animate or more specifically a human first argument could appear in this sentence. As described earlier, in a negated sentence the amount of possible true sentence continuations is even higher. In a sentence with preverbal negation, when the reader/listener encounters the verb, they already have the information that the event described by this verb did not take place. Possibly, they therefore might not have activated the affirmative state of affairs in the first place or at least they might not have obtained a full representation or simulation of the described event.[4] The processing benefit of a preverbal negation would therefore be that it might facilitate a direct representation of the actual state of affairs, without requiring representing (and then inhibiting: see next section) the negated state of affairs. The downside of preverbal negation however is, that it makes predictions relatively hard in general. In line with this assumption is the argument from Nieuwland (2016) who suggests that the incremental incorporation of a negative term into the sentence representation may be particularly difficult when occurring early in the sentence, that is, when the scope is still undetermined. This is however contradicted by Dudschig et al. (2019) who investigated whether additional time to process the negation operator facilitates the integration of negation into the sentence meaning. In one of their experiments, they used an external negation as in *It is not the case that ladybirds are stripy* vs. *Ladybirds are not stripy* in the other experiment. Overall their results suggest that prepending the negation operator (thereby clearly having a wide-scope sentence) does not seem to be neither beneficial nor harmful for the processing of negation as opposed to positioning the negation in a postverbal position within the sentence.

At this point we would like to point out that a large number of prior studies investigating the comprehension of negated sentences make use of copular constructions as, among others, in Wason (1959), Clark and Chase (1972), Villiers and Flusberg (1975), Fischler et al. (1983), Carpenter et al. (1999), Hasson and Glucksberg (2006), Kaup et al. (2006, 2007), Lüdtke et al. (2008), Nieuwland and Kuperberg (2008), Dale and Duran (2011), Wiswede et al. (2013), Orenes et al. (2014, 2015), Dudschig et al. (2019), Haase et al. (2019). Copular verbs[5] such as *to be* bear comparatively little meaning opposed to "full verbs". If we further assume that prediction is largely driven by verb-based information, such as the number and type of arguments (cf. Maess et al., 2016), such copular constructions are to be considered with caution as they construe a link between the subject and the predicate, but they do not denote a real event but rather denote some features or facts about an entity. While copular verbs are convenient for several experimental manipulations, one should be careful when interpreting them with regard to the larger picture of negation comprehension in general.

In sentences with postverbal negation, and hence, after verb-based predictions could be made, it might be easier to predict further linguistic material. However, the reader/listener might already have activated a partial representation of the affirmative state of affairs and therefore might already have activated potential truthmakers of a true affirmative sentence as part of the predictive process. Hence, once the negation is encountered here, the reader/listener needs to adjust their attention toward the set of truthmakers for a negative sentence and are relatively likely to "end up" with two subsequent representations: one of the negated state of affairs and one of the actual state of affairs.

Hence, the placement of the verb after the negation might increase the difficulty of the prediction of the second argument based on the negated verb. On the other hand, the placement of the negation after the verb makes a representation of the underlying affirmative in a first step more likely, especially if one of the arguments (or the only argument) of the verb has already been encountered, making a preactivation of a potential second argument and therefore the prediction and representation of the affirmative scenario more likely. Furthermore, it should be noted that misinterpretations are more likely to persist the longer they have been part of the initial discourse model. This aspect can be transferred to negation comprehension: the later the negation occurs in the sentence, the longer the affirmative meaning forms the initial discourse model, which makes it more likely that the

4 This aspect depends on the underlying approach of language comprehension. Assuming that simulations are a mandatory part of language comprehension, a simulative representation would have to take place in this case as well to obtain a full comprehension of the sentence's meaning. Assuming that the simulative representation is more of an epiphenomenon, negative sentences would not necessarily have to result in a simulation in such cases in which it is clear that the event described by the sentence does not take place. Finally, assuming a symbolic representation, the comprehender can simply symbolically represent the non-existence of the event.

5 A copular verb typically either expresses that the subject and its complement denote the same thing or that the subject has the property denoted by its complement.

initial affirmative representation stays active. This is in line with results from Tian et al. (2016), who show that shortly after the verb in simple negative sentences like *Matt hasn't shut his dads window* participants looked at both the image that was consistent with the content of the underlying affirmative (closed window), and the image consistent with the actual sentence meaning (open window), suggesting that the representation of the underlying affirmative is initially activated, as hypothesized by two-step models of negation comprehension (cf. Kaup et al., 2006; Lüdtke et al., 2008).

Based on comprehenders accurate plausibility judgments as well as a robust P600 ERP effect, Chow et al. (2018) suggested that comprehenders reliably use argument role information for the real-time interpretation of a sentence. Furthermore, they suggest that their predictions, as reflected in an N400, are not immediately sensitive to argument role information. Under that account, comprehenders immediately and reliably use argument role information to interpret thematic relations when the verb appears, but they may not use this information quickly enough to predict the upcoming verb before it appears in the input (Chow et al., 2018). Similarly, the information provided by a negation might not be used quickly enough to adapt the system to the scenario that could be described by a negative sentence since a) the sentence has to be reanalyzed from an affirmative-first to negative polarity, and b) the inference to what is the case has to be made, if possible.

2.2.6 Scope

The activation of alternatives depends on the scope of the negation and therefore on the negated dimension within a sentence. For example, in a sentence like "Holly did not eat a pizza for dinner", it can be the subject (Holly), the action (eating), the object (the pizza) and the specification (for dinner) that can be negated and that, therefore, open up space for alternatives.[6] Tian et al. (2016) have shown that in sentences with clear scope and therefore clearer alternatives, (e.g., *It is John who hasn't ironed his brothers shirt*) incremental comprehension is facilitated.

2.2.7 The Polarity Index

Any sentence is by default indexed as being true until there is evidence to the contrary (cf. Clark and Chase, 1972). In analogy to this truth index, we postulate a *polarity index* assuming that a sentence is parsed under the supposition that it is of affirmative polarity until there is evidence to the contrary. From a logical point of view, it is the use of a negative polarity marker that requires the change of the truth value leading to an additional logical operation and therefore possibly leading to increased processing costs. Yet, that presupposes that we parse a sentence with affirmative polarity first.

This affirmative first hypothesis finds support by studies that show direct effects of the negative marker, especially such effects that indicate rather structural reprocessing (for example, a late positivity in an ERP-study, e.g., Fischler et al. (1983), Lüdtke et al. (2008), Spychalska et al. (2019, 2023)). The idea of certain default structures that are preferred across several languages is not new. Regarding the sequence of subject and object (or actor and undergoer on the semantical level), it is well known and has been proven in several studies that there is a preference to interpret the first entity of a sentence as being the grammatical subject. Among others, this preference results in the so-called subject-object asymmetry, leading to the often observed effect of object relative clauses being more difficult to understand than subject relative clauses, which often is traced back to the so-called "subject-first-preference" (King and Just, 1991; King and Kutas, 1995; Gordon et al., 2001; Bornkessel-Schlesewsky and Schlesewsky, 2009). Accordingly, a preference for the more frequent affirmative polarity can be assumed, resulting in an asymmetry between affirmative and negative sentences.

After having introduced aspects of prediction and how they modulate language comprehension, especially the comprehension of negated sentences, the next section introduces the inhibition of previously activated representations in case a negation is encountered in the linguistic input.

3 Inhibition

3.1 Inhibition in language comprehension

In Section 2.1 we described the possibility of multiple predictions with either equal or graded probability. This leads to the question of what happens with activations that are no longer required, for example, due to multiple predictions that happened in parallel. When prediction is understood in terms of pre-activation leading to memory retrieval of upcoming linguistic material (cf. Brouwer and Hoeks, 2013; Brouwer et al., 2012; Brouwer et al., 2016), as it is the case in this article, it is, therefore, possible that some of these activations turn out to be "unnecessary", namely in cases of a prediction error. However, what happens to these activations? Do they get deactivated[7] once they turn out to be unnecessary? And if so, does this deactivation happen gradually in a similar way to activation being gradual?

Literature on metaphors can provide useful information here. Understanding nominal metaphors such as *My lawyer is a shark* requires an activation of metaphor-relevant information (e.g., *aggressive*, *tenacious*) and a non-activation or, alternatively, deactivation of metaphor-irrelevant information (e.g., *swims*, *has a fin*). Remember that deactivation or inhibition requires an activation (cf. Beltrán et al., 2021). One of the major psycholinguistic theories on metaphor comprehension is the class-inclusion theory (Glucksberg and Keysar, 1990, 1993; Glucksberg et al., 2001; Glucksberg, 2003) according to which a metaphor *X is a Y* is understood as a class-inclusion statement that puts X and Y both into a joint category. Banaruee et al. (2017) suggest that metaphorical classes are created by means of an inhibition-oriented

6 From an information structural point, the default reading in this case, at least in the absence of intonational information, has the subject as a topic and the action, the object and the specification in the comment section.

7 The literature offers various terms for this mechanism, such as "deactivation", "inhibition", and "suppression". It is not entirely clear whether they all refer to the same process. Regarding metaphors, the term "filter" has also been used for this mechanism. We will not distinguish between these expressions here.

mode of comprehension: each metaphorical class is formed by the inhibition of the majority of semantic features and by maintaining a very small set of them. The inhibition mechanisms filter out semantic features. At this point, it is important to distinguish between literal classes and metaphorical classes. Literal classes are largely hierarchical, with the topmost category being defined by a minimum number of semantic features and the bottom categories being defined by specific, more fine-grained features. The same hierarchy does not apply to metaphors. For example, in the metaphor *My job is a jail* the term *jail* does not refer to a category of buildings in the literal sense, but rather to a larger number of restrictive conditions (Banaruee et al., 2017). Assuming that metaphorical classes are understood based on only few semantic features, the remaining features do not need to be activated, and if they become activated they can be filtered out or inhibited. Experimental research on metaphor comprehension has shown that this is indeed the case (e.g., Glucksberg et al. (2001)). Therefore, inhibition is not restricted to the use of negation in a sentence, but rather seems to occur when previous discourse led to the activation of words or features that are no longer needed in the comprehension process.

3.2 Inhibition in negated sentences

Early on MacDonald and Just (1989) showed in a probe recognition and a probe naming task that negated concepts have lower levels of activation in the discourse representation, independent of the position of the negation in the sentence. They presented sentences like *Every weekend, Maria bakes some bread but no cookies for the children.*, followed by recognition or naming tasks of the probe *bread* or *cookies*. Their results furthermore suggested that negation does not inhibit the activation level of related concepts. See Kaup and Zwaan (2003), Kaup et al. (2006) and Mayo et al. (2004) for similar results. Overall, MacDonald and Just (1989) interpret their results as evidence for changes in the activation level of negated concepts via shifts in discourse focus independently from subsequent truth-value computation. In their view, negation reduces the accessibility of information in its scope by means of reduced activation levels for negated information. This aspect can be linked to information structure and discourse representation: negated concepts are less focused in the discourse representation, that is, they might have lower levels of activation than more central (non-negated ones). MacDonald and Just (1989) draw a parallel between selective negation in a small set and reference resolution. They observed that when presenting sentences containing two names the presence of a pronoun shifted the discourse focus toward the referent, and away from the other name resulting in inhibition for the nonreferent name. Hence, in analogy, if one of two coordinated concepts is negated, the negation might decrease the relative prominence of the negated concept. It should be noted that activation levels of concepts could be further altered by stress in spoken language and by modifying the structure of the sentence to render parts of the clause more prominent, e.g., by using cleft constructions (see for example Tian et al. (2016)). At this point, it is important to note that these interpretations in terms of deactivation or inhibition require a representation and therefore activation of the negated as well as the non-negated concepts during language comprehension (cf. Beltrán et al., 2021).

Further support for the argument that negation deactivates concepts in its scope comes from studies investigating the comprehension of negated action-related sentences. Negation reduces differences between action-related and abstract concepts and therefore seems to modulate so-called *embodiment effects*.

For example, Tettamanti et al. (2008) investigated the modulation of negation on the neural representations of action-related and abstract concepts. Hence, they tested whether the impact of negation on the neural responses elicited by the negated propositions is dependent on the semantic field. Their stimuli consisted of action-related (e.g., *Adesso io/non premo il bottone. (Now I/not push the button)*) or abstract (*Ora io/non apprezzo la fedeltà* (*Now I/not appreciate the loyality*)) sentences.

Assuming that negation leads to reduced access to neural representations of the negated proposition for a negative compared to an affirmative action-related sentence, we would expect a reduced activation in the left front-parietal-temporal system, which is the region that is assumed to be activated for the representation of actions (Tettamanti et al., 2008). Instead, for the comparison of affirmative and negative abstract sentences we would expect reduced activity in the posterior cingulate cortex for negative sentences. Tettamanti et al. (2008) report an increased signal in the left ventral inferior frontal gyrus, the left inferior and middle temporal gyri, and the posterior cingulate cortex for abstract compared to action-related sentences. For negative abstract sentences, the authors observe a deactivation of the posterior cingulate cortex, whereas they observe a reduction of activation and connection strength within left-hemispheric front-parietal-temporal regions for the action-related sentences. The effects for action-related vs. abstract sentences were stronger for affirmative than for negative sentences. Furthermore, the activity level in pallid-cortical areas was lower for negated compared to affirmative sentences, irrespective of the semantic type of sentence.

According to Tettamanti et al. (2008), their results support the idea of reduced access to negated information but do not support the assumption of an increased processing load for negative sentences. Furthermore, the results, especially the reduction in the left front-parieto-temporal system, suggest that negated action-related sentences lead to a weaker activation of the action-representation system, supporting the embodied view of language representations. Additionally, they argue that deactivation in the left pallidum is an indication of reduced semantic processing in the case of negative sentences, which is compatible with the reduced accessibility hypothesis. The main result of this line of research on embodiment effects and negation is that the differences between action-related and abstract affirmative sentences disappear in negative sentences. See also Alemanno et al. (2012), Bartoli et al. (2013), and Papeo et al. (2016). As a result some researchers even argue that the meaning of negation is grounded on *disembodiment effects* (Bartoli et al., 2013; Pulvermüller, 2018).

Further support for the argument that the negated representation is activated in a first step and later inhibited and replaced by the representation of the affirmative alternative, as is defended by the two-step theory of negation comprehension (Clark and Chase, 1972; Carpenter and Just, 1975; Kaup et al., 2006, 2007; Lüdtke et al., 2008), comes from studies that included a manipulation of stimuli presentation times in their design where in one condition of for example a sentence-picture verification paradigm the picture is presented early after the sentence and only

about 500 or more milliseconds later in another condition. For example, in Kaup et al. (2007), pictures were presented either 750 ms or 1500 ms after the sentence. For the short-delay condition (750 ms) responses were faster when an affirmative sentence was followed by a picture that matched the actual state of affairs (that is, the affirmative alternative) compared with a picture that mismatched (that is, showing the negated situation). No such difference was found for negative sentences in the short-delay conditions. However, in the late-delay (1500 ms) conditions responses were faster when a negative sentence was followed by a picture that matched the actual state of affairs compared with a picture that mismatched this state of affairs. No such difference was found for the affirmative versions. Kaup et al. (2007) took this difference for the short- and long-delay conditions as an indicator for a two-stage comprehension process in which the comprehenders shift their attention from the negated state of affairs that is then suppressed toward the actual state of affairs at a certain point in time. See also Lüdtke et al. (2008) for an event-related potential study using the same material and the same experimental manipulations.

Direct support for inhibition comes from studies that directly investigated negation and inhibition in Stop-Signal paradigms or Go/No-Go tasks. A recent discovery in this context is the occurrence of a P300-ERP in response to negation. The P300-ERP is a positivity with a typical latency between 250 − 500 ms after the stimulus onset (Polich, 2007) that is commonly elicited in oddball paradigms in response to an infrequent stimulus. It is divided into two subcomponents: The P3a and the P3b. The anterior P3a is sensitive to new stimuli while the posterior P3b is elicited by infrequent target items that are focused or task-related and therefore expected in the experiment (Polich, 2003, 2007). Attention resources (Gray et al., 2004) as well as context updating and working memory updates (Donchin and Coles, 1988) have been argued to be reflected in the P300. Polich (2007) proposes (anterior) stimulus-driven attention processes during task processing to be reflected in the P3a and (posterior) attention-related processes and memory processing to be reflected in the P3b. In addition, he proposes neuroinhibition to be reflected in the P300 whenever a stimulus requires memory operations (Polich, 2007). See Huster et al. (2020) for a metaanalysis. Apart from the P3, the N1 ERP, which is a negative potential occurring around 80 − 120 ms after the stimulus onset, has also been linked to inhibition (see for example Raud and Huster, 2017). Furthermore, several studies have suggested that theta frequency activity plays a key role in inhibitory control (Dippel et al., 2017; Adelhöfer and Beste, 2020; Huster et al., 2020).

In an ERP study employing a Stop-Signal paradigm, Beltrán et al. (2018) investigated whether neural resources used for response inhibition are also used during the comprehension of negated sentences. Participants read affirmative or negative sentences (e.g., *You will cut the bread/You will not cut the bread*)[8] that were embedded in a Stop-Signal paradigm. When the Go cue in the form of an arrow appeared on the screen, participants had to respond with a button press to a recognition question. Instead, when the Stop cue occurred in the form of an acoustic signal, participants were asked to attempt to stop their response. The inhibition-related N1 and P3 ERP components were enhanced by successful inhibition, with the early N1 furthermore being modulated by sentence polarity resulting in the largest amplitudes in successful inhibition for negated sentences. For Go-trials and failed inhibitions, no polarity effects occurred. The results suggest that action-related negative sentences share neural resources with response inhibition.

de Vega et al. (2016) used the same type of action sentences in a Go/NoGo task measuring brain oscillations. They report reduced frontal theta oscillations for NoGo trials in negative compared to affirmative sentences. In light of theta oscillations being considered markers of neural inhibition, the theta oscillations being modulated by negation suggest the use of resources typically used for inhibition.

Subsequently Beltrán et al. (2019) investigated the interplay of inhibition and sentence comprehension combining action-related sentences with non-action sentences in a dual-task paradigm (sentence comprehension + Go/NoGo task). They report that the increased theta power for NoGo trials was affected by negative action and non-action sentences. Hence, the effects in de Vega et al. (2016) and Beltrán et al. (2019) cannot be reduced to being a result of the use of action-related sentences. Instead, the results suggest inhibitory control mechanisms to be part of the comprehension process of negated sentences. More specifically, cognitive control might be responsible for the monitoring of competing representations during language comprehension and for the resolution by inhibiting one representation and activating the other (Beltrán et al., 2019, see also Foroni and Semin, 2013). Beltrán et al. (2021) introduced the Neural Reuse Hypothesis that assumes that negation processing relies on the reuse of general domain inhibitory mechanisms.

Sommer et al. (2021) observe negation-related changes in theta power as well and suggest that negation has an inhibitory nature, but in an interactive nature with its function in context.

However, there are also studies contradicting the idea that inhibition is always involved in the comprehension of negated sentences. As mentioned above, inhibition requires activation, and therefore, studies that suggest that the negated information has not been activated in the first instance (Tian et al., 2010; Dale and Duran, 2011; Orenes et al., 2014, 2015) show that inhibition is not an inherent feature of negation comprehension. For example, when the context allows for a prediction in a negated sentence and hence, for a prediction within the set of *not p*, the representation of the underlying affirmative is not required (Dale and Duran, 2011; Orenes et al., 2014, 2015). Giora et al. (2007, 2009) argue that the deactivation, or rather suppression, of concepts in the study by MacDonald and Just (1989) is a result of the presentation of isolated sentences. In their view, suppression is not a default property of the negative operator and does not necessarily occur in longer discourse. Giora et al. (2007) report that in a discourse consisting of context sentence and target sentences, for example, *The train to Boston was no rocket. The trip to the city was fast though* the target concept *fast* was rather primed than suppressed in such a case compared to a discourse like for example *The train to Boston was no rocket. The old man in the film spoke fast.* Hence, suppression may not always be involved in the comprehension process of negated sentences. This speaks in favor of dynamic accounts of negation comprehension, that is, accounts that allow

8 The experiment was conducted in Spanish; the original sentences are for example *Ahora sí cortarás el pan/Ahora no cortarás el pan.*

for a context-dependent adaptation of the form of processing in negated sentences. Furthermore, it shows that the concurrent presentation of contextual information and target sentences, for example in the Eye-Tracking studies conducted by Orenes et al. (2014, 2015), Tian et al. (2016) are beneficial to further address this topic. Such experimental designs furthermore resemble the use of negation in communication where the choice of alternatives for a negated sentence is typically limited by the prior contextual information in discourse or by the current situation (e.g., choices that are visually present).

4 Discussion and conclusion

Both prediction and inhibition are domain-general cognitive mechanisms. Prediction is an essential part of general language comprehension, including the comprehension of negated sentences. Instead, inhibition seems to be involved in the comprehension of negated sentences, at least in those cases where an already activated representation of the negated situation needs to be inhibited. Prediction during language comprehension has repeatedly been reported to be reflected in the N400 component (DeLong et al., 2005; Pickering and Garrod, 2007, 2013; Altmann and Mirković, 2009; Kutas et al., 2011; Pickering and Gambi, 2018). Hence, simply put, prediction during language comprehension leads to a pre-activation of upcoming linguistic material. The ease of retrieval based on retrieval cues present in the linguistic input is reflected in the N400 (Brouwer et al., 2012, 2016; Brouwer and Hoeks, 2013). Instead, inhibitory mechanisms that lead to a deactivation of representations that became activated based on retrieval cues may be reflected in the P300 component that has been associated with general inhibitory mechanisms Beltrán et al. (2018). Both prediction and inhibition are not constrained to negated sentences. Further research is needed to understand, whether predictive and inhibitory processes in negated sentences follow the same mechanism as they do for affirmative sentences. A recent study by Spychalska et al. (2023) that investigates the comprehension of affirmative and negative sentences preceded by strongly or weakly constraining contextual scenarios shows that for prediction this seems to be the case. This is less clear for inhibition; negation seems to be an explicit cue to inhibit the negated information, whereas metaphor comprehension requires the inhibition of metaphor-irrelevant information, independent of sentence polarity and without an explicit cue to inhibit irrelevant information. It currently also remains unclear whether ”inhibition”, ”suppression” and ”deactivation” are the same mechanism. An alternative possibility might be that inhibition and suppression refer to a rather direct process, e.g., explicitly triggered by a cue such as negation. Instead, deactivation may be a more indirect process during which words and features that became predicted in parallel vanish once the retrieval of the actual linguistic input material happened. According to this distinction, one process (inhibition/suppression) requires some sort of attention switch that occurs at a specific moment during the comprehension (e.g., when the negation marker is encountered), whereas the other process (deactivation) might be more gradual, taking place over time without requiring an explicit switch of attention. Further research is necessary to tackle this speculation.

Because of the temporal overlap of the N400 and the P300, it may be difficult to dissociate effects related to inhibition from effects caused by predictive mechanisms. Studies that attempted to dissociate parallel N400 and P300 effects suggest that an increased P300 might lead to the superficial impression of a strongly reduced N400 (Roehm et al., 2007; Alday and Kretzschmar, 2019). In three ERP studies Roehm et al. (2007) dissociated the processes involved in the comprehension of antonomy relations from the general process involved in the comprehension of lexical-semantic relations. To this end, they compared the comprehension of antonym relations (*black-white*), of related (*black-yellow*), and of unrelated (*black-nice*) word pairs. The study aimed to dissociate influences of semantic relatedness which are independent on the experimental design from predictability processes which are dependent of the experimental design. In experiment 1 word pairs were presented in a sentential context of the form *The opposite of X is Y* and included a sensicality judgment. In experiment 2 and 3 word pairs were presented. While subjects were asked to respond to a lexical decision task in experiment 2, in experiment 3 they had to respond to an antonymy judgment task. Task and experimental design varied across experiments, but the lexical-semantic manipulation was kept constant. The authors report a graded N400 effect that was largest for unrelated words and smallest for antonyms, supporting the assumption that semantic associations are processed automatically. Additionally, the authors report a simultaneous N400 and P300 effect for the antonym condition in highly constrained task environments, leading to the superficial impression of a strongly reduced N400 for antonym pairs. The P300 effect across experiments and participant groups was not only dependent on sentence context but also on individual processing strategies that were employed for successful task performance. See also Delogu et al. (2021) regarding the temporal overlap of ERP components.

Coming back to the aspect of the role of the position of the negative marker with respect to the verb, we could assume two different options that both would have different impacts on prediction and inhibition. In a sentence with a preverbal negation with the structure SVO (subject, verb, object), the first argument (subject) is activated, and predictions are made under the assumption that the sentence is of affirmative polarity. When the negation occurs, it functions as a cue to reverse the polarity of the sentence and to direct the attention toward the complement set (*not p*). When the verb occurs, predictions for the remaining argument can be directly made within the complement set of *not p*. In principle, it can then be expected that predictions for the second argument are already made taking into account the negation. In an ERP study, an N400 at the position of the second argument would then reflect a violation of predictions within the complement set *not p*. However, this assumption stands in contrast to results by Dudschig et al. (2019), who showed that in sentences with an external negation (e.g., *It is not the case that ladybirds are stripy*), the result pattern does not differ compared to sentences where the negation is placed in a postverbal position within the sentences (e.g., *Ladybirds are not stripy*). One possible explanation might be that the semantic similarity between the entities (here the two arguments) in such sentences is a stronger cue than the negation and therefore, preactivation is strongly driven by the first argument, independent of the position of the negation. This may furthermore be reinforced by the use

of a copular verb that does not allow for strong (verb-based) predictions.

Instead, in a sentence with postverbal negation and SVO structure the assumption for the first argument would be identical, that is, the first argument (subject) is activated and predictions are made under the assumption that the sentence is of affirmative polarity. When the verb is processed, further predictions are being made for an affirmative sentence, and the second argument of the verb is predicted (based on the verb and within the set of *p*). A representation of an event is formed when the negation occurs, which functions as a cue to reverse the attention to the complement set, and furthermore requires an inhibition of the already (partially) activated representation. In an ERP study, an N400 at the position of the second argument could then reflect a violation of predictions within the complement set *not p* as well as a P300 for the inhibition of the representation of the underlying affirmative. However, this may be a time-consuming process, and especially without contextual constraints it might be difficult to adapt the predictions fast enough. This might explain results from Fischler et al. (1983) and Wiswede et al. (2013) and might also explain the reduced effects in Haase et al. (2019) where the alternatives were contextually constrained.

Let us remark again that full verbs might be the better stimuli material compared to copular verbs that bear little meaning and often result in relatively odd condition contrasts (such as *A robin is a bird* vs. *A robin is a truck* (Fischler et al., 1983)). In addition, copular constructions often lead to the priming of the second argument through the preactivation of common features with the first argument or even through the use of hyponyms. It may also be useful to control for the animacy of the entities as animate and non-animate entities have distinct features. Such an approach has been taken by Haase et al. (2019). See furthermore Khatin-Zadeh et al. (2019) for the investigation of inhibition in metaphors, dependent on animacy relations. Furthermore, there are cases in which a direct representation of the alternative to the negated situation is possible, e.g., in the case of polar adjectives but also in such cases in which the context provides only a binary choice of referents (cf. Orenes et al., 2014).

We have pointed out several aspects of the predictive and inhibitory mechanisms at hand during the comprehension of negated sentences. Further research combining both prediction and inhibition is necessary to better understand the mechanisms at hand. The overlap of functionally distinct components in the ERPs elicited during language comprehension requires cautious interpretations. Especially regarding the P300 and its relation to inhibition, we would also like to point out that Huster et al. (2020) suggest that effects other than inhibition might be reflected in the often observed P300 in response to what has been interpreted as inhibition. They suggest that other latent processes might be reflected in the P300, such as behavioral adaptations in the context of performance monitoring operations. Furthermore, we would like to emphasize the role of the verb in negated sentences. The results from Maess et al. (2016) and Chow et al. (2018) taken together may be taken as an indication that other processes than negation are time-consuming as well, such as the construction of argument role information that largely depends on the verb which in return largely affects the prediction in a sentence.

Future research needs to address the comprehension of negated sentences and how it is modulated by predictive and inhibitory mechanisms as well as how predictive and inhibitory mechanisms modulate the comprehension of negation. To disentangle these two processes, methods with a high temporal resolution seem most adequate, such as event-related potentials and Eye-Tracking. The use of contextual information not only provides a more natural setting than the presentation of isolated sentences does, but it furthermore allows to restrict the number of potential alternatives. A recent study by Spychalska et al. (2023) addresses the investigation of negation comprehension of affirmative and negative sentences preceded by logically comparable visual scenarios that are either strongly constraining for a specific noun occurring in the target sentence, or weakly constraining. They report an N400 on the target noun in weakly compared to strongly constraining conditions, for both sentences polarities. They argue that the often observed Truth by Polarity interaction, as reflected in the N400 ERP, is dependent on predictability, for affirmative as well as for negative sentences. The study furthermore suggests that inhibitory processes triggered by negation may be reflected in the P300 ERP. Farshchi et al. (2023) show that in the auditory modality, negative sentences seem to be processed less effortfully than in the visual modality. Haase et al. (2019) have shown that constraining the number of alternatives by using hyponyms of a joint hyperonym, that is, introducing a constraint regarding semantic features, seemed to slightly facilitate negation comprehension. However, this study made use of copular constructions. Future research with target sentences containing full verbs and contextual embedding is required to further investigate the temporal dynamics of prediction and inhibition in negated sentences. Further studies investigating the auditory modality would provide valuable insights into these processes under more realistic temporal conditions. Furthermore, the auditory presentation of stimuli allows for a modulation of alternatives by variations in the intonation. Ultimately, these designs may be combined with Go/NoGo tasks to directly investigate the inhibition processes in more detail. Apart from these experimental investigations, theoretical models and simulations of the semantic networks and modulations within them can address how activations spread across items and features in sentences with negation, making use of knowledge about predictive and inhibitory processes, including those addressed in this article.

Author contributions

VH: Conceptualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. We acknowledge support by the DFG Open Access Publication Funds of the Ruhr-Universität Bochum.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Adelhöfer, N., and Beste, C. (2020). Pre-trial theta band activity in the ventromedial prefrontal cortex correlates with inhibition-related theta band activity in the right inferior frontal cortex. *Neuroimage* 219, 117052. doi: 10.1016/j.neuroimage.2020.117052

Alday, P. M., and Kretzschmar, F. (2019). Speed-accuracy tradeoffs in brain and behavior: Testing the independence of P300 and N400 related processes in behavioral responses to sentence categorization. *Front. Hum. Neurosci.* 13, 285. doi: 10.3389/fnhum.2019.00285

Alemanno, F., Houdayer, E., Cursi, M., Velikova, S., Tettamanti, M., Comi, G., et al. (2012). Action-related semantic content and negation polarity modulate motor areas during sentence reading: an event-related desynchronization study. *Brain Res.* 1484, 39–49. doi: 10.1016/j.brainres.2012.09.030

Altmann, G. T. M., and Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cogn. Sci.* 33, 583–609. doi: 10.1111/j.1551-6709.2009.01022.x

Banaruee, H., Khoshsima, H., Khatin-Zadeh, O., and Askari, A. (2017). Suppression of semantic features in metaphor comprehension. *Cogent Psychol.* 4, 1409323. doi: 10.1080/23311908.2017.1409323

Bar, M. (2009). The proactive brain: Memory for predictions. *Philosoph. Trans. Royal Soc. B: Biol. Sci.* 364, 1235–1243. doi: 10.1098/rstb.2008.0310

Bartoli, E., Tettamanti, A., Farronato, P., Caporizzo, A., Moro, A., Gatti, R., et al. (2013). The disembodiment effect of negation: negating action-related sentences attenuates their interference on congruent upper limb movements. *J. Neurophysiol.* 109, 1782–1792. doi: 10.1152/jn.00894.2012

Beltrán, D., Liu, B., and de Vega, M. (2021). Inhibitory mechanisms in the processing of negations: a neural reuse hypothesis. *J. Psycholinguist. Res.* 50, 1243–1260. doi: 10.1007/s10936-021-09796-x

Beltrán, D., Morera, Y., García-Marco, E., and de Vega, M. (2019). Brain inhibitory mechanisms are involved in the processing of sentential negation, regardless of its content. Evidence from EEG theta and beta rhythms. *Front. Psychol.* 10, 1782. doi: 10.3389/fpsyg.2019.01782

Beltrán, D., Muñetón-Ayala , M., and de Vega, M. (2018). Sentential negation modulates inhibition in a stop-signal task. Evidence from behavioral and ERP data. *Neuropsychologia* 112, 10–18. doi: 10.1016/j.neuropsychologia.2018.03.004

Bornkessel-Schlesewsky, I., and Schlesewsky, M. (2009). *Processing Syntax and Morphology: a Neurocognitive Perspective*. Oxford: Oxford University Press.

Bornkessel-Schlesewsky, I., and Schumacher, P. B. (2016). "Towards a neurobiology of information structure," in *The Oxford Handbook of Information Structure*, eds. C. Féry, and S. Ishihara. (Oxford: Oxford University Press), 581–598.

Brouwer, H., Crocker, M., Venhuizen, N., and Hoeks, C. (2016). A neurocomputational model of the N400 and the P600 in language processing. *Cognitive Sci.* 12461, 1–35. doi: 10.1111/cogs.12461

Brouwer, H., Fitz, H., and Hoeks, J. (2012). Getting real about semantic illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Res.* 29, 127–143. doi: 10.1016/j.brainres.2012.01.055

Brouwer, H., and Hoeks, C. (2013). A time and place for language comprehension: mapping the N400 and the P600 to a minimal cortical network. *Front. Human Neurosci.* 7, 758. doi: 10.3389/fnhum.2013.00758

Bubic, A., von Cramon, Y., and Schubotz, R. (2010). Prediction, cognition and the brain. *Front. Hum. Neurosci.* 4, 25. doi: 10.3389/fnhum.2010.00025

Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mechanisms: evidence from event-related brain potentials. *Brain Lang.* 98, 159–168. doi: 10.1016/j.bandl.2006.04.005

Carpenter, P. A., and Just, M. A. (1975). Sentence comprehension: a psycholinguistic processing model of verification. *Psychol. Rev.* 82, 45–73. doi: 10.1037/h0076248

Carpenter, P. A., Just, M. A., Keller, T., Eddy, W., and Thulborn, K. (1999). Time course of fMRI-activation in language and spatial networks during sentence comprehension. *Neuroimage* 10, 216–224. doi: 10.1006/nimg.1999.0465

Chambers, C. D., Garavan, H., and Bellgrove, M. A. (2009). Insights into the neural basis of response inhibition from cognitive and clinical neuroscience. *Neurosci. Biobehav. Rev.* 33, 631–646. doi: 10.1016/j.neubiorev.2008.08.016

Chow, W.-Y., Lau, E., Wang, S., and Phillips, C. (2018). Wait a second! Delayed impact of argument roles on on-line verb prediction. *Lang. Cogn. Neurosci.* 33, 803–828. doi: 10.1080/23273798.2018.1427878

Clark, H. H., and Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cogn. Psychol.* 3, 472–517. doi: 10.1016/0010-0285(72)90019-9

Dale, R., and Duran, N. (2011). The cognitive dynamics of negated sentence verification. *Cogn. Sci.* 35, 983–996. doi: 10.1111/j.1551-6709.2010.01164.x

de Vega, M., Morera, Y., León, I., Beltrán, D., Casado, P., and Martín-Loeches, M. (2016). Sentential negation might share neurophysiological mechanisms with action inhibition. Evidence from frontal theta rhythm. *J Neurosci.* 36, 6002–6010. doi: 10.1523/JNEUROSCI.3736-15.2016

Delaney-Busch, N., Morgan, E., Lau, E., and Kuperberg, G. R. (2019). Neural evidence for Bayesian trial-by-trial adaptation on the N400 during semantic priming. *Cognition* 187, 10–20. doi: 10.1016/j.cognition.2019.01.001

Delogu, F., Brouwer, H., and Crocker, M. W. (2021). When components collide: spatiotemporal overlap of the N400 and P600 in language comprehension. *Brain Res.* 1766, 147514. doi: 10.1016/j.brainres.2021.147514

DeLong, K. A., Urbach, T. P., and Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nat. Neurosci.* 8, 1117–1121. doi: 10.1038/nn1504

Dippel, G., Mückschel, M., Ziemssen, T., and Beste, C. (2017). Demands on response inhibition processes determine modulations of theta band activity in superior frontal areas and correlations with pupillometry –implications for the norepinephrine system during inhibitory control. *Neuroimage* 157, 575–585. doi: 10.1016/j.neuroimage.2017.06.037

Donchin, E., and Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behav. Brain Sci.* 11, 357. doi: 10.1017/S0140525X00058027

Dudschig, C., Mackenzie, I. G., Maienborn, C., Kaup, B., and Leuthold, H. (2019). Negation and the N400: investigating temporal aspects of negation integration using semantic and world-knowledge violations. *Lang. Cogn. Neurosci.* 34, 309–319. doi: 10.1080/23273798.2018.1535127

Farshchi, S., Andersson, A., van de Weijer, J., and Paradis, C. (2023). Brain responses to negated and affirmative meanings in the auditory modality. *Front. Hum. Neurosci.* 17, 1079493. doi: 10.3389/fnhum.2023.1079493

Fauconnier, G. (1999). Creativity, simulation, and conceptualization. *Behav. Brain Sci.* 22, 615. doi: 10.1017/S0140525X99282143

Federmeier, K., Wlotko, E., De Ochoa-Dewald, E., and Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Res.* 18, 75–84. doi: 10.1016/j.brainres.2006.06.101

Fischler, I., Bloom, P., Childers, D., Roucos, S., and Perry, N. (1983). Brain potentials related to stages of sentence verification. *Psychophysiology* 20, 400–409. doi: 10.1111/j.1469-8986.1983.tb00920.x

Foroni, F., and Semin, G. R. (2013). Comprehension of action negation involves inhibitory simulation. *Front. Hum. Neurosci.* 7, 209. doi: 10.3389/fnhum.2013.00209

Freunberger, D., and Roehm, D. (2016). Semantic prediction in language comprehension: Evidence from brain potentials. *Lang. Cogn. Neurosci.* 31, 1193–1205. doi: 10.1080/23273798.2016.1205202

George, M. S., Mannes, S., and Hoffman, J. E. (1997). Individual differences in inference generation: An ERP analysis. *J. Cogn. Neurosci.* 9, 776–787. doi: 10.1162/jocn.1997.9.6.776

Giora, R., Fein, O., Aschkenazi, K., and Alkabets-Zlozover, I. (2007). Negation in context: a functional approach to suppression. *Discour. Proc.* 43, 153–172. doi: 10.1080/01638530709336896

Giora, R., Heruti, V., Metuki, N., and Fein, O. (2009). "When we say no we mean no": Interpreting negation in vision and language. *J. Pragmat.* 41, 2222–2239. doi: 10.1016/j.pragma.2008.09.041

Glucksberg, S. (2003). The psycholinguistics of metaphor. *Trends Cogn. Sci.* 7, 92–96. doi: 10.1016/S1364-6613(02)00040-2

Glucksberg, S., and Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychol. Rev.* 97, 3–18. doi: 10.1037//0033-295X.97.1.3

Glucksberg, S., and Keysar, B. (1993). *How Metaphors Work*. Cambridge: Cambridge University Press.

Glucksberg, S., Newsome, M. R., and Goldvarg, Y. (2001). Inhibition of the literal: filtering metaphor-irrelevant information during metaphor comprehension. *Metaphor Symbol* 16, 277–298. doi: 10.1080/10926488.2001.9678898

Gordon, P. C., Hendrick, R., and Johnson, M. (2001). Memory interference during language processing. *J. Exp. Psychol.* 27, 1411–1423. doi: 10.1037//0278-7393.27.6.1411

Gray, H. M., Ambady, N., Lowenthal, W. T., and Deldin, P. (2004). P300 as an index of attention to self-relevant stimuli. *J. Exp. Soc. Psychol.* 40, 216–224. doi: 10.1016/S0022-1031(03)00092-1

Haase, V., Spychalska, M., and Werning, M. (2019). Investigating the comprehension of negated sentences employing world knowledge: an event-related potential study. *Front. Psychol.* 10, 2184. doi: 10.3389/fpsyg.2019.02184

Hagoort, P., Hald, L. A., Bastiaansen, M. C. M., and Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science* 304, 438–441. doi: 10.1126/science.1095455

Hasson, U., and Glucksberg, S. (2006). Does understanding negation entail affirmation? An examination of negated metaphors. *J. Pragmat.* 38, 1015–1032. doi: 10.1016/j.pragma.2005.12.005

Horn, L. R. (1989). *A Natural History of Negation*. Chicago: University of Chicago Press.

Huster, R. J., Messel, M. S., Thunberg, C., and Raud, L. (2020). The P300 as marker of inhibitory control –fact or fiction? *Cortex* 132, 334–348. doi: 10.1016/j.cortex.2020.05.021

Kaup, B., and Dudschig, C. (2020). "Understanding negation," in *The Oxford Handbook of Negation* (Oxford: Oxford University Press), 635–655.

Kaup, B., Lüdtke, J., and Zwaan, R. A. (2006). Processing negated sentences with contradictory predicates: is a door that is not open mentally closed? *J. Pragmat.* 38, 1033–1050. doi: 10.1016/j.pragma.2005.09.012

Kaup, B., Lüdtke, L., and Zwaan, R. (2007). "The experiental view of language comprehension: How is negated text information represented?," in *Higher Level Language Processes in the Brain: Inference and Comprehension Processes*, eds. F. Schmalhofer, and C. Perfetti (Mahwah: Lawrence Erlbaum), 255–288.

Kaup, B., and Zwaan, R. A. (2003). Effects of negation and situational presence on the accessibility of text information. *J. Exp. Psychol.* 29, 439–446. doi: 10.1037/0278-7393.29.3.439

Kay, P. (1971). Taxonomy and semantic contrast. *Language* 47, 866. doi: 10.2307/412161

Khatin-Zadeh, O., Banaruee, H., and Yazdani-Fazlabadi, B. (2019). Suppression in metaphor comprehension: a perspective from distributed models of conceptual representation. *Neuroquantology* 17, 2. doi: 10.14704/nq.2019.17.1.1919

King, J., and Just, M. A. (1991). Individual differences in syntactic processing: the role of working memory. *J. Mem. Lang.* 30:580–602. doi: 10.1016/0749-596X(91)90027-H

King, J. W., and Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *J. Cogn. Neurosci.* 7, 376–395. doi: 10.1162/jocn.1995.7.3.376

Kuperberg, G. R., and Jaeger, T. F. (2015). What do we mean by prediction in language comprehension? *Lang. Cogn. Neurosci*. 31, 32–59.

Kuperberg, G. R., and Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Lang. Cogn. Neurosci.* 31:32–59. doi: 10.1080/23273798.2015.1102299

Kuperberg, G. R., Paczynski, M., and Ditman, T. (2011). Establishing causal coherence across sentences: An ERP study. *J. Cogn. Neurosci.* 23, 1230–1246. doi: 10.1162/jocn.2010.21452

Kutas, M., DeLong, K. A., and Smith, N. J. (2011). "A look around at what lies ahead: prediction and predictability in language processing," in *Predictions in the Brain: Using Our Past to Generate a Future*, ed. M. Bar (Oxford: Oxford University Press), 190–207.

Kutas, M., and Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science* 4:463–470. doi: 10.1016/S1364-6613(00)01560-6

Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647. doi: 10.1146/annurev.psych.093008.131123

Kutas, M., and Hillyard, S. A. (1980). Reading between the lines: Event-related brain potentials during natural sentence processing. *Brain Lang.* 11, 354–373. doi: 10.1016/0093-934X(80)90133-9

Kutas, M., and Van Petten, C. (1994). "Psycholinguistics electrified: Event-related potential investigations," in *Handbook of Psycholinguistics*, ed. M. A. Gernsbacher (San Diego: Academic Press), 83–143.

Lau, E. F., Holcomb, P. J., and Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single word contexts. *J. Cogn. Neurosci.* 25, 484–502. doi: 10.1162/jocn_a_00328

Lüdtke, J., Friedrich, C., De Filippi, M., and Kaup, B. (2008). Event-related potential correlates of negation in a sentence-picture verification paradigm. *J. Cogn. Neurosci.* 20, 1355–1370. doi: 10.1162/jocn.2008.20093

MacDonald, M., and Just, M. A. (1989). Changes in activation levels with negation. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 633–642. doi: 10.1037//0278-7393.15.4.633

Maess, B., Mamashli, F., Obleser, J., Helle, L., and Friederici, A. D. (2016). Prediction signatures in the brain: Semantic pre-activation during language comprehension. *Front. Hum. Neurosci*. 10, 591. doi: 10.3389/fnhum.2016.00591

Mayo, R., Schul, Y., and Burnstein, E. (2004). "I am not guilty" vs "I am innocent": Successful negation may depend on the schema for its encoding. *J. Exp. Soc. Psychol.* 40, 433–449. doi: 10.1016/j.jesp.2003.07.008

Münte, T. F., Schiltz, K., and Kutas, M. (1998). When temporal terms belie conceptual order. *Nature* 395, 71–73. doi: 10.1038/25731

Naumann, R., and Petersen, W. (2021). A theoretical framework for a hybrid view of the N400. *Front. Psychol.* 12, 678020. doi: 10.3389/fpsyg.2021.678020

Nieuwland, M. S. (2016). Quantification, prediction and the online impact of sentence truth-value: Evidence from event-related potentials. *J. Exp. Psychol. Learn. Mem. Cogn.* 42, 316–334. doi: 10.1037/xlm0000173

Nieuwland, M. S., and Kuperberg, G. R. (2008). When the truth isn't too hard to handle: An event-related potential study on the pragmatics of negation. *Psychol. Sci.* 19, 1213–1218. doi: 10.1111/j.1467-9280.2008.02226.x

Orenes, I., Beltran, D., and Santamaria, C. (2014). How negation is understood: Evidence from the visual world paradigm. *J. Mem. Lang.* 74, 36–45. doi: 10.1016/j.jml.2014.04.001

Orenes, I., Moxey, L., Scheepers, C., and Santamaría, C. (2015). Negation in context: Evidence from the visual world paradigm. *Quart. J. Exp. Psychol.* 69, 1082–1092. doi: 10.1080/17470218.2015.1063675

Papeo, L., Hochmann, J.-R., and Battelli, L. (2016). The default computation of negated meanings. *J. Cogn. Neurosci.* 28, 1980–1986. doi: 10.1162/jocn_a_01016

Payne, J. (1985). "Negation. language typology and syntactic description," in *Clause Structure*, ed. T. Shopen (Cambridge: Cambridge University Press), 197–242.

Pickering, M. J., and Gambi, C. (2018). Predicting while comprehending language: a theory and review. *Psychol. Bull.* 144, 1002–1044. doi: 10.1037/bul0000158

Pickering, M. J., and Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends Cogn. Sci.* 11, 105–110. doi: 10.1016/j.tics.2006.12.002

Pickering, M. J., and Garrod, S. (2013). An integrated theory of language production and comprehension. *Behav. Brain Sci.* 36, 329–347. doi: 10.1017/S0140525X12001495

Polich, J. (2003). "Theoretical overview of P3a and P3b," in *Detection of Change: Event-Related Potential and fMRI Findings*. (Boston: Kluwer Academic Press), 83–98.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clini. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019

Pulvermüller, F. (2018). Neural reuse of action perception circuits for language, concepts and communication. *Prog. Neurobiol.* 160, 1–44. doi: 10.1016/j.pneurobio.2017.07.001

Rabovsky, M. (2020). Change in a probabilistic representation of meaning can account for N400 effects on articles: a neural network model. *Neuropsychologia* 143, 107466. doi: 10.1016/j.neuropsychologia.2020.107466

Rabovsky, M., Hansen, S. S., and McClelland, J. L. (2018). Modelling the N400 brain potential as change in a probabilistic representation of meaning. *Nat. Human Behav.* 2, 693–705). doi: 10.1038/s41562-018-0406-4

Raud, L., and Huster, R. J. (2017). The temporal dynamics of response inhibition and their modulation by cognitive control. *Brain Topogr.* 30, 486–501. doi: 10.1007/s10548-017-0566-y

Roehm, D., Bornkessel-Schlesewsky, I., Rösler, F., and Schlesewsky, M. (2007). To predict or not to predict: influences of task and strategy on the processing of semantic relations. *J. Cogn. Neurosci.* 19, 1259–1274. doi: 10.1162/jocn.2007.19.8.1259

Rosch, E. (1975). Cognitive representations of semantic categories. *J. Exp. Psychol.* 104, 192–233. doi: 10.1037//0096-3445.104.3.192

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., and Boyes-Braem, P. (1976). Basic objects in natural categories. *Cogn. Psychol.* 8, 382–439. doi: 10.1016/0010-0285(76)90013-X

Rosch, E. H. (1973). "On the internal structure of perceptual and semantic categories," in *Cognitive Development and Acquisition of Language* (Amsterdam: Elsevier), 111–144.

Schumacher, P. B. (2012). "Context in neurolinguistics: time-course data from electrophysiology," in *What is a Context? Linguistic Approaches and Challenges*, eds. R. Finkbeiner, J. Meibauer, and P. B. Schumacher (Amsterdam/Philadelphia: John Benjamins), 33–53.

Schumacher, P. B. (2014). Content and context in incremental processing: "The ham sandwich" revisited. *Philos. Stud.* 168, 151–165. doi: 10.1007/s11098-013-0179-6

Sommer, J., Hansen-Schirra, S., Nagels, A., and He, Y. (2021). *Processing Linguistic Negation With and Without Truth-Value Evaluation: Evidence From N400 and Theta Oscillations*. Charlottesville: Center for Open Science.

Spychalska, M., Haase, V., Kontinen, J., and Werning, M. (2019). "Processing of affirmation and negation in contexts with unique and multiple alternatives: evidence from event-related potentials," in *Proceedings of the 41st Annual Conference of the Cognitive Science Society* (Montreal, QC: Annual Conference of the Cognitive Science Society), 2845–2851.

Spychalska, M., Haase, V., and Werning, M. (2023). To predict or not to predict: the role of context constraint and truth-value in negation processing. *PsyArXiv [Preprints]*. doi: 10.31234/osf.io/gtme6

Swaab, T. Y., Ledoux, K., Camblin, C. C., and Boudewyn, M. A. (2012). "Language-related ERP components," in *The Oxford Handbook of Event-Related Potential Components*, eds. S. J. Luck, and E. S. Kappenman (Oxford: Oxford University Press), 397–440.

Taylor, W. L. (1953). "Cloze Procedure": a new tool for measuring readability. *J. Quart*. 30, 415–433. doi: 10.1177/107769905303000401

Tettamanti, M., Manenti, R., della Rosa, P., Falini, A., Perani, D., Cappa, S., et al. (2008). Negation in the brain: modulating action representations. *Neuroimage* 43, 358–367. doi: 10.1016/j.neuroimage.2008.08.004

Tian, Y., Breheny, R., and Ferguson, H. (2010). Why we simulate negated information: a dynamic pragmatic account. *Q. J. Exp. Psychol.* 63, 2305–2312. doi: 10.1080/17470218.2010.525712

Tian, Y., Ferguson, H., and Breheny, R. (2016). Processing negation without context - why and when we represent the positive argument. *Lang. Cogn. Neurosci.* 31, 683–698. doi: 10.1080/23273798.2016.1140214

Van Petten, C., and Luka, B. J. (2012). Prediction during language comprehension: benefits, costs, and ERP components. *Int. J. Psychophysiol.* 83, 176–190. doi: 10.1016/j.ijpsycho.2011.09.015

Verhagen, A. (2010). *Construal and Perspectivization*. Oxford: Oxford University Press.

Villiers, J. G. D., and Flusberg, H. B. T. (1975). Some facts one simply cannot deny. *J. Child Lang.* 2, 279–286. doi: 10.1017/S0305000900001100

Wason, P. C. (1959). The processing of positive and negative information. *Q. J. Exp. Psychol.* 11, 92–107. doi: 10.1080/17470215908416296

Wason, P. C. (1965). The contexts of plausible denial. *J. Verbal Learn. Verbal Behav.* 4, 7–11. doi: 10.1016/S0022-5371(65)80060-3

Wiswede, D., Koranyi, N., Müller, F., Langner, O., and Rothermund, K. (2013). Validating the truth of propositions: behavioral and ERP indicators of truth evaluation processes. *Soc. Cogn. Affect. Neurosci.* 8, 647–653. doi: 10.1093/scan/nss042

Zanuttini, R. (2001). "Sentential negation," in *The Handbook of Contemporary Syntactic Theory*, eds. M. Baltin, and C. Collins. Hoboken, NJ: Blackwell Publishers. doi: 10.1002/9780470756416.ch16