Explicit and implicit negation, negative polarity, and levels of semantic representation

Ming Xiang, Julian Grove, Anastasia Giannakidou

Linguistics Department, University of Chicago

Corresponding author:

Ming Xiang

Language Processing Lab

Linguistics Department, University of Chicago, IL, Chicago, 60637

mxiang@uchicago.edu

phone: 773-702-8023

fax: 773-834-0924

Abstract

Most previous studies on negation have generally only focused on sentential negation (*not*), reporting cost and delay associated with computing its meaning. In an ERP study (Experiment 1), we make use of the negation-sensitivity of negative polarity items (NPIs) and examine the time course of processing different kinds of negation. Four kinds of NPI-licensing environments were examined: the negative determiner *no*, the negative determiner *few*, the focus marker *only*, and emotive predicates (e.g., *surprised*). While the first three contribute a negative meaning via assertion (explicit negation), the last gives rise to a negative inference via non-asserted content (implicit negation). Under all these environments, an NPI elicited a smaller N400 as well as a smaller late anterior negativity, compared to an unlicensed NPI, suggesting that negation, regardless of its source, is rapidly computed online, contrary to previous findings. However, we also observed that explicit negative meaning (e.g., as contributed by the first three licensors) and implicit negative meaning (contributed by emotive predicates) were integrated into the grammatical representation in different ways, leading to a difference in the P600, and calling for a separation of distinct levels of semantic integration during sentence processing. The qualitative differences between these conditions were also replicated in a self-paced reading study (Experiment 2).

Introduction

Polar opposition—negation vs. affirmation—is a fundamental distinction in human language. In English, an affirmative sentence such as *John came to school* can be denied by sentential negation *didn't*, as in *John didn't come to school*; a predicate such as *mortal*, can be negated by attaching *not* or a negative affix to it, i.e., *immortal* or *not mortal*. Negation is one of the distinctive properties of human language (Horn, 2001): every natural language includes at least one device that can express the negation of an affirmative constituent. Even in rudimentary linguistic systems such as home sign negation is one of the first structure building operations to emerge during language creation (Franklin, Giannakidou, and Goldin-Meadow 2011).

All else being equal, the semantic computation of negative sentences seems to be more complex than that of their affirmative counterparts, since negative statements involve an extra step of semantic processing, along with extra morphological or syntactic structure. Moreover, negation presents challenges for semantic and syntactic computation not only because it is an additional layer of meaning and structure to process, but also because there are many different ways to express negation. *Not many students came to school* is approximately synonymous with *Few students came to school*. Likewise, *John didn't believe Mary would win* expresses a similar negative meaning to *John doubted Mary would win*. The necessity of identifying different types of negative expressions, based on morphosyntactic, semantic, or pragmatic cues, enhances the processing complexity of negation (see Horn (2001) for a detailed and lucid discussion of the issues of complexity, or markedness, of negation).

Even though negation is extremely common in everyday communication, the time course of comprehending negation remains poorly understood. Two issues are of particular interest in this area. First, one striking result consistently observed is that the processing of sentential negation *not* is particularly difficult and involves a slow time course (Wason, 1959; 1961; Clark and Chase, 1972; Fischler, Bloom, Childers, Roucos and Perry, 1983; Kounios and Holcomb, 1992). This aspect of negation makes it stand out as an exception within the large body of literature showing that semantic comprehension is highly incremental—fine-grained semantic (and pragmatic) information from different sources is rapidly accessed and processed online, forming the basis of context-updating and forward semantic prediction (e.g., Altmann and Steedman, 1988; MacDonald, Pearlmutter, and Seidenberg, 1994; Tanenhaus, Spivey-Knowlton, Eberhard, and Sedivy, 1995; van Berkum, 2009). Second, since most of this research has focused on the negative form *not*, we know very little about how other types of negation are processed. In the current study, we make use of the negation-sensitive property of the closed-class items known as "negative polarity items" (such as *ever*) to investigate how negative meaning is extracted from a number of different negative environments. These environments fall under two general classes that we call explicit (asserted) and implicit (non-asserted) negation.

Explicit and implicit negation

Natural language has a rich landscape of negative expressions. There are a number of dimensions we can use to classify negative expressions into groups. In the current paper, following Clark (1976)[1], we make a distinction between negation in the *asserted* meaning and negation in the *non-asserted* content (also see Horn (1996)); also following Clark (1976), we call the first group *explicit* negation, and the second *implicit* negation. Under Clark's classification, explicit negation in English includes expressions like *scarcely*, *hardly*, *few*, *seldom*, *little*, and *only*, as well as more obviously negative expressions like *no*, *not*, and *never*. Implicit negation, on the other hand, includes expressions like *forget*, *fail*, *doubt*, and *deny* (see also Fodor, Fodor, and Garrett, (1975)).[2] It is already clear from these examples that *explicit* negation is not a label for *morphologically* explicit (or overt) negation, though overt negation is indeed a core member of this category. Rather, the contrast between explicit versus implicit negation relies on which level of semantic representation, i.e., assertion or non-assertion, negation appears at, a distinction we come to below.

Any given utterance conveys an array of meanings. In the widely used Gricean and neo-Gricean frameworks (Grice, 1975; Stalnaker, 1978, Horn 2001, Geurts, 2010), assertion conveys the logical meaning of a sentence, and non-asserted meaning is thought of as *pragmatic* meaning. Logical semantic meaning determines the literal meaning of a sentence, i.e., the truth conditions and *entailments* of a sentence; pragmatic meaning, on the other hand, includes inferences beyond entailments, specifically presuppositions, conversational implicatures, and conventional implicatures. For example, if I say *Mary's children are blond* I am asserting that Mary's children are blond, the sentence entails that

[1] Clark (1976) was specifically making a distinction between asserted negation and negation in the presupposition. It should be pointed out that presupposition is only one type of non-asserted meaning.

[2] We agree with Clark (1976) on the general distinction between asserted and non-asserted negation. We do not necessarily adopt his specific classification of verbs like *doubt, deny*, etc., but we leave the discussion on these verbs open since the current study does not address these verbs.

Mary's children are blond, and it is true if they are, and false if they are not. The sentence also conveys the information that Mary has children, and this is, classically, a presupposition of the sentence, not an entailment. A common diagnostic to distinguish presupposition from entailment is that presupposition survives under negation, but entailment does not. For example, with a negative sentence like *Mary's children are not blond*, it is longer entailed that Mary's children are blond. But the presupposition that Mary has children remains intact. The precise division of labor between asserted and non-asserted meaning is a central issue in the study of linguistic meaning (for recent overviews, see Potts (2005, 2007), Tonhauser, Beaver, Roberts, and Simons, (2013)).

We can define two classes of negation based on their source of negative meaning. If negation is expressed as part of the asserted meaning of an utterance, i.e., if it is an entailment, it is ***explicit*** negation; if it belongs to the non-asserted meaning (i.e., presupposition or implicature), it is ***implicit*** negation. Overt negation, such as *no* and *not*, mark grammatical negation and obviously affect the assertion, as we just saw. They constitute explicit negation. But it is important to note that explicit negation does not necessitate that negation is morphologically overt. Expressions such as *few*, *scarcely*, *hardly*, *seldom*, and *little*, although not morphologically realized as negative, behave nevertheless syntactically and semantically negative under a number of well known, and by now classic, diagnostics (Klima, 1964; Horn, 2001; Postal, 2005; etc.). For instance, *few*, *scarcely*, *hardly*, *seldom*, and *little* can be followed by a conjunct modified by *neither*, but not by *so*. Moreover, they may also co-occur in a conjunct with *either*, but not with *too*. Some examples are given below:

So/*Neither*-diagnostic

(1) a. Those students passed the exam, and so/*neither did the teachers.

b. *No* students passed the exam, and *so/neither did the teachers.

c. *Few* students passed the exam, and *so/neither did the teachers.

d. Those students *hardly* passed the exam, and *so/neither did the teachers.

Too/either diagnostic

(2) a. The students left, and all the teachers left too/*either.

b. The students left, and *none* of the teachers stayed *too/either.

c. The students left, and *few* of the teachers stayed *too/either.

d. The students left, and the teachers *hardly* stayed *too/either.

At the same time, it has also been noted that these syntactic diagnostics are sufficient but not necessary properties of explicit negation. For some cases of explicit negation, although a negative meaning is asserted, the syntactic tests above are not applicable. Consider, e.g., the *exclusive* focus particle *only*. By saying *Only John read the article*, one asserts content equivalent to that asserted by the exceptive sentence *Nobody other than John read the article.* It is generally agreed that this negative exclusive component is part of the asserted meaning of *only* (and of negative exceptives in general; Horn, 1996, 2002; Atlas, 1993; Beaver and Clark 2008)[3], even though *only* does not have negative morphology, and is not syntactically negative based on the tests above.[4]

---

[3] What is not agreed upon is whether the meaning that *John read the article* is also part of the assertion of the sentence *Only John read the article*. Opinions differ here, with those that believe it part of the assertion (Atlas, 1993) and those that believe it is a presupposition (Horn, 1996, 2002; Beaver and Clark, 2008). This debate is not important for our purposes here, but see Giannakidou (1998, 2006) on how it may impact NPI-licensing cross-linguistically.

[4] There are good reasons why the syntactic tests in (1) and (2) do not apply to *only* (or negative exceptives in general, such as *nobody other than*). For example, one cannot say **Only Bill read the newspaper*, *and John did either*. The particles *either* or *neither* are additive – they would require somebody other than Bill to read the newspaper – which clashes with the negative

This brief discussion shows that semantically asserted negation does not map uniformly onto syntactic or morphological negation. Some instances of explicit negation contain overt negative morphology (*no*, *nobody other than*); some contain no overt morphology but pass syntactic diagnostics of negation (*few*); and yet others are neither morphologically nor syntactically negative, but nevertheless <u>assert</u> a negative meaning (*only*). We call all these cases in which negation is an entailment of the sentence "explicit" negation, regardless of their morphosyntactic realization.

Implicit negation, on the other hand, involves negative meaning whose source is pragmatic (presupposition or implicature). For current purposes, we consider the class of "emotive" predicates, which trigger negative inferences, though their negative content is not asserted. Emotive verbs depict certain emotions or attitudes (hence the term "emotive") towards the content of an embedded clause, which is presupposed to be true (hence the term "emotive factive"; Kartunnen, 1971; Kiparski and Kiparski, 1970). Examples of emotive predicates include *be amazed*, *be sorry*, *be surprised*, *be lucky*, *be disappointed*, *be irritated*, *regret*, etc. Consider the sentences given below:

(3) a. John was amazed that the tofu was so delicious.

b. John was lucky that he passed the exam.

c. John was surprised that he got the last ticket to the game.

Take (3a) as an example. The sentence obviously makes an affirmative assertion: it asserts that John experienced a particular psychological state (e.g., amazement) about the fact that the tofu was delicious. Although this is all that is asserted by the sentence, a person who hears an utterance like this is likely to draw a negative inference that John

---

exceptive meaning of *only*. So, the fact that *only* fails the *either* test does not tell us anything about its negativity, it is merely a case where the test cannot be applied.

didn't expect the tofu to be delicious. This negative inference has been characterized in the literature as an implicature (Linebarger, 1980), or presupposition (Baker, 1970; Giannakidou, 2006, to appear). Baker (1970) describes the negativity of emotive predicates by saying that they express "a relation of contrariness between a certain fact and some mental or emotional state. For example, we say that we are *surprised* when a certain fact does not conform to our expectations; *relieved* when it does not conform to our fears; *disappointed* when it is not in line with our hopes; and *lucky*, if it is not in line with some standard set of probabilities."

To summarize, negative meaning can arise from two sources: either from the assertion (what is said) or from the non-asserted content (presuppositions or implicatures). We call the former "explicit negation", and the latter "implicit". The current study looks at how these different types of negation are computed online.

Incremental comprehension, negation, and the N400

Previous processing studies on negation have largely focused on sentential negation *not*, as in *A is NOT B* constructions. Sentence verification tasks (Wason, 1959; 1961; Clark and Chase, 1972) have shown that negative sentences are generally more difficult for subjects to verify than their affirmative counterparts. Studies that have employed online comprehension techniques, such as ERPs (Fischler, Bloom, Childers, Roucos and Perry, 1983; Kounios and Holcomb, 1992), have also shown that negation does not seem to come into play soon enough to influence the N400 amplitude of an upcoming target word. Sentence pairs like *A robin is*/*is not a bird* (Fischler et al., 1983), despite obvious truth value differences, did not produce N400 differences on the critical word *bird*, suggesting lexical semantic associations (e.g., 'robin'-'bird'), instead of the truth value of the sentence, modulated the N400 amplitude in this case. The general hypothesis that has been adopted to account for these findings is that the comprehension of a negative proposition is decomposed into stages: the affirmative subject-predicate relation is processed first, and after that, the negative logical relation is processed (Clark and Chase, 1972; Wason, 1961; Gough, 1965; Trabasso, Rollins, and Sharghnessy, 1971), leading to extra cost, as well as delayed processing of negation.

Such a hypothesis, however, is surprising given the large body of research that has consistently shown semantic knowledge to be rapidly integrated into online comprehension (e.g., Altmann and Steedman, 1988; Tanenhaus, Spivey-Knowlton, Eberhard, and Sedivy, 1995). If the comprehension of negation were indeed delayed until after the rest of the proposition is processed, this would undermine the highly incremental view of semantic comprehension. The view that negation constitutes a counterexample to the strong incremental hypothesis, however, is challenged by the ERP results of Nieuwland and Kuperberg (2008), who showed that the observed processing cost for sentential negation may arise from difficulty in accommodating pragmatic conditions that are necessary prerequisites for uttering a negative statement ("plausible denial", Wason, 1965). Once the relevant pragmatic conditions on denial are met, sentential negation turns out to be incrementally processed and influences the N400 on an upcoming word (see also similar proposals from a sentence verification task in Greene (1970)).

Although most previous studies have focused on the sentential negation expressed by *not*, there has also been a certain amount of discussion of other types of negation. The general observation seems to be that these other types of negation are costly, as well, and are not immediately accessed in online processing. In an ERP study, Urbach and Kutas (2010) found that the negative quantifier *few* and the negative adverb *rarely* were neither immediately processed nor completely delayed in their processing. True sentences involving negation, like *Few farmers grow worms* produced larger N400s on *worms* than the false sentence *Few Farmers grow crops* does on *crops*, countering what one might expect if negation were immediately accessed to facilitate the processing of the upcoming words in a sentence. At the same time, however, the N400 difference in the pair above was smaller than the affirmative pair *Most farmers grow worms/crops*, suggesting that the negative quantifier *few* is not totally ignored either. Qualitatively similar ERP patterns were also found for *rarely*. These results are consistent with previous behavioral results in Just and Carpenter (1971), in which a sentence verification task showed extra processing cost on negative expressions like *few*, *scarcely*, and *hardly*.

To summarize, previous studies are not totally conclusive as to whether or not the processing of negation is delayed until the point at which the processing of an affirmative proposition is complete. A number of factors may have influenced the results and interpretation of these studies. First, as shown by Nieuwland and Kuperberg (2008), pragmatic conditions heavily influence the online comprehension of a negative statement (also see Tian, Breheny, and Ferguson (2010)). A denial is only appropriate in contexts which assume the corresponding affirmative statement (the "plausible denial" condition, Wason, 1965). *A robin isn't a bird* and *A robin isn't a vehicle* are both negative, but negation in the second sentence is "pragmatically unlicensed", since it is not plausible to expect that *A robin is a vehicl*e is true in the first place. Second, since the majority of studies have only focused on the sentential negation *not*, it is not totally clear yet what generalizations one can draw in regard to the processing profile of different types of negation. And finally, although many researchers have discussed their results in terms of whether the processing of negation is delayed or not, a closer look at the experimental designs actually suggests that they do not directly assess this question. We discuss this last point in more detail below.

In previous ERP studies, the evaluation of how negation is processed has often been carried out in tandem with that of how the truth or falsity of a proposition is evaluated with respect to the real world. Normally, the baseline conditions set up a pair of affirmative propositions, and the subject-predicate relation determines whether the statement is true or false, as in *A robin is a bird*/*vehicle*, or *Most farmers grow crops*/*worms*. Then, negation is applied to these two conditions to create two negative conditions, in which the truth/falsity is reversed compared to the first two statements. ERP responses are measured at the predicate, i.e., *bird*/*vehicle*, or *crops*/*worms*. A reversed N400 on the predicates in the negative pair of conditions is taken to be evidence that negation can be processed rapidly and incrementally online, whereas if the N400 difference on the negative pair patterns with the affirmative baseline conditions (i.e., if it is not reversed), this would suggest delayed processing of negation. This reasoning, however, overlooks the important point that the incremental comprehension of negation is only a necessary, but not a sufficient condition to drive a reversed N400 effect online.

Generally speaking, the amplitude of the N400 evoked by an incoming word indexes the degree to which that word's semantic features match semantic features that have been pre-activated by its context at the time it is encountered (Kutas and Federmeier, 2011; Lau et al., 2008). The "semantic match" between the upcoming word and the current context, and hence the N400 amplitude, can be modulated at multiple different levels of semantic relations. In some situations, the context might be constrained enough that it guides people to make active forward prediction about the upcoming words (e.g., Federmeier and Kutas, 1999; Delong et al., 2005); in other cases, however, the N400 may be modulated simply by the lexical semantic associations between words (e.g., the bag of words model). What factors can modulate N400 amplitude, and how they interact with each other, remain questions under intense debate (see Kuperberg (2013) for a recent overview). In the context of negation, whether or not negation is computed in time is only one of the relevant factors that could have influenced the N400 amplitude. In particular, the failure to obtain a reduced N400 on the true negated sentence *A robin isn't a vehicle* compared to the false one *A robin isn't a bird* could result from at least three different sources, or any combination of them: (i) the delayed negation hypothesis: negation isn't processed until the corresponding affirmative proposition is processed completely; (ii) there is a closer lexical semantic association between *robin* and *bird* than between *robin* and *vehicle*, and this competes with the discourse level semantic representation to modulate the N400; (iii) prior to the critical word *vehicle* the negative context might not have been constraining enough to encourage any semantic prediction of the upcoming word. In fact, the "pragmatically licensed" negation in Nieuwland and Kuperberg (2008) may have helped to reverse the online N400 effect by aiding online prediction in the true negative sentences (e.g., *With proper equipment, scuba-diving isn't very <u>dangerous</u>...*).

Given these considerations, the results in previous studies overwhelmingly showing that negation fails to aid the processing of upcoming words don't directly address the question of whether or not the online computation of negation itself, including different kinds of negation, is delayed. In the current study, instead of manipulating the subject-predicate relation, as has been done previously, we make use of the negation sensitive property of the NPI (negative polarity item) word *ever*, a closed-class word, to gauge the time course of computing negation. Since *ever* is closed-class, it can largely circumvent issues of low-level lexical semantic associations (e.g., *robin-bird*); and, since its distribution is conditioned based on grammatical factors, it can avoid processing effects based on real-world knowledge. It therefore helps in conducting a controlled test of the time course of processing negation alone. Another advantage of the current study is that the word *ever* is sensitive to both explicit and implicit negation, which allows us to test multiple different expressions of negation in the same design and on the same population.

Negative polarity items and previous ERP findings

Negative polarity items (NPIs), such as *ever* and *any,* as their name suggests, are lexical items that need to be *licensed* by negation.[5] There is a large linguistic literature discussing the syntactic, semantic, and pragmatic mechanisms that support NPI licensing (Ladusaw, 1979; Zwarts, 1986, 1995; Baker, 1970; Linebarger, 1987; Giannakidou, 1998, 2006, 2011; Hoeksema, 1994, 2012; von Fintel, 1999; Krifka, 1995; Chierchia, 2006; Kadmon and Landman, 1993; Laka, 1994). For current purposes, we will focus on four types of NPI licensors that contain some sort of negative feature, explicit or implicit. These licensors are the negative determiners *no* and *few*, the exclusive focus particle *only*,

[5] Although negation is the most robust licensing environment cross-linguistically for NPIs, NPIs can also appear in non-negative contexts that are non-veridical (Giannakidou, 1998, Zwarts, 1995). We do not discuss non-negative contexts in this paper.

and emotive predicates such as *surprised*, *amazed*, etc. As discussed above, the first three words contain explicit negation, but emotives provide an example of implicit negation. An example of NPI licensing under these expressions is given in (4) below. We also included a control example (4e) in which the NPI is not licensed, resulting in an ungrammatical sentence.

(4) It is hard to train a dog.

a. No dogs Andrew owns have ever responded to commands.

b. Few dogs Andrew owns have ever responded to commands.

c. Only dogs Andrew owns have ever responded to commands.

d. Andrew is surprised that the dogs he owns have ever responded to commands.

e. *The dogs Andrew owns have ever responded to commands.

In the (a) and (b) examples above, the NPI *ever* is licensed by the negative meaning contained in the negative quantifiers *no* and *few*. In (c), it is licensed by the exclusive component of *only* (*Only* NP VP asserts [*Nobody else other than* NP] VP.). In (d), the emotive *surprised* presupposes (Giannakidou, 2006, to appear) or implicates (Linebarger, 1980), rather than asserts a negative meaning, and *ever* in this case is licensed via this non-asserted implicit negation (Baker, 1970; Linebarger, 1987; Giannakidou, 2006).[6]

Besides semantic differences, the licensors in (a-d) above also differ along morphological and syntactic dimensions. The licensor *no* is the canonical morphological expression of negation, in addition to being both syntactically (bearing the syntactic

[6] There are also approaches under which emotive predicates as *semantic* licensors that license NPIs through downward entailment (e.g. von Fintel, 1999, Strawson downward entailment). Results reported in this study actually pose challenges to such approach, since this approach would likely to make the prediction that emotive predicate should be processed just like other semantic licensors, contrary to what we found. Going into a detailed discussion will lead us too far from the current focus, and we will not pursue this discussion further in this paper.

feature [+Neg]) and semantically negative; *few* is also syntactically and semantically negative, but does not contain an overt negative morpheme; *only* asserts a negative meaning through its exceptive component (*nobody other than),* but it is neither morphologically nor syntactically negative. Finally, emotive verbs are neither syntactically nor morphologically negative, and they only contribute negation via non-asserted content. We present the differences between these expressions in Table 1.

Given that it is the negative meaning extracted from these expressions that licenses an NPI in English, the online time course of NPI licensing under these licensors provides important evidence to address the question of whether or not negation can be rapidly processed online, and, moreover, whether or not the time course of interpreting negation varies depending on its source. In the current study, we employ highly time-sensitive ERP measures to investigate the time course of processing different negative licensors. The particular ERP components of interest are the N400 and the P600.

Table 1: The source of negation for the four different NPI licensors

	Asserted negation			Non-asserted negation
	Morphologically negative	Syntactically negative	Negation in semantic assertion	Presupposed or conventionally implicated
No	✔	✔	✔	–
Few	–	✔	✔	–
Only	–	–	✔	–
Emotive	–	–	–	✔

<u>*N400 and NPI licensing*</u>

The N400 is a negative-going waveform that peaks at approximately 400ms, with a primarily centro-posterior scalp distribution. The amplitude of the N400 evoked by an incoming word indexes the degree to which that word's semantic features match semantic features that have been pre-activated by its context at the time it is encountered (Kutas and Federmeier, 2011). The term "pre-activation" has often been associated with active prediction of specific lexical items. But, we use the term in a more neutral sense: we use it to refer to the activation of relevant semantic features, regardless of whether active prediction or expectation of the upcoming word is at work, ahead of encountering the full linguistic input. In the context of NPI licensing, given that negation is cross-linguistically the most robust licensor, it is reasonable to assume that the abstract lexical semantic features of an NPI, such as *ever*, include the [+Neg] feature. During the incremental comprehension of a sentence, if a semantic [+Neg] feature is compositionally derived prior to encountering *ever*, it should lead to a reduced N400 on the NPI word. We want to note that we use '[+Neg]' to represent a semantic negation compositionally derived at the global discourse level, instead of a lexical feature associated with a particular lexical licensor. As we show in Table 1, not all licensors contain a lexical semantic or morphosyntactic negative feature. But crucial for our purposes here is that all of the licensors in Table 1 contribute *a negative sentential meaning* in one way or another. If such a negative meaning were computed prior to the NPI word, one would expect N400 reduction on the NPI word.

The first ERP study on NPI licensing is reported in Shao and Neville (1998), who found a larger frontal negativity between 300-500ms on unlicensed *ever* than on its grammatical control *never*, as in the sentences *Max says he has *ever/never been to a* *birthday party*. However, one potential drawback of this study is that different lexical items (*ever* and *never*) were compared. In another series of studies on German (Saddy, Drenhaus, and Frisch, 2004; Drenhaus et al., 2005, 2006; Drenhaus, Blaszczak, and Schutte, 2007), a reduced N400 with a central maximum was found on the German NPI *jemals* ('ever') when it was licensed by negation, as in *No man who had a beard was ever happy*, compared to the ungrammatical counterpart when *ever* was not licensed. Similar N400 effects were also found for Dutch (Yurchenko et al., 2013) and Turkish NPIs (Yanilmaz and Drury, 2013) and in an MEG study by Tesan, Johnson, and Crain (2012). These findings are in line with our hypothesis that a pre-activated negation feature can facilitate the lexical processes implicated in interpreting the NPI. Note that only the negative expressions *no* and *not* were used in these studies. Therefore, although these findings at least suggest that the negative feature associated with *no* and *not* was successfully computed prior to encountering the NPI, it isn't clear whether or not this result can be extended to other types of negative expressions.

Interestingly, another study by Steinhauer, Drury, Portner, Walenski, and Ullman (2010) did not find an N400 difference between licensed and unlicensed *ever*[7], but one crucial difference between their study and the others mentioned above is that Steinhauer et al. (2010) had a larger set of licensors in their stimuli, including various negative licensors such as *not*, *without*, *rarely*/*hardly*, and also licensors that are not negative per se, but non-veridical, such as *every*, *before*, *whether*, and yes-no questions. It is possible that the N400 reduction on the NPI *ever* can only be triggered by pre-activated negative features (see Xiang et al. (2013) for more discussion), and therefore, that the N400 effect

[7] Multiple NPIs were tested in Steinhauer et al. 2010. N400 difference was only observed for "at all", but not for "ever" or "any".

in Steinhauer et al. (2010) could have been wiped out by the use of both negative and non-negative licensors. Finally, it is also worth noting that the complexity of the experimental stimuli seems to affect whether or not an N400 effect will emerge. Xiang, Dillon, and Phillips (2009) found no N400 effect for English *ever* under either licensors *no* or *few*. This may have resulted from the fact that in this study, there was a long and complex relative clause intervening between the licensor and the NPI.

The existing ERP literature does seem to suggest that it is plausible that a pre-activated semantic feature [+Neg] can facilitate the processing of an upcoming NPI, leading to a reduced N400 on the NPI itself. The N400 amplitude on the NPI, therefore, can help gauge the time course of the computation of various kinds of negative information. For the four different licensors we examine in this study—*no*, *few*, *only*, and emotive factives—if the negative information on these licensors has been successfully computed by the time an NPI is encountered, we expect a reduced N400 on the NPI under all of them relative to the ungrammatical condition in which there is no licensor. Conversely, if these licensors trigger different degrees of N400 reduction, then important information is provided about the different time courses of computing different types of negation.

<u>*P600 and NPI licensing*</u>

The majority of the studies reviewed above also reported a posteriorly distributed P600-like late positivity effect, which was larger for unlicensed NPIs than for licensed ones[8]. The P600 effect was originally associated with syntactic processing, since it is

[8] To our knowledge, the only two studies that did not find a P600 are Saddy et al. (2004) and Yurchenko et al. (2013). The original data from Saddy et al. (2004) was reanalyzed in Drenhaus

reliably elicited by syntactic errors (Hagoort, Brown, and Groothusen, 1993; Osterhout and Holcomb, 1992) or grammatical but syntactically complex constructions (Kaan et al., 2000; Osterhout et al., 1994; Phillips et al., 2005; Gouvêa, Phillips, Kazanina, and Poeppel, 2010); on the other hand, there is a growing body of work on the "semantic P600" effect (Kim and Osterhout, 2005; Kuperberg, 2007; Van de Meerendonk, Kolk, Chwilla, and Vissers, 2009; Bornkessel-Schlesewsky and Schlesewsky, 2008; Paczynski and Kuperberg, 2012; Chow and Phillips, 2013; Brouwer, Fitz, and Hoeks, 2012), showing that words that are semantically implausible within their context can also elicit a large P600. Although the precise functional interpretation of the P600 is yet to be determined, a broad generalization that has emerged is that it reflects costs associated with a processing stage in which information from different sources is integrated into one coherent representation (Friederici and Weissenborn, 2007; Kuperberg, 2007; Bornkessel and Schlesewsky, 2008; van Petten and Luka, 2012). Increased P600 amplitudes signal the detection of an integration error or integration difficulty, including costs associated with the process of reanalysis.

In the particular context of NPI licensing, multiple streams of information—syntactic, semantic, and pragmatic—are recruited to construct a grammatical representation that can license NPIs. In an ungrammatical sentence that does not license NPIs, the comprehension system fails to integrate an NPI into the current grammatical representation, and therefore produces a large P600. If the P600 broadly indexes the integration effort with which an NPI is licensed, it provides a useful tool to examine

et al. (2006) using a symbolic resonance analysis and a hidden P600 was discovered. For Yurchenko et al. (2013), the authors acknowledged that the lack of a P600 may be due to insufficient power in the data, as well as, potentially, to task-specific effects.

whether or not various kinds of negation are recruited in different ways for the grammatical purpose of NPI licensing. Specifically, if asserted negation (e.g., *no*, *few*, and *only*) and implied/non-asserted negation (e.g., emotive predicates) are adopted by the comprehension system as equally viable licensors, they should present similar P600 profiles relative to an unlicensed NPI.

Combining observations from the N400 and the P600 time windows, we will be able to construct a complete picture as to when and how negation is computed and used for grammatical purposes. In particular, the N400 amplitude reveals information about whether or not a negative meaning is established incrementally in context; the P600 amplitude assesses whether negative meaning, if available, can be immediately adopted to serve the grammatical function of NPI licensing.

Experiment 1

Methods

Stimuli Creation

One-hundred-and-fifty items like (4) were constructed, each with five conditions. An example is given in Table 2. All items contained a context sentence, which was followed by a target sentence containing the critical word *ever*. For each item, five conditions were created by varying the determiner on the subject NP of the target sentence for conditions a, b, c, and e, and by embedding the target sentence in the complement of a factive predicate for condition d. The context sentence and the remainder of the target sentence were kept identical across conditions. The first condition contained a subject NP with the quantificational determiner *no* (as in *no third-graders*…); the second condition contained *few* (as in *few third-graders*…); the third condition contained *only* (as in *only third-graders*…); the fourth condition embedded the target sentence in the complement of a factive predicate like *strange* (as in *It's strange that third-graders*…); and the fifth condition contained either a definite determiner (as in *The third-graders*…) or a bare plural (as in *Third-graders*…) as an ungrammatical control, in which the NPI *ever* is not licensed. The fourth condition containing emotive predicates used a range of different lexical items as stimuli, among which the seven most frequent ones were *amazed*, *amazing*, *surprised*, *surprising*, *lucky, glad*, and *shocked*. These predicates formed 90% of the items in this condition.

Table 2: An example of the experimental stimuli and the acceptability rating for each condition on a 1-7 scale.

Context:	The teacher brought a tarantula to class.	Acceptability rating (1-7)
No	a. **No** third-graders had ***ever*** seen one before.	5.6 (0.21)
Few	b. **Few** third-graders had ***ever*** seen one before.	5.7 (0.19)
Only	c. **Only** third-graders had ***ever*** seen one before.	5.2 (0.23)
Emotives	d. She was **surprised** that third-graders had ***ever*** seen one before.	5.2 (0.17)
Unlicensed	e. Third-graders had **ever** seen one before.	2.2 (0.26)

Standard deviations shown in parentheses

The full set of items was divided into five lists so that each item appeared once per list in one of its five conditions and so that an equal number of items for each condition appeared in each list. In addition to the experimental items, each of the five lists contained a set of one hundred filler items, which was the same set across the five lists. These fillers mirrored the construction type of the critical experimental items, such that twenty fillers started with *no*, twenty with *few*, twenty with *only*, twenty with emotives, and twenty with a definite determiner or bare plural. The fillers, however, did not contain the critical word *ever*. The purpose of this was to prevent subjects from building an association between a licensor and the NPI, thereby creating a strategy of anticipating the critical word *ever* when encountering one of the four licensors. To keep participants focused during the experimental session, we included comprehension questions for 80 of the trials (about 30% of the total trials). Among the total of 250 sentences in each list, a comprehension question with a yes-or-no answer was created for fifty of the experimental items and thirty of the filler items. Forty questions had correct "yes" answers, and forty had correct "no" answers.

Offline Acceptability

Before the ERP session, we normed our stimuli for acceptability on Amazon.com's Mechanical Turk. Twenty-five participants were recruited online. All were native English speakers, according to a background questionnaire administered before the survey, and the survey itself was restricted to only appear in the U.S. Each participant received one of the five lists described above, except that no filler items were included. Participants were told that each trial would consist of a pair of sentences, and they were instructed to rate the second sentence of each pair based on the first sentence that sets it up. Acceptability was rated on a 1 to 7 scale, with 1 as least acceptable and 7 as most acceptable. The rating results are provided in Table 2. All conditions with a licensor were rated significantly higher than the ones without a licensor ((all $ts>8$; $ps<.0001$). In addition, among licensed conditions, sentences with licensors *no* and *few* were rated higher than sentences with licensor *only* and factives (ts>3; ps<.01).

Participants

Forty-one native English speakers participated in the ERP study. All participants were recruited from either the undergraduate body at the University of Chicago or from the greater Chicago area. All were between the ages of 18 and 35 years old. Participants received either $10 per hour or course credit for their participation. All participants signed a written consent form in accordance with the guidelines of the University of Chicago IRB.

Stimulus Presentation

Participants sat in a quiet, dimly lit room. Stimuli were presented on the presentation monitor in black font, centered on a white background. Before each trial began, the text "Ready…" displayed on the screen, cueing participants to press the center button on the response pad in order to initiate the trial. The trial began with the context sentence centered on the screen. Participants again pressed the center button to initiate presentation of the target sentence. The target sentence began with a crosshair ("+") that displayed for 500ms, followed by a 150ms blank screen. After the blank screen began the word-by-word presentation of the target sentence. Each word of the target sentence displayed for 400ms in the center of the screen, followed by a 150ms interstimulus interval (ISI), except for the sentence-final word (together with the sentence-final period), which displayed for 1000ms. On trials containing a comprehension question, the question appeared directly after the sentence final word. If there was no comprehension question, participants were taken to a screen with the text “Ready…” to indicate the beginning of the next trial. If participants answered a comprehension question incorrectly, they received a message on the screen that said “Oops, wrong answer…” before being taken to the next trial. Before beginning the experiment, participants completed seven practice trials, three of which had a comprehension question. Overall, two-hundred-and-fifty trials were presented to each participant. The full set of trials was divided into eight blocks with roughly thirty trials per block. Participants were allowed to take a short break in between blocks.

EEG Recording

EEG responses were recorded from 32 electrodes (Brain Product, see Figure 1 for montage). Two additional pairs of electrodes were placed above and below the left eye and at the outer canthus of both eyes, in order to monitor vertical and horizontal eye movements, respectively. The EEG signal was referenced to the average of all the electrodes online, and was re-referenced to the average of the two mastoids (TP9, TP10 in Figure 1) offline. The EEG signal was continuously sampled at 1000 Hz, amplified and bandpassed at 40hz, and the impedance was kept below 5 kOhm.

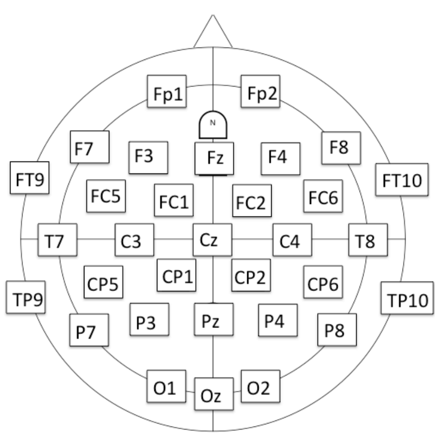

Figure 1: Channel layout.
Midline: Anterior (Fp1, Fp2 , Fz); Frontal (FC1, FC2, Cz); Posterior (CP1, CP2, Pz); Parietal (O1, O2, Oz).
Peripheral: Left Frontal (F7, F3, FC5); Right Frontal (F4, F8, FC6); Left Posterior (CP5, CP1, P3); Right Posterior (CP6, P4, P8).

Data analysis

For the ERP analysis, three subjects were removed from the data, due to excessive ocular and muscular artifacts, and one additional subject was removed due to low comprehension accuracy (<10%). For the remaining thirty-seven subjects, averaged ERPs, cleaned of ocular and muscular artifacts and time-locked to the critical words, were calculated relative to a 200ms pre-stimulus baseline. At the critical word, based on visual inspection, we carried out analyses for three consecutive time windows that showed the most prominent effects: 300-400ms, which encompassed the peak of the N400 effect, and 500-650ms, which encompassed the P600 effect, and a later 700-900ms window that showed a late anterior negativity effect.

For each time window, two initial omnibus ANOVAs were carried out for the mid-regions and peripheral regions separately. For the mid-region ANOVA, the mid-region electrodes were divided into four ROI regions, each containing three electrodes: anterior (Fp1, Fp2, Fz), frontal (Fc1, Fc2, Cz), posterior (Cp1, Cp2, Pz), and parietal (O1, Oz, O2). The within-subject variables were Licensor (five levels, *no*, *few*, *only*, emotive factives, unlicensed), Region (four levels), and electrodes (three levels). For the peripheral region ANOVA, there were also four 3-electrode ROIs: left frontal (F7, F3, Fc5), left posterior (Cp5, Cp1, P3), right frontal (F4, F8, Fc6), and right posterior (Cp6, P4, P8). The within-subject variables were Licensor (5 levels), Hemisphere (2 levels, left or right), Region (2 levels, frontal or posterior), and electrodes (3 levels). We carried out follow-up analyses for smaller regions whenever the omnibus ANOVA revealed an interaction between Licensor and Region or Hemisphere (see below). All analyses were carried out in the statistical package R (R Development Core Team, http://www.R-project.org).

Results

<u>*Behavioral results*</u>

With comprehension questions, all of the subjects had an overall accuracy larger than 88%. The average accuracies for the five conditions were: *no*: 94%; *few*: 94%; *only*: 92%; emotive factives: 92%, and unlicensed: 90%.

<u>*Event related potentials*</u>

N400: 300-400ms

The mid-region omnibus ANOVA showed a main effect of Licensor ($F(4,144)=3.87$, $p<.01$), and also an interaction between Licensor and Region ($F(12, 432)=2.7$, $p<.01$). The peripheral omnibus ANOVA showed a main effect of Licensor ($F(4,144)=4.51$, $p<.01$), and an interaction between Licensor and Region ($F(4,144)=2.46$, $p<.05$), but the interaction between Hemisphere, Licensor, and Region, and the interaction between Hemisphere and Licensor were not significant ($Fs<1.5$, $ps>.3$). Since both analyses showed an interaction between Licensor and Region, we carried out ANOVAs for each ROI in the mid-region and in the peripheral region. We present the results in Table 3.

As shown in Table 3, within the 300-400ms window, the effect of Licensor was significant for most of the ROIs across the whole scalp. Since mid-region and peripheral regions showed largely similar results, for planned pairwise comparisons we combined the mid-anterior, mid-frontal, left-anterior and right anterior ROIs into a 12-electrode anterior-frontal region; similarly, we also combined the mid-posterior, mid-parietal, left and right posterior ROIs into a 12-electrode posterior-parietal region. Within each of these two areas, planned pairwise comparisons were carried out between the unlicensed condition and each of the four licensed conditions. The results are presented in Table 4, and waveforms in Figure 2. We also plot all five conditions together in Figure 3.

Table 3. ANOVA F-values for the effect of Licensor in each ROI in the mid-region and peripheral region.

F(4, 144)	300-400ms	500-650ms	700-900ms
Mid-region			
Frontal	2.4*	--	--
Central	4.8**	--	--
Posterior	3.8**	2.9*	--
Parietal	2.1 ^	3.4*	--
Peripheral Region			
Left-Anterior	4.3**	--	2.7*
Left-Posterior	2.1 ^	2.7*	--
Right-Anterior	3.2*	--	1.98 ^
Right-Posterior	3.0*	2.9*	--

^ $p<.1$; * $p<.05$; ** $p<.01$;

At the anterior-frontal site, the target NPI under the four licensors *no*, *few*, *only*, and emotive factives all showed reduced negativity compared to the unlicensed condition (Table 4). At the posterior-parietal site, *no*, *few*, and emotive factives also showed reduced negativity compared to the unlicensed condition. The difference between *few* and the unlicensed condition is slightly weaker in the posterior-parietal site: it only approached significance in the large 12-electrode area (t(36)=1.7, p<.1), but this difference became significant on a subset of the posterior electrodes (when pooling together P3, P4, Pz, CP1, CP2, CP5, CP6, t(36)=2.1, p<.05). The N400 difference between *only* and the unlicensed condition is also more frontally distributed: it is only marginally significant on a subset of the posterior electrodes (t(36)=1.8, p<.08). Overall, our result suggests that the four grammatical conditions show qualitatively very similar patterns in the N400 time window. We carried out further post-hoc paired comparisons between the four grammatical conditions, and found no differences between any of them—neither in the anterior-frontal site, nor in the posterior-parietal site (Tukey test, all ps>.2, Figure 3).

Table 4: t-values for planned pairwise comparisons between each licensor and the unlicensed condition. Mid-regions and peripheral regions are combined into two large regions: central-frontal and posterior-parietal, with 12 electrodes in each area.

t(36)	300-400ms	500-650ms
Central-frontal		
No vs. Unlicensed	3.1**	--
Few vs. Unlicensed	2.9**	--
Only vs. Unlicensed	2.5*	--
Emotives vs. Unlicensed	2.1*	--
Posterior-parietal		
No vs. Unlicensed	2.0*	2.2*
Few vs. Unlicensed	1.7^	1.8 ^
Only vs. Unlicensed	--	2.7**
Emotives vs. Unlicensed	3.3**	--

^ p<.1; * p<.05; **p<.01;

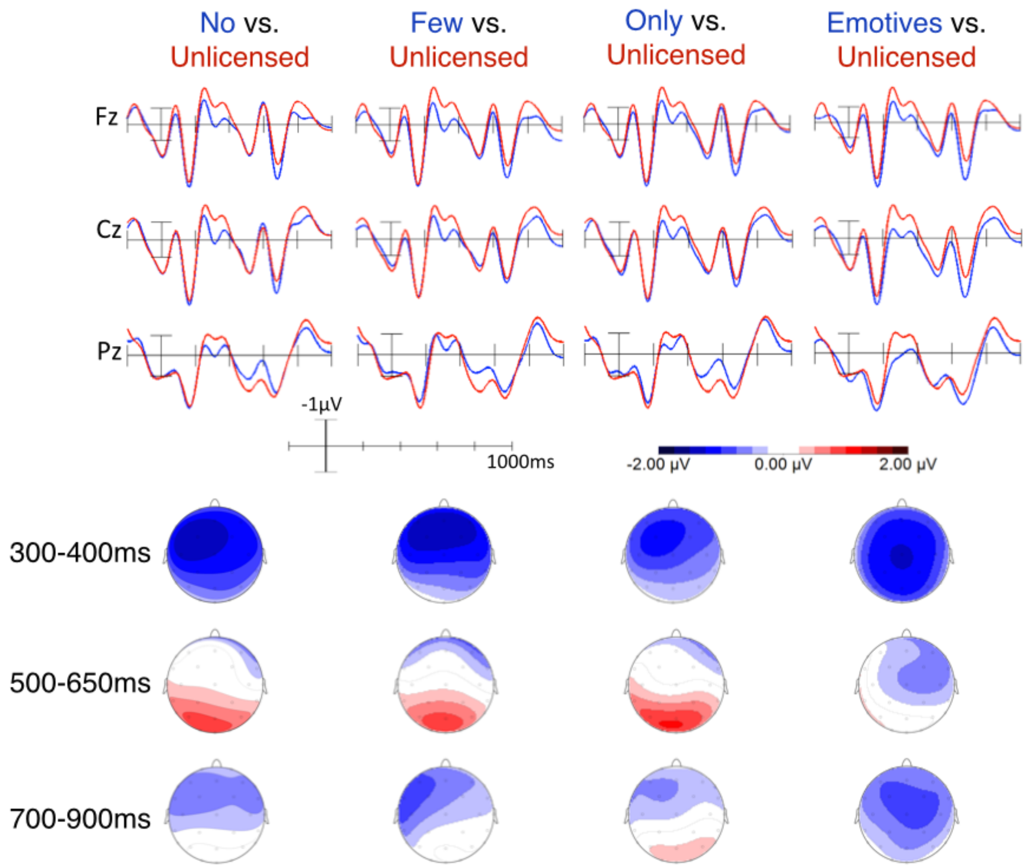

Figure 2: Top: waveforms that compare each of the grammatical licensor with the ungrammatical condition, on the critical word *ever*. The blue line represents a grammatical licensor, and the red line represents the ungrammatical condition. **Bottom**: topographical maps from three time windows. Each plot represents the difference resulting from subtracting a grammatical condition from the ungrammatical condition.

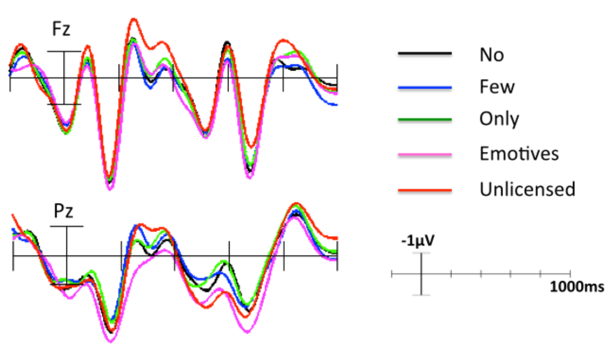

Figure 3: Waveforms from two representative midline channels: frontal Fz and posterior Pz, with all five conditions shown together.

P600: 500-650ms

At the later time window 500-650ms, the mid-region omnibus ANOVA did not show any main effect of Licensor ($F(4,144)=1.3$, $p>.2$), but there was an interaction between Licensor and Region ($F(12, 442)=3.28$, $p<.001$). The peripheral omnibus ANOVA did not reveal any main effect for Licensor ($F(4,144)=0.7$, $p>.5$), but there was also an interaction between Licensor and Region ($F(4, 144)=5.5$, $p<.001$); there were no other interactions involving Licensor ($Fs<1.5$, $ps>.2$).

The ANOVA results for each ROI are presented in Table 3. As shown there, the effect of Licensor in this time window is mostly posteriorly distributed, in posterior and parietal ROIs, but not in frontal or anterior ones. For planned pairwise comparisons, we again divided the scalp into a large anterior-frontal site (12 electrodes) and a large posterior-parietal site (12 electrodes), in the same way we did for the earlier 300-400ms time window. Within each site, we again compared the unlicensed condition with each of the licensed conditions separately. The results are presented in Table 4, and the waveforms are shown in Figure 2. In the anterior-frontal area, none of the comparisons revealed any differences ($ts(36)<1.5$, $ps>.1$). In the posterior-parietal area, the unlicensed condition showed a larger positivity relative to conditions with the licensor *no* and the licensor *only*. The difference between the unlicensed condition and *few* was marginally significance in the large 12-electrode area ($t(36)=1.8$, $p<.1$), and a closer look showed that, on a subset of the parietal electrodes (6 electrodes, P3, P4, Pz, O1, O2, Oz), the condition with *few* did elicit a reliably smaller positivity than the unlicensed condition ($t(36)=2.1$, $p<.05$). We found no difference between the emotive predicates and the unlicensed condition, either in the large 12-electrode posterior-parietal area ($t(36)=0.04$, p>.9) or in the subset of the parietal electrodes (t(36)=0.3, p>.7). We again carried out post-hoc Tukey tests for pair comparisons between all the grammatical licensors. There were no differences between *no*, *few*, and *only* in this time window---neither in the anterior-frontal regions, nor in the posterior-parietal regions (all ps>.8). The emotive condition is not different from the other three licensors in the anterior-frontal region (all ps>.8), but in the posterior-parietal region, it elicited larger P600 than the other three licensors (see Figure 3, emotives vs. *no*, p<.05; emotives vs. *few*, p<.08; emotives vs. *only*, p<.05).

Late anterior negativity: 700-900ms

In the late time window 700-900ms, the mid-region omnibus ANOVA did not show any main effect of Licensor (F(4,144)=1.6, p>.1), nor an interaction between Licensor and Region (F(12, 432)=1.28, p>.2). The peripheral omnibus ANOVA revealed no effect for Licensor (F(4,144)=1.9, p>.1), but there was an interaction between Licensor and Region (F(4, 144)=2.77, p<.05); there were no other interactions involving Licensor (Fs<1.5, ps>.2). ANOVAs within each ROI (Table 3) confirmed that, for mid-regions, the effect of Licensor was not significant for any of the ROIs (Fs<2, ps>.1). For the peripheral region, however, there was an effect of Licensor for the two anterior areas, but not the posterior ones. Since there was no effect of Hemisphere in the omnibus ANOVA, we combined the left and right peripheral anterior ROIs to form a 6-electrode anterior region, and also combined the left and right peripheral posterior ROIs to form a 6-electrode posterior region, and then carried out planned comparisons between the ungrammatical condition and the other four grammatical conditions for both regions. The results are presented in Table 5. As expected, there was no effect in the peripheral posterior region. In the peripheral anterior region, the licensed conditions under licensor *no, few*, and emotives showed a smaller anterior negativity compared to the unlicensed condition. The condition under *only* also showed a trend of a smaller negativity compared to the unlicensed condition, but it didn't reach significance ($t(36)=1.6$, $p=.1$). The post-hoc paired comparisons showed no differences between any of the four grammatical conditions (all ps>.1, Tukey tests), with the only exception that *only* had a slightly larger negativity than emotives ($p<.1$).

Table 5: t-values for planned pairwise comparisons between each licensor and the unlicensed condition, during the late 700-900ms time window, for the peripheral regions.

700-900ms; t(36)	Peripheral Anterior	Peripheral Posterior
No vs. Unlicensed	2.6*	--
Few vs. Unlicensed	2.2*	--
Only vs. Unlicensed	--	--
Emotives vs. Unlicensed	2.4*	--

* p<.05;

Summary and discussion of Experiment 1

In summary, all conditions that contain a legitimate licensor showed a qualitatively similar N400 reduction during the 300-400ms time window at the critical NPI, relative to the unlicensed condition. They also by and large showed a reduced anterior negativity compared to the unlicensed condition during the late 700-900ms time window (modulo the weaker effect on *only*). However, during the P600 time window, although conditions under *no*, *few*, and *only* showed qualitatively similar patterns involving a smaller P600 amplitude relative to the unlicensed condition, the emotive predicate condition yielded a P600 as large as the P600 in the unlicensed condition.

This pattern of results sets implicit negation (e.g., emotive predicates) apart from explicit negation (e.g., *no*, *few*, and *only*) in terms of their time course of processing. Most crucially, while both types of negative conditions are grammatical and are processed similarly in the early N400 time window and the late 700-900ms window, only the emotive condition showed a larger P600 effect that is indistinguishable from the ungrammatical condition. Before we discuss further what these results suggest for the processing of negation, a potential concern is that the emotive predicate condition involves a different structure compared to other conditions—i.e., that this is the only condition in which the NPI *ever* and its licensor are not contained in the same clause. There are different reasons why this may become a concern. First, one may wonder whether the fact that the critical word *ever* in the emotive condition appeared at a later linear position in the sentence, compared to the other conditions, may have affected the results. Word position has been shown to mainly impact the N400 amplitude. For example, Van Petten & Kutas (1990) observed reduced N400 amplitude along the course of a sentence. This effect has mainly been associated with strengthened sentential anticipation of the upcoming word. In the current design, the critical NPI word was not cued by any contextual expectations/constraints, making it less at risk of the word position effect; and in our results, the most salient difference between the emotive condition and other grammatical conditions was not on the N400 amplitude, but the P600. Second, it is also possible that the structural complexity of the emotive condition made it more difficult to process than the other grammatical conditions, leading to enhanced P600 amplitude on the emotive condition. A similar concern was raised in Steinhauer et al. (2010), which argued that any differences one observes between different licensing environments are not necessarily associated with NPI licensing per se, but could also be attributed to the independent differences among different environments. To rule out this possibility, we need to compare these licensing environments when they contain NPIs and when they do not, and examine whether the observed effects are unique to the NPI-present conditions. Finally, one may also ask whether the P600 effect on the emotive condition is due to potential component overlap—the N400 difference between the emotive and the ungrammatical condition is slightly more extensive over the scalp than other conditions (Figure 2), which may have overlapped with the P600 activities in the posterior region, masking any P600 difference between the ungrammatical and the emotive conditions. Component overlap poses methodological issues for standard ERP techniques to detect or eliminate, but if qualitatively similar results could be replicated with other experimental technique/paradigms, it would minimize the possibility that component overlap significantly contributed to the current results. We address all of these concerns in Experiment 2 below.

Experiment 2

Experiment 2 has two primary goals. First, it examines whether the basic pattern of results in Experiment 1 could be replicated in a different behavioral paradigm, i.e., self-paced reading; second, it assesses whether the additional processing cost found on the emotive condition is due to its NPI licensing properties or to the other possible sources of processing complexity mentioned above. We conducted a 2 x 5-design self-paced reading study. The first five conditions are the same as in Experiment 1; the other five conditions were modified from the first five by removing the NPI word "*ever*". An example is shown in Table 6 below.

Table 6: An example item set for Experiment 2, with "/" indicating the separating of regions in the self-paced reading paradigm.

Context:	The teacher brought a tarantula to class.
No	a/f. No/ third-graders/ had/ (*ever*)/ seen/ one/ of/ those/ before.
Few	b/g. Few/ third-graders/ had/ (*ever*)/ seen/ one/ of/ those/ before.
Only	c/h. Only/ third-graders/ had/ (*ever*)/ seen/ one/ of/ those/ before.
Emotives	d/i. She/ was surprised/ that/ third-graders/ had/ (*ever*)/ seen/ one/ of/ those/ before.
Unlicensed (No Negation)	e/j. Third-graders/ had/ (*ever*)/ seen/ one/ of/ those/ before.

For the five **NPI conditions** (i.e., conditions that contain "*ever*"), we make two predictions based on the findings in Experiment 1. First, starting from the critical word "*ever*", the reading time on the unlicensed condition (e) should be longer than the licensed conditions (a-d); this reflects the enhanced processing cost of detecting the ungrammaticality in the unlicensed condition (*e*). Second, the four licensed conditions are expected to initially show similar reading time, and all of them will be read faster than the unlicensed condition; but the emotive condition (d) should also evoke longer reading time at a later time point, diverging from the other three conditions (a-c). If these two predictions are borne out, we consider it a replication of the basic findings in Experiment 1. The five **Plain conditions** (i.e. conditions without "*ever*") then serve as the baseline control conditions to assess whether the predicted effects in the NPI conditions could be explained by reasons other than their NPI licensing properties. In order for the predicted effects in the NPI conditions to be attributed only to NPI related processing, the same effects should be *absent* in the Plain conditions.

Method

Material

Sixty sets of experimental items were created, each with 10 conditions. An example is shown in Table 6. These items were all taken from the material used in Experiment 1; some of them were slightly modified to create a sufficient number of post-critical regions for the purpose of data analysis on spill-over regions. The critical word for the NPI conditions was defined as the NPI word "*ever*". For the Plain conditions, since the NPI word was removed from the stimuli sentence, we defined the critical word as the one that immediately following "*ever*" in the original NPI condition. In the example in Table 6, the critical word for the Plain condition is "*seen*". In this way, the pre-critical regions for the NPI conditions and their corresponding Plain conditions were made completely identical. In addition to the critical items, there were also sixty filler items.

Participants and Procedure

Sixty-four native English speakers (ages ranging from 18 to 30 years old) participated in our study. Participants read through each sentence word by word at their own pace. To keep them focused, for about 60% of the trials, after the last word of each sentence, we asked a simple comprehension question. Participants were instructed to answer this question by pressing one of the two answer keys (Y or N) on the keyboard.

Results

Among the sixty-four participants, four were excluded due to low comprehension accuracy (<80%). Data analysis was performed on the results of the sixty remaining participants. Prior to the data analysis, reading times longer than 2 standard derivations from the mean was excluded. We present in Table 7 and Figure 4 raw reading times for four different regions: one region before the critical word (CW-1), the critical word (CW), and two regions after the critical word (CW+1, CW+2). For data analysis, we first log-transformed the raw reading times. Next, to control for word length and word position effects, we took the whole data set and performed a linear regression with word length and word position as predictors. The residuals from the linear regression, i.e., the residualized log-transformed reading times, are the dependent variable for all the statistical analyses.

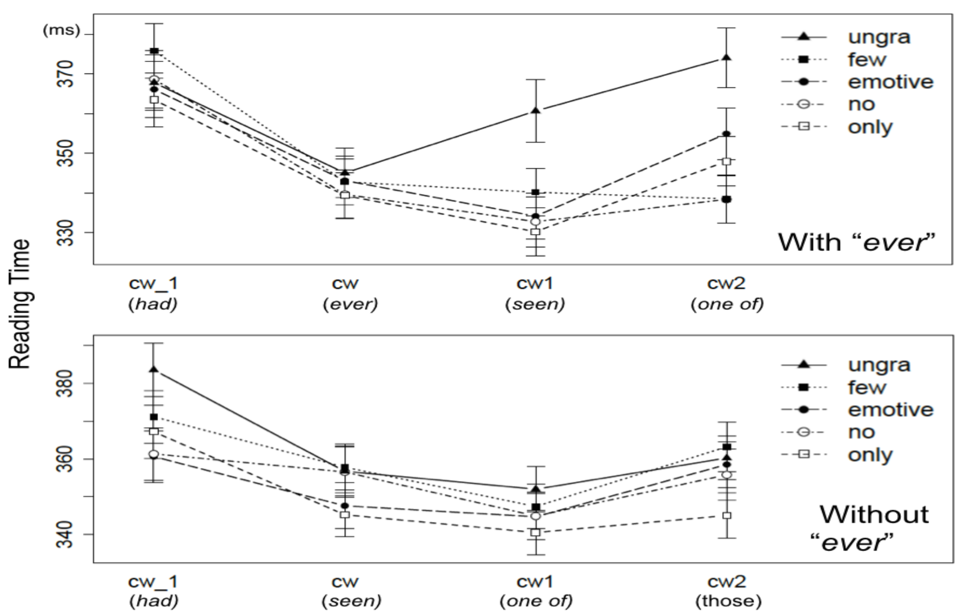

Figure 4: Raw reading times from Experiment 2.

Table 7: Raw reading times(ms) for Experiment 2, with standard errors in parentheses

Raw RTs (ms)	CW-1	CW	CW+1	CW+2
With "ever"	*(had)*	*(ever)*	*(seen)*	*(one of)*
No	369 (13)	340 (11)	333 (12)	338 (12)
Few	376 (12)	343 (12)	340 (12)	339 (12)
Only	364 (13)	339 (11)	330 (13)	348 (12)
Emotive	366 (12)	343 (12)	334 (12)	355 (12)
No negation	368 (13)	345 (12)	361 (15)	374 (14)
Without "ever"	CW-1	CW	CW+1	CW+2
	(had)	*(seen)*	*(one of)*	*(those)*
No	361 (13)	357 (13)	345 (13)	356 (13)
Few	371 (13)	358 (12)	347 (12)	363 (12)
Only	367 (13)	345 (11)	341 (12)	345 (11)
Emotive	361 (13)	348 (12)	345 (12)	359 (12)
No negation	384 (13)	357 (13)	352 (12)	360 (11)

Linear mixed-effects models were conducted for data analysis. Our primary interest is examining whether the following two sets of comparisons show similar or different results for the NPI and the Plain conditions: (i) the difference between the No-Negation condition (i.e., the unlicensed condition if an NPI is present) and the other four conditions with a licensor; and (ii) the difference between the emotive predicate condition and the other three types of licensors (i.e., *no, few*, and *only*). In order to directly assess the effects of these two comparisons, we set them up as two orthogonal contrasts with Helmert coding (Venables and Ripley, 1999; Vasishth & Broe, 2011), and these two contrasts were then included in the mixed-effects models as two fixed effect predictors. Separate models were conducted for the NPI and the Plain conditions. For the NPI conditions, the first contrast examines ***the effect of Negation***, in which the ungrammatical condition is contrasted with the four grammatical conditions (i.e., *e.* vs. *a*, *b*, *c*, *d*); the second contrast examines ***the effect of the Emotive Predicate***, in which the emotive predicate condition is contrasted with the other three grammatical conditions (i.e., *d.* vs. *a*, *b*, *c*). For the Plain conditions, the same two sets of contrasts were set up, i.e., the effect of negation: condition *j* vs. *f*, *g*, *h*, *i*; and the effect of the emotive predicate: condition *i* vs. *f*, *g*, *h*. We also included the reading time from the two immediately preceding regions (Spill-over region 1 and Spill-over region 2) as two fixed effect predictors in the mixed-effects models, given that self-paced reading time at any given region is usually influenced by the reading time at previous regions (i.e., the spill-over effect). For the random effect structure, we included random intercepts for both subjects and items, and maximal random slopes of the two user-defined contrasts above were also included whenever the model could converge. Statistical analyses were carried out using the lmerTest package in R.

On any given region, effects from the two previous regions were always highly significant (all ps<.0001) in all of our models. Since these effects are not our primary interest, we won't discuss them further. In Table 8, we present results for the two crucial contrasts on the CW region and the two following regions (CW+1 and CW+2)[9]. With the five **NPI conditions**, no effect was found on the CW "*ever*". The effect of grammaticality (i.e., Contrast 1 in Table 8, longer RTs on the No-Negation condition) did not emerge until the region after the CW, and this effect also continued into the CW+2 region. Since it is not uncommon for the self-paced reading paradigm to delay the effect on the critical word to spill-over regions, we consider the effect on the CW+1 region (and the region after) to be largely in line with our prediction that the participants were sensitive to the

---

[9] We also conducted a more traditional mixed-effects model on the whole data set, with NPI (2 levels: NPI or no-NPI) and Licensor (5 levels) as the fixed effect predictors (again controlling for spill-over effects). No significant effect was found on the CW. On the CW+1 region, there is a *NPI x Licensor* interaction when all ten conditions were considered (p<.05), but no interaction when only the eight conditions with licensors (i.e. excluding conditions e&j) were considered (p>.9). Crucially, the latter interaction was significant on the CW+2 region (p=.01). This is completely in line with the results reported in Table 8.

grammaticality contrast between the unlicensed condition and the other four licensed conditions. Furthermore, our results also showed that the emotive predicate condition triggered longer RTs than other grammatical conditions, but not until the CW+2 region (Contrast 2 in Table 8). The relative timing difference between Contrast 1 and Contrast 2 in Table 8 is of most interest for our purpose—at the early stages of processing, all four licensed conditions pattern similarly, and all of them differ from the unlicensed condition; but at a later point, the emotive condition also revealed more processing cost than the other three licensed conditions. This pattern of results replicates the basic findings from Experiment 1, even though the ERP results in Experiment 1 are more fine-grained, given the more precise time course information and the more specific inferences provided by different ERP components (see the General Discussion section for discussion). Equally important for current purposes is that no effect was found for the five **Plain conditions** on any of the regions we examined, eliminating the possibility that the effects observed on the NPI conditions were associated with position of the NPI or structural complexity of the sentences containing it.

Table 8: Results on the two crucial contrasts for Experiment 2

	Contrast 1: No-negation vs. (no, few, only, emotives)				Contrast 2: Emotives vs. (no, few, only)			
With "ever"	Estimate	Std. Error	t-value	p-value	Estimate	Std. Error	t-value	p-value
CW	0.003	0.01	0.28	0.78	-0.003	0.01	-0.29	0.78
CW+1	-0.05	0.01	-3.39	**0.001****	-0.01	0.01	-0.54	0.59
CW+2	-0.04	0.02	-2.43	**0.02***	-0.02	0.01	-2.37	**0.02***
Without "ever"								
CW	0.01	0.01	1.01	0.31	0.01	0.01	0.92	0.36
CW+1	-0.01	0.01	-0.71	0.48	-0.004	0.01	-0.38	0.71
CW+2	-0.02	0.01	-1.49	0.14	-0.01	0.01	-1.53	0.13

Summary of Experiment 2

Granted that information obtained from the self-paced reading paradigm is less fine-grained than ERP results, the findings in Experiment 2 are nevertheless informative in two important aspects. First, they show that the basic findings in Experiment 1 are replicable in a different experimental paradigm. This point is important in ruling out one account of the increased P600 seen with the emotive factives in Experiment 1. If the larger P600 were an artifact due to component overlap between the N400 and the P600 time windows at the posterior scalp site, one should not see similar processing cost replicated in a self-paced reading paradigm. Second, the self-paced reading results rule out the possibility that the observed effects among the NPI conditions should be attributed to independent structural or contextual differences among different conditions. If structural complexity or position of the NPI were relevant to the observed differences between conditions, the same difference should have been observed regardless of the presence of an NPI.

Admittedly, plenty of caution needs to be taken in drawing parallel relations between ERP and self-paced reading results. The electrophysiological response and the behavioral reading time response are generated, at least partially, by different mechanisms. But the fact that the same costs, with the same relative timing, are observed in the NPI conditions from both the ERP data and the reading time data suggests that they are comparable measures to examine the online processing of NPI comprehension. On this reasoning, the five Plain conditions from the behavioral task give clues about the interpretation of the ERP findings.

General Discussion

In this study, we capitalize on the negation-sensitive property of NPIs to assess the time course of comprehending different kinds of negative expressions. In particular, we looked at how the negative information from *no*, *few*, *only*, and emotive predicates is extracted. We showed, first, that during the N400 time window, all these licensors helped to trigger a reduced N400 on the NPI word *ever*, compared to the unlicensed condition, suggesting that the negative meaning of *no*, *few*, *only*, and emotive predicates is successfully processed by the time the word *ever* is encountered. Second, during the P600 time window, we observed a difference between *asserted* negation (i.e., an entailment with *no*, *few*, and *only*) and negation that is part of the non-asserted content (e.g., a presupposition of emotive predicates). On the word *ever,* the unlicensed condition elicited a larger positivity only relative to conditions containing asserted negation (e.g., *no*, *few*, and *only*), whereas the emotive factives did not show any positivity difference compared to the unlicensed condition. And finally, at the late 700-900ms time window, the four licensors demonstrated qualitatively similar patterns, showing a smaller anterior negativity relative to the unlicensed condition. Results from the self-paced reading paradigm in Experiment 2 largely replicated the general pattern of results from Experiment 1. Crucially, the additional control conditions in Experiment 2 provided further evidence that the observed effects on the NPI conditions were indeed due to the NPI licensing properties (i.e., negation) of different conditions.

The findings presented in this study have implications both for how negation is accessed and processed, and for how polarity items are licensed in online processing. We discuss both of these issues below. Since the ERP results provided more fine-grained information to bear upon these questions, and given that the self-paced reading results are completely in line with the ERP results, our discussion below will largely focus on the ERP findings.

Negative meaning is computed immediately with no delay

The first conclusion that we draw from the current results is that the meaning of negation, regardless of how negation is expressed, is computed very early. The finding regarding the N400 effects in Experiment 1 provided strong evidence for this. As mentioned in the introduction, we assume that N400 amplitude indexes the degree of semantic feature match between the upcoming word and the previous context. Reduced N400 on an upcoming word generally indicates that some relevant semantic feature of the word has already been activated prior to the appearance of the word. In many discourse situations, such "pre-activation" of a semantic feature is due to active expectation driven by a discourse model in which a comprehender's understanding of the discourse context, combined with their stored real-world knowledge, encourages top-down predictions about what the upcoming word/event should be. In the current study, it is unlikely that the discourse context encouraged the anticipation of the specific target word "ever" itself. However, a discourse context containing *no*, *few*, *only*, or emotive predicates (i.e., the four NPI licensors) will give rise to a negative meaning, explicitly or implicitly, therefore activating a [+Negation] feature prior to the appearance of the NPI word *ever*. This pre-activation leads to the facilitated processing of the target word.

These four licensors, as we mentioned earlier, encode negative information at different levels of representation, ranging from morphology to pragmatics (Table 1).

Since they all trigger a reduced N400 on the target NPI word, this suggests that the computation of negation is NOT delayed until a complete affirmative proposition has been comprehended, regardless of how the negative information is presented. In particular, our results suggest that rapid incremental computation is carried out, not only for the obvious instances of negation like *no*, but also for instances of negation without overt negative morphology (e.g., *few* and *only*), and for those instances in which the negation itself is not even part of the asserted meaning (e.g., emotive predicates).

Contrary to previous studies that have concluded that negation is not immediately accessed online (e.g., Fischler et al., 1983; Kounios and Holcomb, 1992), our results suggest that the delayed effect observed in previous studies arose not because negation, per se, as a grammatical or semantic relation, is particularly difficult to compute. This is in line with conclusions from Nieuwland and Kuperberg (2008) and Tian, Breheny, and Ferguson (2010). As discussed above, previous studies often conflate the question of how quickly negation can be computed with the question of how quickly information from different sources, including world knowledge, interact to facilitate the processing of the upcoming word. By making NPIs the target word, the current study manages to assess the first issue without interference from the second, and we show fast computation and immediate sensitivity to negation. To better understand the delayed effects observed in previous ERP studies on negation, we suggest that future research needs to address the question of how quickly the semantic association based on real-world conceptual knowledge can be suppressed or inhibited when this information is in conflict with what discourse information suggests. As mentioned earlier, information from multiple levels of representation can affect the processing of an upcoming word, and therefore modulate the amplitude of the N400. This includes both semantic associations based on real world knowledge that are stored in long-term memory, *and* structured discourse information that is incrementally accrued while a situation model based on context is constructed. In some situations, these two streams of information align with each other, and they modulate N400 amplitude in the same way. When negation is present, however, the stored knowledge about relations between entities and events may conflict with the compositionally derived meaning of a sentence or a discourse. For instance, in the sentence "*A robin is not a bird*", the long-term memory association "robin = bird" and the sentence meaning "robin ≠ bird" push N400 amplitudes in different directions. Even though negation in this sentence might be processed in a timely fashion, the effect of negation on the N400 at the word *bird* (e.g., arising from the falsity of the sentence) may only surface when a facilitating effect from real world associations is sufficiently suppressed or inhibited (e.g., the lexical association between *robin* and *bird*).

Two distinct mechanisms for NPI licensing

Although different types of negation can be rapidly processed by the early N400 time window, they appear to be prioritized in different ways for the grammatical purpose of licensing NPIs. In particular, during the P600 time window, we observed that participants relied only on asserted negation (i.e., negation from *no*, *few*, and *only*) to license NPIs; whereas non-asserted negation (i.e., emotive predicates) was not treated the same at this stage. As a result, emotive predicates appeared to have similar P600 profiles to the unlicensed condition.

We think the contrast between the N400 and P600 time windows underlines a deeper distinction in the functional interpretation of the N400 and P600 components. The N400 amplitude simply indexes the degree of overlap between the semantic features that are activated prior to the target word and the lexical semantic features that the target word contains. Therefore, its modulation reflects the cost of accessing and retrieving the lexical representation of the upcoming word: the more relevant the features that have been pre-activated in the previous context, the easier it is to process the upcoming word, and hence the more reduction of the N400 amplitude. The N400 itself does *NOT* index the process of integrating different sources of information into a linguistic representation, whose truth, as well as real-world plausibility and coherence, can be evaluated. The P600, on the other hand, *does* broadly index an integration stage in which the linguistic representation of the current proposition is updated with the addition of the upcoming word. In the context of NPI licensing, a number of authors have suggested that the P600 time window supports the logical-semantic processes required for NPI licensing, or the mapping between a combinatorial syntactic process and a compositional semantic process (e.g., Steinhauer et al., 2010).

Our findings in the P600 time window support the existence of two distinct mechanisms for NPI licensing (e.g., Giannakidou (1998; 2006)'s 'dual mode' of licensing), and these two different mechanisms are applied in online processing with a different time course. Under one mechanism, an NPI is licensed by negation in the syntactic-semantic representation. This is the licensing mechanism for licensors like *no*, *few*, and *only*, all of which give rise to negation via syntactic or semantic means (or both). Under the other mechanism, an NPI is licensed if a negative inference can be generated pragmatically.[10] This licensing mechanism is relevant for emotive licensors. The syntactic-semantic licensing mechanism is the first licensing mechanism that the comprehension system consults during the P600 time window, and environments that do not contain negation in the syntactic-semantic representation, including both the ungrammatical condition and the emotive predicate condition, showed an enhanced P600, indicating integration difficulty. The failure of the emotive predicates to license NPIs in the P600 time window, coupled with the fact that the negative inferences from emotive predicates have been successfully computed by the N400 time window, strongly suggests that the pragmatic licensing mechanism is not immediately applied online for NPI licensing by the P600 time window even though a pragmatically inferred negation is available.

The actual licensing effect of the pragmatic licensing mechanism, in the context of emotive predicates, is delayed until after the P600 time window, during the 700-900ms window. In this time window, on the critical NPI, we observed an increased anterior negativity associated with the unlicensed condition, relative to all other licensed conditions. This suggests that, by this time window, the emotive predicate was finally recognized by the comprehension system as an appropriate licensor, on a par with other grammatical licensors. This increased anterior negativity on the ungrammatical condition was slightly more pronounced on the left anterior site (Table 3). An increased late anterior negativity, with a left dominance, was also observed for unlicensed NPIs in Steinhauer et al. (2010), who called it the L-LAN effect. Following Steinhauer et al.

[10] It must also be noted that not all negative inferences could license NPIs, see discussions in (Linebarger 1987, Horn 1989, 2002, Giannakidou 2006). A negative presupposition, as is the case with emotives, is part of their conventional meaning, hence it will have strong licensing potential. A mere contextual inference, on the other hand, will have less licensing potential (see Giannakidou 2006 for more details).

(2010), we suggest that the increased late anterior negativity on the unlicensed condition was driven by the working memory cost associated with maintaining different levels of representation (Kluender and Kutas, 1993; Hagoort and Brown, 1994; Kaan and Swaab, 2003; King and Kutas, 1995), when the comprehension system eventually failed to integrate them into a coherent representation that can license NPIs; although we note that the memory related LAN effect in the literature tends to be more prolonged than what we observed here (or what was found in Steinhauer et al., 2010). On the other hand, the four legitimate licensors elicited relatively smaller working memory costs, since the overall integration is successful.

Conclusions

This paper investigates the time course of processing negation by examining how the negative polarity item *ever* is processed in different types of negative environments. Our results show that negative information from different sources, e.g., whether it is part of the asserted or non-asserted content, is accessed equally rapidly in online processing. At the same time, however, we observe that asserted negation, i.e., negation in the syntactic-semantic representation, is applied immediately to grammatically license NPIs; pragmatically inferred negation, on the other hand, is only adopted at a later processing stage as a secondary NPI-licensing mechanism.

Acknowledgements

References:

Atlas, J. D. (1993) The importance of being only: Testing the neo-gricean versus neo-entailment paradigms. *Journal of Semantics* 10, 301–318.

Altmann, G. T., & Steedman, M. (1988) Interaction with context during human sentence processing. *Cognition*, 30, 191–238.

Baker, C. L. (1970) Double Negatives, *Linguistic Inquiry* 1, 169-186.

Beaver, D. and Clark, D. (2008) *Sense and Sensitivity: How Focus Determines Meaning.* Blackwell, Oxford.

Beltrama, A. and M. Xiang. (2012) Is good better than excellent? An experimental investigation on scalar implicatures and gradable adjectives. In E. Chemla, V. Homer & G. Winterstein (eds) *Proceedings of Sinn und Bedeutung* 17, 81-98.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on “semantic P600” effects in language comprehension. Brain Research Reviews, 59, 55-73.

Bott, Lewis, & Noveck, Ira A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. Journal of Memory and Language, 51, 437-457.

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about Semantic Illusions: Rethinking the functional role of the P600 in language comprehension. Brain Research, 1446, 127-143

Breheny, Richard, Katsos, Napoleon, & Williams, John. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. Cognition, 100, 434-463. doi: 10.1016/j.cognition.2005.07.003

Chierchia, C. (2006). Broaden your views. Implicatures of domain widening and the ‘logicality’ of language. *Linguistic Inquiry*, 37, 535–590.

Chow, W. Y. and Phillips, C. (2013). No semantic illusion in the “Semantic P600” phenomenon: ERP evidence from Mandarin Chinese. *Brain Research*, 1506, 76-93.

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

Clark, H. H. (1976). *Semantics and comprehension*. The Hague: Mouton.

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

Doran, R.; Ward, G., Larson, M., McNabb, Y. and Baker, R. (2012) A novel experimental paradigm for distinguishing between WHAT IS SAID and WHAT IS IMPLICATED. *Language* 88, 1, 124-154.

Drenhaus, H., Saddy, D., & Frisch, S. (2005). Processing negative polarity items: When negation comes through the backdoor. In S. Kepser & M. Reis (Eds.), *Linguistic evidence—empirical, theoretical, and computational perspectives* (pp. 145–165). Berlin/New York: Mouton de Gruyter.

Drenhaus, H., beim Graben, P., Saddy, D., & Frisch, S. (2006). Diagnosis and repair of negative polarity constructions in the light of symbolic resonance analysis. Brain and Language,96, 255–268.

Drenhaus, H., Blaszczak, J., & Schutte, J. (2007). Some psycholinguistic comments on NPI licensing. In: E. Puig-Waldmuller (Ed.), Proceedings of Sinn und Bedeutung,11 (pp. 180–193). Barcelona, Spain: Universitat Pompeu Fabra

Federmeier, K.D., Kutas, M. (1999). A rose by any other name: long-term memory structure and sentence processing. Journal of Memory and Language, 41, 469-495.

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

Fodor, J. D., Fodor, J. A. and Garrett, M. F. (1975) The psychological unreality of semantic representations, Linguistic Inquiry, 6, 515-531.

Franklin, A., Giannakidou, A., & Goldin-Meadow, S. (2011) Negation and structure building in a home sign system. *Cognition*, *118(3)*, 398-416.

Friederici, A. D. and Weissenborn, J. (2007) Mapping sentence form onto meaning: The syntax-semantic interface". *Brain Research* 1146: 50–8.

Geurts, B. (2010). *Quantity Implicatures*. Cambridge University Press.

Giannakidou, A. (1998). *Polarity sensitivity as (non)veridical dependency*. Amsterdam, Philadelphia: John Benjamins.

Giannakidou, A. (2006). Only, emotive factives, and the dual nature of polarity dependency. *Language*, 82, 575–603.

Giannakidou, A. (2011). Negative and Positive Polarity Items. In von Heusinger, Klaus, Claudia Maienborn and Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning* de Gruyter, 1660–1712

Giannakidou (to appear) The subjunctive as *evaluation*— Nonveridicality, epistemic subjunctive, and emotive-as-expressive. In: Blaszack, J., Klimek-Jankowska, D., Mygdalski, K., and Giannakidou, A. (eds) *Tense, Mood, and Modality: New Perspectives on Old Questions.* University of Chicago Press.

Gough, P. B. (1965). Grammatical transformations and speed of understanding. *Journal of Verbal Learning & Verbal Behavior, 4*, 107-111.

Gouvêa, A.C., Phillips, C., Kazanina, N. & Poeppel, D. (2010). The linguistic processes underlying the P600. Language and Cognitive Processes, 25 (2), 149-188.

Greene, J. M. (1970). The semantic function of negatives and passives. *British Journal of Psychology, 6*, 17-22.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. Morgan (Eds.), Syntax and semantics 3: Speech acts, pp. 41–58. New York: Academic Press. Reprinted in and cited from Grice (1989:22-40).

Grodner, Daniel J., Klein, Natalie M., Carbary, Kathleen M., & Tanenhaus, Michael K. (2010). "Some," and possible all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. Cognition, 116, 42-55.

Hagoort, P., Brown, C.M. & Groothusen, J. (1993) The syntactic positive shift as an ERP measure of syntactic processing. Language and Cognitive Processes, 8, 439-483.

Hagoort, P., & Brown, C. M. (1994). Brain responses to lexical ambiguity resolution and parsing. In C. Clifton, Jr., L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 45-80). Hillsdale,NJ: Erlbaum.

Hoeksema, J. (1994). On the grammaticalization of negative polarity items. In: S. Gahl, A. dolbey & C. Johnson (eds.). *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* (BLS) 20. Berkeley, CA: Berkeley Linguistics Society, 273–282.

Hoeksema, J. (2012). On the Natural History of Negative Polarity Items. *Linguistic Analysis* 38, 1/2, 3-33.

Hoeksema, J., and den Ouden, D.B. (2005). Positief- en negatief-polaire bepalingen van graad: Een empirisch onderzoek. Tabu: Bulletin voor taalwetenschap, 34(3/4), 129–144.

Horn, L. R. (1996). Exclusive company: Only and the dynamics of vertical inference. *Journal of Semantics* 13, 1–40.

Horn, L. (2001) *A Natural History of Negation*. Chicago: University of Chicago Press.

Horn, Laurence R. (2002) Assertoric inertia and NPI licensing. Proceedings of the Annual Meeting of the Chicago Linguistic Society. 38. University of Chicago.

Klima, E. S. (1964). Negation in English. In J. A. Fodor & J. J. Katz (Eds.), *The structure of language*. Englewood Cliffs, NJ: Prentice-Hall.

Huang, Yi Ting, & Snedeker, Jesse. (2009). Online interpretation of scalar quantifiers: Insight into the semantics--pragmatics interface. Cognitive Psychology, 58, 376-415.

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

Kaan, E., Harris, A., Gibson, E. & Holcomb, P.J. (2000) The P600 as an index of syntactic integration difficulty. Language and Cognitive Processes, 15, 159-201.

Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analy- sis: An electrophysiological differentiation. Journal of Cognitive Neuroscience, 15, 98–110.

Kartunnen, L. 1971. "Some observations on factivity." *Papers in Linguistics* 4:55-69.

Kim, A. & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52(2)*, 205-225.

King, J. W., & Kutas, M. (1995). Who did what when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience, 7*, 376- 395.

Kiparsky, P. and C. Kiparsky. 1968. "Fact." In M. Bierwisch and K. E. Hei- dolph (eds.), *Progress in Linguistics,* pp. 143-173. The Hague: Mouton.

Kadmon, N., & Landman, F. (1993). Any. *Linguistics and Philosophy*, 16, 353–422.

Kiparksy, P. and C. Kiparsky, 1970. Fact. In M. Bierwich and K. Heilolph (eds), *Progress in Linguistics*, 143-173. Mouton.

Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from ERP's on the processing of unbounded dependencies. Journal of Cognitive Neuroscience, 5(2), 196–214.

Kounios, J., & Holcomb, P. J. (1992). Structure and process in semantic memory: Evidence from event-related brain potentials and reaction times. *Journal of Experimental Psychology: General, 121*, 459-479.

Krifka, M. (1995). The semantics and pragmatics of polarity items in assertion. *Linguistic Analysis* 15, 209–257.

Kuperberg GR. (2013) The pro-active comprehender: What event-related potentials tell us about the dynamics of reading comprehension. In: *Unraveling the Behavioral, Neurobiological, and Genetic Components of Reading* Comprehension (pp. 176-192). Miller, B., Cutting, L., & McCardle, P (Eds): Baltimore: Paul Brookes Publishing.

Kuperberg, G.R., 2007. Neural mechanisms of language comprehension: challenges to syntax. Brain Res. 1146, 23–49.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*, 161-163.

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.

Ladusaw, W. A. (1979). *Negative polarity items as inherent scope relations*. Ph.D. Dissertation, University of Texas at Austin.

Laka, I. (1994). *On the Syntax of Negation*. Outstanding Dissertations in Linguistics Series, New York-London: Garland Publishing Co.

Lau EF, Phillips C, Poeppel D. (2008). A cortical network for semantics: [de]constructing the N400. Nat. Rev. Neurosci. 9:920–33.

Linebarger, M. (1987). Negative polarity and grammatical representation. *Linguistics and Philosophy*, 10, 325–387.

MacDonald, M. C., Pearlmutter, N. J. and Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676-703.

Münte, T. H., Heinze, H.J. and Mangun, G. R. (1993). Dissociation of brain activity related to syntactic and semantic aspects of language. Journal of Cognitive Neuroscience, 5, 335-344.

Munte,T.F., and Heinze,H.J. (1994). ERP negativities during syntactic processing of written words. In H.J.Heinze, T.F.Munte,& G.R.Mangun (Eds.), Cognitive electrophysiology. Boston, MA: Birkhauser.

Neville, H., Nicol, J. L., Barss, A., Forster, K.I. and Garrett, M. F. (1991) Syntactically based sentence processing classes: Evidence from event-related brain potentials. Journal of Cognitive Neuroscience, 3, 151-165.

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, 1213-1218.

Nieuwland, Mante S., Ditman, Tali, & Kuperberg, Gina R. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. Journal of Memory and Langauge, 63, 324-346.

Osterhout, L., & Holcomb, P.J. (1992). Event-related potentials elicited by syntactic anomaly. Journal of Memory and Language, 31, 785-806.

Osterhout, L. Holcomb, P. J. & Swinney, D.A. (1994) Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. Journal of Experimental Psychology: Learning, Memory and Cognition, 20, 786-803.

Paczynski, M. and Kuperberg, G.R. (2012) Multiple influences of semantic memory on sentence processing: distinct effects of semantic relatedness on violations of real-world event/state knowledge and animacy selection restrictions. *Journal of Memory and Language*, 67, 4, 426-448.

Phillips, C., Kazanina, N., & Abada, S. (2005). ERP effects of the processing of syntactic long-distance dependencies. Cognitive Brain Research, 22, 407-428.

Politzer-Ahles S, Fiorentino R, Jiang X, Zhou X (2013) Distinct neural correlates for pragmatic and semantic meaning processing: an event-related potential investigation of scalar implicature processing using picture-sentence verification. Brain Res 1490: 134–152.

Postal, P. (2002) Skeptical linguistic essays. New York: Oxford University Press.

Potts, C. (2005) *The Logic of Conventional Implicatures*. Oxford University Press.

R Development Core Team (2012). R: A language and environment for statistical computing. http://www.R-project.org.

Saddy, D., Drenhaus, H., & Frisch, S. (2004). Processing polarity items: Contrasting licensing costs. Brain and Language, 90, 495–502.

Shao, J., & Neville, H. (1998). Analyzing semantic processing using event-related potentials. Newsletter for the Center for Research in Language, 11(5), 3–20.

Stalnaker, 1978, Assertion', in P. Cole (ed), *Syntax and Semantics 9, New York: New York Academic Press*, 315–32.

Steinhauer, S.; Drury, J.; Portner, P.; Walenski, M. and Ullman, M. (2010). Syntax, concepts, and logic in the temporal dynamics of language comprehension: Evidence from event-related potentials. *Neuropsychologia*, 48: 1525–1542.

Tesan, G., Johnson, B. W., & Crain, S. (2012). How the brain responds to any: An MEG study. Brain and Language, 120, 66–72.

Tian, Y., Breheny, R., Ferguson, H. J. (2010). Why we simulate negated information: A dynamic pragmatic account. *Quarterly Journal of Experimental Psychology* 63(12), 2305-2312.

Tonhauser, J., D. Beaver, C. Roberts, and M. Simons (2013), Towards a taxonomy of projective content, *Language* 89(1): 66-109.

Trabasso, T., Rollins, H., & Shaughnessy, E. (1971). Storage and verification stages in processing concepts. Cognitive Psychology, 2, 239–289.

Urbach, T. and Kutas, M. (2010) Quantifiers more or less quantify on-line: ERP evidence for partial incremental interpretation. Journal of Memory and Language, 63, 158-179.

van Berkum, J.J.A. (2009). The neuropragmatics of "simple" utterance comprehension: An ERP review. In U. Sauerland & K. Yatsushiro (Eds.), *Semantics and* *pragmatics: From experiment to theory* (pp. 276–316). Basingstoke: Palgrave Macmillan.

van de Meerendonk, N., Kolk, H.H.J., Chwilla, D.J., Vissers,C.T.W.M., 2009. Monitoring in language perception. Lang. Ling.Compass 3, 1211–1224

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

Van Tiel, B., van Miltenburg, E., Zevakhina, N. and Geurts, B. (2013) Scalar diversity. Talk presented at the 5th Experimental Pragmatic conference, Utrecht.

Vasishth, S., and Broe, M. (2011). The Foundations of Statistics: A Simulation-based Approach. Berlin: Springer.

Venables, W. N., and Ripley, B. D. (1999). Modern Applied Statistics with S-PLUS. NewYork, NY: Springer. doi: 10.1007/978-1-4757-3121-7

von Fintel, K. (1999). NPI-licensing, strawson-entailment, and context-dependency. Journal of Semantics, 16, 97–148.

Wason, P.C (1959).The processing of positive and negative information. Quarterly Journal of Experimental Psychology, 11, 92-107.

Wason, P.C. (1961). Response to affirmative and negative binary statements. British Journal of Psychology, 52, 133-142.

Wason, P. C. (1965). The contexts of plausible denial. Journal of Verbal Learning and Verbal Behavior, 4, 7–11.

Xiang, L., Dillon, B., & Phillips, C. (2009). Illusory licensing effects across dependency types: ERP evidence. Brain and Language, 108, 40–55.

Yanilmaz, A., and Drury, J. (2013). “Licensing in Turkish: an ERP study,” in Poster Presented at the 20th Cognitive Neuroscience Society Meeting (San Francisco, CA).

Yurchenko. A., den Ouden, D.B., Hoeksema, J., Dragoy, O., Hoeks, J., Stowe, L., et al., (2013) Processing polarity: ERP evidence for differences between positive and negative polarity, *Neuropsychologia* 5, 132–141.

Zwarts, F. (1986). CategorialeGrammatica en Algebraische Semantiek. Een Studie Naar Negatie en Polariteit in het Nederlands. Ph D thesis, University of Groningen.

Zwarts, F. (1995). Nonveridical con- texts. Linguist. Anal. 25, 286–312.