Journal of Neurolinguistics 38 (2016) 71–88

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

Research paper

Semantic and pragmatic processes in the comprehension of negation: An event related potential study of negative polarity sensitivity

Ming Xiang*, Julian Grove, Anastasia Giannakidou

Linguistics Department, University of Chicago, United States

ARTICLE INFO

Article history:
Received 7 July 2015
Received in revised form 1 November 2015
Accepted 9 November 2015
Available online xxx

Keywords:
Negation
Semantics and pragmatics
Negative polarity
Discourse processing
Assertion and non-assertion
ERP

ABSTRACT

Most previous studies on negation have generally only focused on sentential negation (*not*), but the time course of processing negative meaning from different sources remains poorly understood. 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 semantic assertion (explicit negation), the last gives rise to a pragmatic negative inference via non-asserted content (implicit negation). Under all these environments, an NPI elicited a smaller N400 compared to an unlicensed NPI, suggesting that negation, regardless of its source, is rapidly computed online. However, we also observed that explicit negative meaning (i.e., semantic, as contributed in the assertion) and implicit negative meaning (contributed by pragmatic inferences) were integrated into the grammatical representation in different ways, leading to a difference in the P600, and calling for a separation of semantic and pragmatic integration during sentence processing (and NPI licensing). The qualitative differences between these conditions were also replicated in a self-paced reading study (Experiment 2).

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Polar opposition—negation vs. affirmation—is one of the distinctive properties of human language (Horn, 2001): every natural language includes at least one device that expresses the negation of an affirmative constituent. 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*. 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. 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*

* Corresponding author. Language Processing Lab, Linguistics Department, University of Chicago, Chicago, IL 60637, United States.
E-mail address: mxiang@uchicago.edu (M. Xiang).

http://dx.doi.org/10.1016/j.jneuroling.2015.11.001
0911-6044/© 2015 Elsevier Ltd. All rights reserved.

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.

Even though negation is extremely common in everyday communication, the comprehension process of negation remains poorly understood. Most previous investigations have focused almost exclusively on the negative form *not* (Clark & Chase, 1972; Fischler, Bloom, Childers, Roucos, & Perry, 1983; Kounios & Holcomb, 1992; Nieuwland & Kuperberg, 2008; Tian, Breheny, & Ferguson, 2010; Wason, 1959, 1961), and we know relatively little about how other types of negation are processed. In the current study, we make use of the negation-sensitive property of a closed-class of 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.

1.1. 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 of an expression 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. 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 elaborate on below.

Any given utterance conveys an array of meanings. In the widely used Gricean and neo-Gricean frameworks (Geurts, 2010; Grice, 1975; Horn, 2001; Stalnaker, 1978), assertion conveys the logical meaning of a sentence, i.e., the truth conditions and *entailments* of a sentence; and non-asserted meaning is thought of as *pragmatic* meaning, i.e., inferences beyond entailments, including presuppositions, conversational implicatures, and conventional implicatures (Potts, 2005; Tonhauser, Beaver, Roberts, and Simons, 2013). We can define two classes of negation based on the 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*, marks grammatical negation and obviously affects the assertion. *No* and *not* constitute explicit negation. But it is important to remember that explicit negation does not necessitate that negation is morphologically overt. As noted already, expressions such as *few*, *scarcely*, *hardly*, *seldom*, and *little*, although not morphologically negative, behave nevertheless syntactically and semantically as negative under a number of well known diagnostics (Klima, 1964; Horn, 2001 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*. At the same time, it has also been noted that syntactic diagnostics alone are not sufficient for all kinds of explicit negation. 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 & Clark, 2008),[3] even though *only* does not have negative morphology, and is not syntactically negative based on the tests mentioned above.[4]

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 (e.g., *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 assert 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 (presuppositional or implicature). For current purposes, we consider the class of "emotive" predicates, which trigger negative inferences, though their

[1] Clark (1976) was specifically making a distinction between asserted negation and negation in the presuppositions of an expression. 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.

[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 (Beaver & Clark, 2008; Horn, 1996, 2002).

[4] There are good reasons why some of the syntactic tests 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 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.

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 “factive”; Kartunnen, 1971; Kiparski & 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:

(1) 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 (1a) 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 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, 1998, 2006). 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.

1.2. 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 (Clark & Chase, 1972; Wason, 1959, 1961) 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 et al., 1983; Kounios & 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. Furthermore, 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 (also see Just & Carpenter, 1971). 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 & Chase, 1972; Gough, 1965; Trabasso, Rollins, & Shaughnessy, 1971; Wason, 1961), leading to extra cost, as well as delayed processing of negation. The view that the processing of negation is delayed, however, is challenged by the ERP results of Nieuwland and Kuperberg (2008), who showed that once the relevant pragmatic conditions on using negation are met (e.g., “plausible denial”, Wason, 1965), sentential negation is incrementally processed and influences the N400 on an upcoming word (see also similar proposals from a sentence verification task in Greene, 1970; and Tian et al., 2010).

The discussion on the processing of negation, however, has primarily focused on the sentential negation *not*. It is therefore not totally clear yet what generalizations one can draw with regard to the processing profile of different types of negation. The current study examines the question of whether or not the time course of processing negation varies for different types of negation. 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* to gauge the time course of computing negation. Since *ever* is a closed-class item, the distribution of which is conditioned on grammatical factors, it can largely circumvent issues of low-level lexical semantic associations or interference from real world knowledge (e.g., *robin-bird*). The word *ever* is also 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.

1.3. Negative polarity items and previous ERP findings

1.3.1. Negative polarity items (NPIs)

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, 2012; von Fintel, 1999; Krifka, 1995; Chierchia, 2006; Kadmon & Landman, 1993; Laka, 1994). For current purposes, we will focus on

[5] Although negation is the most robust licensing element cross-linguistically, NPIs can also appear in non-negative contexts that are non-veridical (Giannakidou, 1998; Zwarts, 1995), i.e., in questions.

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

	Asserted negation			Non-asserted negation
	Morphologically negative	Syntactically negative	Negation in semantics (assertion)	Presupposed or implicated
No	✓	✓	✓	–
Few	–	✓	✓	–
Only	–	–	✓	–
Emotive	–	–	–	✓

four types of NPI licensors that contain some sort of negative feature, explicit or implicit. Examples of NPI licensing under these four types of licensors are given in (2) below. We also included a control example, (2e), in which the NPI is not licensed, resulting in an ungrammatical sentence.

(2) 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 expressed by the negative quantifiers *no* and *few*. In (c), it is licensed by the exclusive component of *only*. 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 non-asserted implicit negation (Baker, 1970; Giannakidou, 2006; Linebarger, 1987).[6]

Besides semantic differences, the licensors in (a–d) above also differ along morphological and syntactic dimensions. The licensor *no* is one of the canonical expressions of negation; *few* is 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 summarize 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 (assertion versus non-assertion). 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.

1.3.2. N400 and NPI licensing

The N400 is a negative-going waveform that peaks at approximately 400 ms, 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 & Hillyard, 1984; Lau, Phillips, & Poeppel, 2008; Kuperberg, 2013; Kutas & 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. Crucially, in the current case, although not all NPI licensors contain a lexical semantic or morphosyntactic negative feature (e.g., 2a–2d), they all 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.

In a series of studies on German (Drenhaus, Blaszczak, & Schutte, 2007; Drenhaus, Saddy, & Frisch, 2005, 2006; Saddy, Drenhaus, & Frisch, 2004), a reduced N400 with a central maximum was found on the German NPI *jemals* (‘ever’) when it was licensed by negation, compared to an ungrammatical counterpart in which *ever* was not licensed. Similar N400 effects were found for Dutch (Yurchenko et al., 2013), Turkish (Yanilmaz & Drury, 2013) and English NPIs (Shao & Neville, 1998), as well as 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. But only the negative

[6] There is also an approach where emotive predicates are *semantic* licensors by being Strawson downward entailing (e.g. von Fintel, 1999). This approach may make the prediction that emotive predicate should be processed just like other semantic licensors. The results reported in our study pose challenge to this approach.

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, and therefore, that the N400 effect 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 emerges. 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 suggests that a pre-activated semantic feature [+Neg] can facilitate the processing of an up-coming 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 verbs—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.

1.3.3. P600 and NPI licensing

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 reliably elicited by syntactic errors (Hagoort, Brown, & Groothusen, 1993; Osterhout & Holcomb, 1992) or grammatical but syntactically complex constructions (Gouvêa, Phillips, Kazanina, & Poeppel, 2010; Kaan, Harris, Gibson, & Holcomb, 2000; Osterhout, Holcomb, & Swinney, 1994; Phillips, Kazanina, & Abada, 2005); on the other hand, there is a growing body of work on the "semantic P600" effect (Bornkessel-Schlesewsky & Schlesewsky, 2008; Brouwer, Fitz, & Hoeks, 2012; Chow & Phillips, 2013; Kim & Osterhout, 2005; Kuperberg, 2007; van de Meerendonk, Kolk, Chwilla, & Vissers, 2009; Paczynski & Kuperberg, 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 (Bornkessel-Schlesewsky & Schlesewsky, 2008; Friederici & Weissenborn, 2007; Kuperberg, 2007; Van Petten & 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 the 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 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 in comparison 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 directly adopted to serve the grammatical function of NPI licensing.

2. Experiment 1

2.1. Methods

Stimuli creation

One-hundred-and-fifty items like (2) 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

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

[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, beim Graben, Saddy, and Frisch (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.

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.	5.8 (0.32)
Few	b. **Few** third-graders had ***ever*** seen one before.	5.9 (0.31)
Only	c. **Only** third-graders had ***ever*** seen one before.	5.5 (0.36)
Emotives	d. She was **surprised** that third-graders had ***ever*** seen one before.	5.5 (0.34)
Unlicensed	e. Third-graders had **ever** seen one before.	2.6 (0.38)

Standard deviations shown in parentheses.

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 an emotive 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 an emotive 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.

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 with fifty participants, who were self-reported native English speakers. Only IP addresses within the US were allowed to participate in the norming task. 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 < 0.0001). In addition, among licensed conditions, sentences with licensors *no* and *few* were rated higher than sentences with licensor *only* and factives (ts > 3; ps < 0.01).

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 (14 females). All were between the ages of 18 and 35 years old (mean = 24).

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. Each trial began with the context sentence centered on the screen. Participants pressed the center button of a controller to initiate presentation of the target sentence. The target sentence began with a crosshair ('+') that displayed for 500 ms, followed by a 150 ms blank screen. After the blank screen began the word-by-word presentation of the target sentence. Each word of the target sentence displayed for 400 ms in the center of the screen, followed by a 150 ms interstimulus interval (ISI), except for the sentence-final word (together with the sentence-final period), which displayed for 1000 ms. On trials containing a comprehension question, the question appeared directly after the sentence final word. 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.

EEG recording

EEG responses were recorded from 32 electrodes (BrainProducts, see Fig. 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

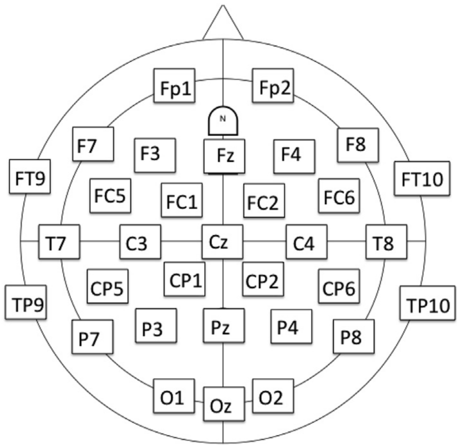

Fig. 1. Channel layout. **Midline**: Frontal (Fp1, Fp2, Fz); Central (FC1, FC2, Cz); Posterior (CP1, CP2, Pz); Parietal (O1, O2, Oz). **Peripheral**: Left Frontal (F7, F3, FC5); Right Frontal (F4, F8, FC6); Left Posterior (CP5, P7, P3); Right Posterior (CP6, P4, P8).

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 Fig. 1) offline. The EEG signal was continuously sampled at 1000 Hz, with a band pass of 0.1–70 Hz, and the impedance was kept below 5 kOhm.

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 (artifact rejection rate: no 16%; few 17%; only 16%; emotive 17%; ungrammatical 16%), were calculated relative to a 200 ms pre-stimulus baseline. At the critical word, we carried out two different analyses. First, all five conditions, including the four grammatical and one ungrammatical condition, were analyzed together. Second, we performed an additional analysis that only included the four grammatical conditions.

For the first analysis that included all conditions, based on visual inspection, we carried out analyses for two consecutive time windows that showed the most prominent effects: 250–400 ms, which encompassed the peak of the N400 effect, and 500–700 ms, which encompassed the P600 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: frontal (Fp1, Fp2, Fz), central (Fc1, Fc2, Cz), posterior (Cp1, Cp2, Pz), and parietal (O1, Oz, O2). The within-subject variables were Licensor (five levels, *no*, *few*, *only*, emotives, 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, P7, 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).

For the second analysis which focused primarily on the four grammatical conditions, the analysis procedure was largely similar to the procedures described above, except that a slightly different set of time windows was analyzed. We describe the details of this analysis below.

2.2. Results

2.2.1. Behavioral results

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

2.2.2. Event related potentials

2.2.2.1 The comparison between the grammatical and the ungrammatical conditions. N400: 250–400 ms. The mid-region omnibus ANOVA showed a main effect of Licensor ($F(4,144) = 4.7$, $p < 0.01$), and also an interaction between Licensor and Region ($F(12, 432) = 3.7$, $p < 0.001$). The peripheral omnibus ANOVA showed a main effect of Licensor ($F(4,144) = 4.0$, $p < 0.01$), and an interaction between Licensor and Region ($F(4,144) = 4.3$, $p < 0.01$), but the interaction between Hemisphere, Licensor, and Region, and the interaction between Hemisphere and Licensor were not significant ($Fs < 1.6$, $ps > 0.1$). Since both analyses

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

F(4, 144)	250–400 ms	500–700 ms
Mid-region		
Frontal	2.7*	1.5
Central	5.1***	1.5
Posterior	5.2***	3.8**
Parietal	4.2**	6.6***
Peripheral region		
Left-frontal	4.6**	1
Right-frontal	2.8*	1.6
Left-posterior	2	4.9***
Right-posterior	3.3*	5.5***

$^{\wedge}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Table 4
t-values for planned pairwise comparisons between each licensor and the unlicensed condition, in each ROI.

	No vs. unlicensed		Few vs. unlicensed		Only vs. unlicensed		Emotive vs. unlicensed	
T(36)	250–400 ms	500–700 ms	250–400 ms	500–700 ms	250–400 ms	500–700 ms	250–400 ms	500–700 ms
Mid-region								
Frontal	**2.3***	1.5	**2.8****	**2.1***	**2.5***	1.2	1.6	1.5
Central	**3.2****	0.6	**2.9****	0.07	**3.0****	0.2	**3.9*****	1.7
Posterior	**2.2***	**2.3***	1.3	**2.2***	1.8$^{\wedge}$	**2.3***	**4.2*****	0.4
Parietal	0.01	**3.7*****	1.1	**2.8****	0.1	**3.9*****	**2.5***	1
Peripheral-region								
Left-Frontal	**3.5****	0.2	**3.4****	1.4	**3.0****	0.05	2.0$^{\wedge}$	0.2
Right-Frontal	**32****	1.4	**2.7***	1.9$^{\wedge}$	**2.5***	1.8$^{\wedge}$	1.9$^{\wedge}$	1.9$^{\wedge}$
Left-Posterior	1.5	**3.1****	0.5	**2.5***	0.8	**3.7*****	**2.6***	1.4
Right-Posterior	1.6	**3.7*****	1.3	**2.7***	1.1	**2.9****	**3.4****	0.1

$^{\wedge}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

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 250–400 ms window, the effect of Licensor was significant for most of the ROIs across the whole scalp. For each of the ROIs, planned pairwise comparisons were carried out between the unlicensed/ungrammatical condition and each of the four licensed conditions (Table 4). The waveforms for the pairwise comparisons are presented in Fig. 2, and in Fig. 3 we present all five conditions together with two representative electrodes.

At the frontal and the central cites, for both middle and peripheral regions, the target NPI under the four licensors *no*, *few*, *only*, and emotive verbs all showed reduced negativity compared to the unlicensed condition (Table 4). The difference between the grammatical and the ungrammatical conditions was absent in the posterior and parietal areas for licensors *no*, *few and only*; but the emotives still showed reduced negativity compared to the unlicensed condition in the posterior-parietal area. The topographic maps in Fig. 2 confirm that the N400 reduction on the four grammatical conditions (relative to the ungrammatical one) is qualitatively similar across conditions in the central frontal area, but the emotives demonstrate additional negativity reduction in the posterior region.

P600: 500–700 ms. At the later time window 500–700 ms, the mid-region omnibus ANOVA did not show any main effect of Licensor ($F(4,144) = 1.9$, $p > 0.1$), but there was an interaction between Licensor and Region ($F(12, 432) = 4.6$, $p < 0.001$), driven by the fact that the effect of Licensor is mostly posteriorly distributed, with significant effects in posterior and parietal ROIs, but not in frontal or central ones (Table 3). The peripheral omnibus ANOVA did not reveal any main effect of Licensor ($F(4,144) = 1.3$, $p > 0.2$), but there was also an interaction between Licensor and Region ($F(4, 144) = 8.1$, $p < 0.001$), again driven by the fact that the effect of Licensor was only present in peripheral posterior but not frontal ROIs (Table 3). There was no effect of Hemisphere.

For each ROI, we again compared the unlicensed condition with each of the licensed conditions separately (Table 4 and Fig. 2). In the frontal and central areas, none of the comparisons revealed any differences ($ts(36)<1.5$, $ps > 0.1$). In the posterior and parietal areas, the unlicensed condition showed a larger positivity relative to conditions with the licensor *no*, *few*, and *only*. But strikingly, we found no difference between the emotive and the unlicensed conditions (Table 4).

To summarize, the comparison between the licensed and the unlicensed NPIs reveals two main findings. First, we replicated previous findings in the literature that the processing cost on unlicensed NPIs (relative to the licensed ones) is mainly reflected as increased N400 and P600 amplitude.[9] We discuss the specific interpretations of these two components in

[9] Visual inspection on Fig. 2 suggests an enhanced central-frontally distributed late negativity (800–900 ms) associated with the ungrammatical condition. But it didn't reach significance in any ROIs. Among all the previous ERP studies on NPIs, only Steinhauer et al. (2010) reported a late negativity for the unlicensed NPIs.

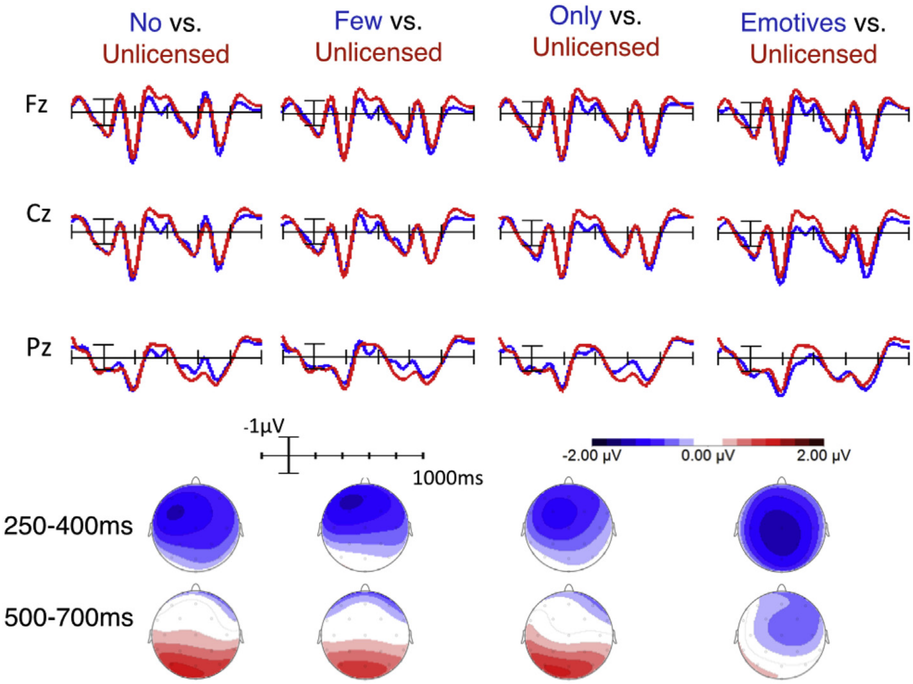

Fig. 2. Top: Waveforms that compare each of the grammatical licensors with the ungrammatical condition, on the critical word *ever*. The blue line represents a grammatical licensor, and the red line represents the ungrammatical condition. **Bottom**: topographical maps from two time windows. Each plot represents the difference resulting from subtracting a grammatical condition from the ungrammatical condition. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

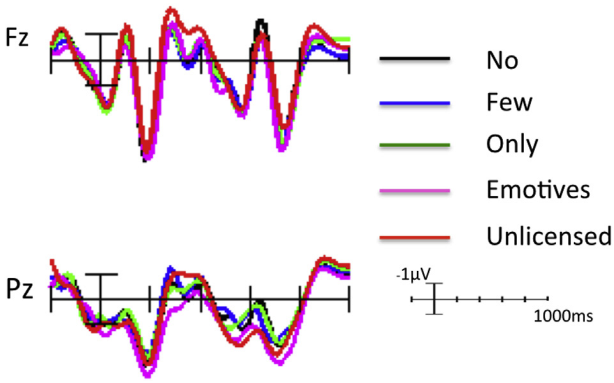

Fig. 3. Waveforms from two representative midline channels: frontal Fz and posterior Pz, with all five conditions shown together.

the General Discussion section. Second, a crucial difference emerged between NPIs licensed under the explicit negation (*no*, *few*, and *only*) and those licensed under the implicit negation (emotives). During the 250–400 ms time window, although all licensed NPIs showed a reduced N400 relative to the unlicensed condition, the N400 reduction under *no*, *few*, and *only* is mainly distributed in the central-frontal regions, while the N400 reduction under emotives is more extensive, extending from the central-frontal areas into the posterior-parietal regions. Furthermore, during the 500–700 ms time window, the emotive condition, different from all other licensed conditions, showed a P600 amplitude that is as large as that of the unlicensed condition in the posterior-parietal regions. Visual inspection on channel Pz in Fig. 3 suggests that the difference between the emotives and other grammatical conditions in the posterior-parietal regions is mainly driven by the fact that the emotive condition elicited an enhanced positivity that started very early and continued into the later time window. To more precisely quantify the difference between the emotives and the other three grammatical conditions, we carried out an additional analysis of just the four grammatical conditions.

2.2.2.2 Comparison between the explicit and the implicit negation. The waveforms and topographic maps that compare the emotive condition with the other three grammatical conditions are presented in Fig. 4. Visual inspection suggests that the emotive condition elicited an extensive positivity with an early onset in the posterior-parietal areas. We analyzed three consecutive time windows: 0–200 ms, 200–400 ms, and 450–750 ms. The ROI division and the ANOVA procedure is the same as in the analysis presented earlier. Since we found no significant effect in the 0–200 ms time window (see Table 5), we only focus on the 200–400 ms and the 450–750 ms time windows in our report below.

200–400 ms. The mid-region omnibus ANOVA did not find a main effect of Licensor ($F(3,108) = 2.1$, $p > 0.1$). There was a significant Licensor x ROI interaction ($F(9,324) = 2.0$, $p < 0.05$), driven by the fact that an effect of Licensor was found in the posterior and the parietal regions, but not the frontal and central regions (Table 5). The peripheral omnibus ANOVA found neither a main effect of Licensor ($F(3,108) = 0.3$, $p > 0.7$), nor an interaction involving ROI or Hemisphere (Fs < 2, ps > 0.1). Planned pairwise comparisons (Table 6) further confirmed that, in the central posterior and the parietal regions, the emotive condition elicited larger positivities than the *no, few*, and *only* conditions. Moreover, post-hoc pairwise comparisons found no differences among the explicit negation conditions *no, few*, and *only* (all ps > 0.3).

450–750 ms. The mid-region omnibus ANOVA found a main effect of Licensor ($F(3,108) = 2.8$, $p < 0.05$), and a significant Licensor × ROI interaction ($F(9,324) = 2.0$, $p < 0.05$), driven by the fact that an effect of Licensor was significant in the posterior and the parietal regions, and marginally significant in the central region, but not significant in the frontal ROI (Table 5). The peripheral omnibus ANOVA didn't find a main effect of Licensor (($F(3,108) = 1.6$, $p > 0.1$), but there was a Licensor x ROI interaction ($F(3,108) = 3.4$, $p < 0.05$), driven by the significant effect of Licensor in the Left and Right posterior ROIs (with the effect in the left posterior ROI marginally significant, Table 5). There was no effect of Hemisphere.

As shown by the planned pair comparisons in Table 6, between 450 and 750 ms, the emotive condition elicited a larger positivity than the other three licensors in both mid- and peripheral posterior and parietal regions. The positivity reduction on *no* and *few* also extended into the mid-central ROI. Further post-hoc comparisons between the three explicit negation conditions (with licensors *no*, *few*, and *only*) again did not reveal any significant differences among them (all ps > 0.3).

2.3. Summary and discussion of Experiment 1

In summary, the findings in Experiment 1 revealed both similarities and differences between the NPIs licensed under *explicit* negation and those licensed under *implicit* negation. They are similar in the sense that all conditions that contain a legitimate licensor, explicit and implicit negation alike, showed a qualitatively similar central-frontally distributed N400 reduction during the 250–400 ms time window at the critical NPI, relative to the unlicensed condition. However, during the P600 time window, the implicit negation (i.e., the emotive condition) elicited larger positivities than the other three explicit negation conditions, both in an early time window 200–400 ms and at a later time window 450–750 ms, with the latter time window largely overlapping with the enhanced P600 effect elicited by the ungrammatical condition.

Before we discuss further what these results implicate for the processing of negation (and the licensing of NPIs), we want to first address a potential concern: that the emotive predicate condition involves a different structure from the other conditions—i.e., that it 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 be 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. In fact, as we will discuss later, it is possible to reinterpret the enhanced positivity on the emotive condition in the early 200–400 ms window as a reduced negativity due to the linear position of the critical word in this condition (e.g., Van Petten & Kutas, 1990). But the more important question is whether or not the enhanced P600 effect in the later time window (450–750 ms) on the emotive condition can be attributed to a structural property of this construction. For example, it is possible that the structural complexity (embedded clause) of the emotive condition made it more difficult to process than the other grammatical conditions, leading to an enhanced P600 amplitude on the emotive condition. A similar concern was raised in Steinhauer et al. (2010), who argued that any differences one observes between different licensing environments are not necessarily associated with NPI licensing per se, but could 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. Since a full design of this sort requires a total of ten conditions (five with NPIs, and five without), which is impractical for an ERP study, we adopted a self-paced reading paradigm in Experiment 2 to address this question.

3. Experiment 2

Experiment 2 has two primary goals. First, it examines whether the basic pattern of results in Experiment 1 can 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 × 5-design self-paced reading study. The first five conditions are the same as

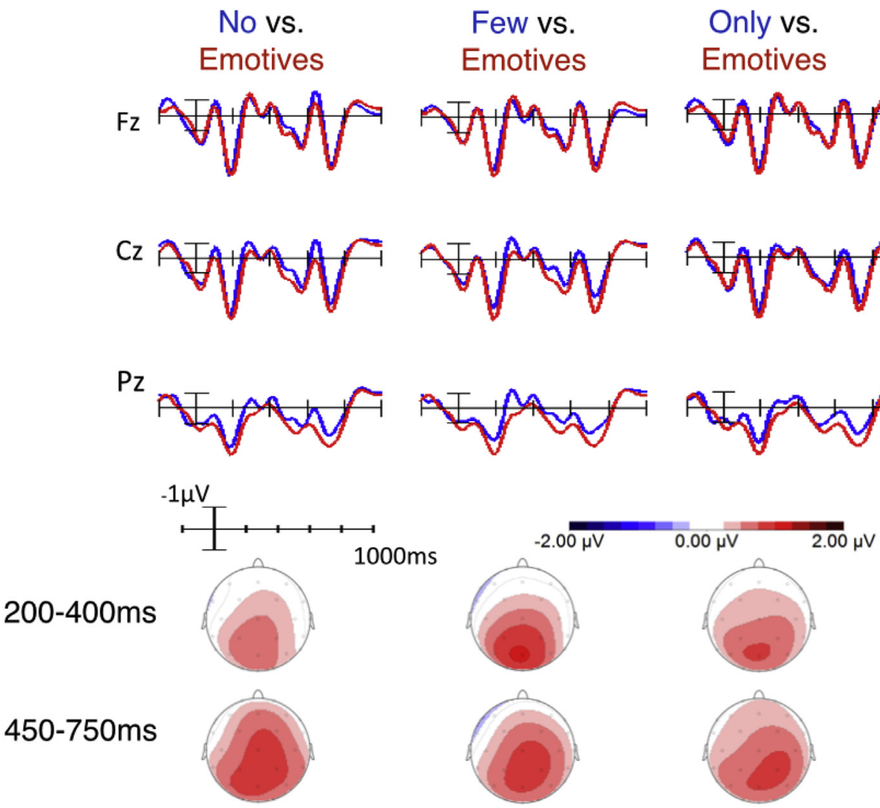

Fig. 4. Top: Waveforms that compare each of the licensors *no*, *few*, and *only* with the emotive condition, on the critical word *ever*. The red line represents the emotive condition. **Bottom**: Topographical maps from two time windows. Each plot represents the difference resulting from subtracting the *no/few/only* condition from the emotive condition. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
ANOVA F-values for the effect of Licensor in each ROI in the mid-region and peripheral region. Only the four grammatical conditions were included in the analysis.

F(3, 108)	0–200 ms	200–400 ms	450–750 ms
Mid-region			
Frontal	0.07	0.09	0.5
Central	0.5	0.8	2.4^{Δ}
Posterior	2.0	$\mathbf{4.4}^{**}$	$\mathbf{3.8}^{*}$
Parietal	2.1	$\mathbf{5.8}^{**}$	$\mathbf{3.6}^{*}$
Peripheral region			
Left-Frontal	0.2	0.5	1.1
Right-Frontal	0.03	0.03	0.3
Left-Posterior	0.8	1.3	2.5^{Δ}
Right-Posterior	0.4	1.9	$\mathbf{4.4}^{**}$

$^{\Delta}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Table 6
t-values for planned pairwise comparisons between licensors *no*, *few*, *only*, and the emotive condition, in each ROI.

	No vs. emotives		Few vs. emotives		Only vs. emotives	
t(36)	200–400 ms	450–750 ms	200–400 ms	450–750 ms	200–400 ms	450–750 ms
Mid-region						
Frontal	0.08	0.4	0.4	0.6	0.04	0.6
Central	1.1	$\mathbf{2.5}^{*}$	1.4	$\mathbf{2.1}^{*}$	1.4	1.6
Posterior	$\mathbf{2.2}^{*}$	$\mathbf{2.7}^{*}$	$\mathbf{3.1}^{**}$	$\mathbf{2.8}^{**}$	$\mathbf{3.5}^{**}$	$\mathbf{2.7}^{*}$
Parietal	$\mathbf{2.9}^{**}$	$\mathbf{3.3}^{**}$	$\mathbf{4.0}^{***}$	$\mathbf{2.2}^{*}$	$\mathbf{2.5}^{*}$	$\mathbf{2.3}^{*}$
Peripheral region						
Left-Frontal	0.6	0.5	1	1.8^{Δ}	0.2	0.06
Right-Frontal	0.9	0.9	1.6	0.3	1.6	0.7
Left-Posterior	0.04	$\mathbf{2.3}^{*}$	0.2	1.7^{Δ}	0.1	$\mathbf{2.3}^{*}$
Right-Posterior	1.7^{Δ}	$\mathbf{3.9}^{***}$	1.8^{Δ}	$\mathbf{2.2}^{*}$	$\mathbf{2.3}^{*}$	$\mathbf{2.4}^{*}$

$^{\Delta}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Table 7
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.

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 7.

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, among the four licensed conditions, we also expect the emotive condition (d) to show enhanced reading times compared to the other three grammatical conditions. Moreover, it is possible that all four licensed conditions initially show similar processing costs, and that the additional processing cost on the emotive condition only appears 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*) may then serve as the baseline control conditions to assess whether the predicted effects in the NPI conditions can be explained by factors other than their NPI licensing properties. For example, if the observed effects in the NPI conditions are driven by the fact that the emotive condition involves an embedded clause and the critical word is in a different linear position in the sentence, then a similar effect should also arise in the Plain conditions. If the difference in effects between emotives and the other three licensors only appear among the conditions containing NPIs, then we can conclude that they were driven by NPI-licensing per se.

3.1. Method

Material

Sixty sets of experimental items were created, each with 10 conditions. An example is shown in Table 7 . These items were all taken from the materials 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 immediately following *ever* in the original NPI condition. In the example in Table 7 , 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.

3.2. 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 two standard derivations from the mean was excluded. We present in Fig. 5 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.

Linear mixed-effects models were conducted for data analysis. Parallel to the crucial comparisons in the ERP data analysis, our primary interest in this experiment was examining whether or not the following two sets of comparisons showed 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*). To directly assess the effects of these two comparisons, and to avoid excessive multiple comparisons, we set up the crucial comparisons as two contrasts with Helmert coding (Vasishth & Broe, 2011; Venables & Ripley, 1999), and these two contrasts were then included in the mixed-

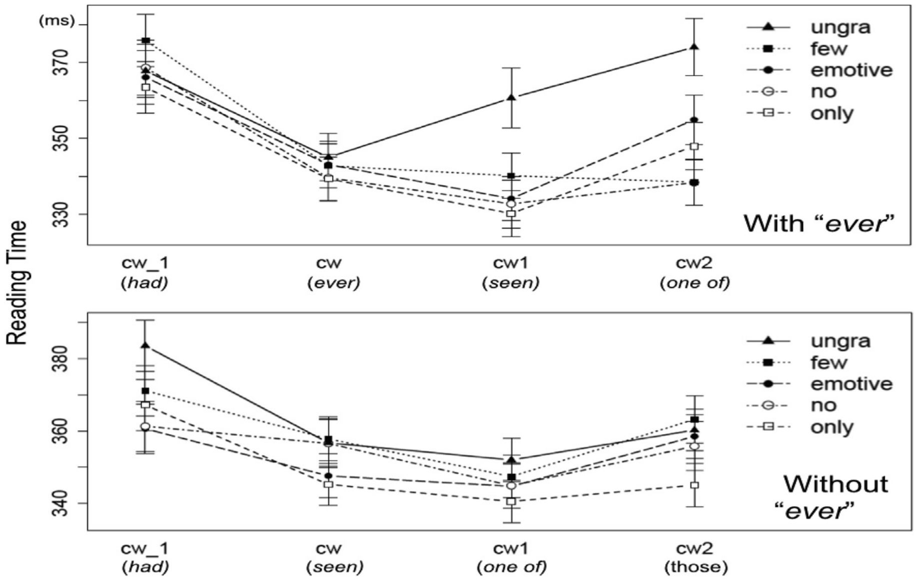

Fig. 5. Raw reading times from Experiment 2.

effects models as two fixed effect predictors. Separate models were created 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*, and *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*, and *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*, *and i*; and the effect of the emotive predicate: condition *i* vs. *f*, *g*, *and h*. For the analysis on any given region, we also included the reading time from the two immediately preceding regions 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-effects 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 would 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 $ps < 0.0001$) 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).[10] 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 show delayed effects from the critical word on the 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 participants were sensitive to the 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. It is also worth noting that even though by and large the emotive condition is more costly than the other grammatical conditions at CW+2 (Contrast 2 in Table 8), pair comparisons between the emotives and each of the grammatical licensors only found significant difference between the emotives and *no*/*few* ($ps < 0.05$), but not with *only* ($p = 0.5$). This is in line with the offline acceptability judgments reported earlier, in which *only* and the emotives were both judged slightly less acceptable than the *no* and *few* conditions. But this is distinct from the ERP results on the critical word *ever*, in which *no*, *few* and *only* patterned similarly, and were all less costly than the emotives. We do not have any hypothesis as to why the behavior of the *only* condition varies depending on the task. It is possible that *only*,

[10] We also created a more traditional mixed-effects models for the whole data set, with NPI (2 levels: NPI or no-NPI) and Licensor (5 levels) as the fixed-effects predictors (again controlling for spill-over effects). No significant effect was found on the CW. On the CW+1 region, there was an *NPI x Licensor* interaction when all ten conditions were considered ($p < 0.05$), but no interaction when only the eight conditions with licensors (i.e., excluding conditions e&j) were considered ($p > 0.9$). Crucially, the latter interaction was significant on the CW+2 region ($p = 0.01$). This is completely in line with the results reported in Table 8.

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)			
	Estimate	Std. error	t-value	p-value	Estimate	Std. error	t-value	p-value
With "ever"								
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

being a focus marker, could demand additional processing effort at late stage of the sentence processing, leading to longer RTs at spill over regions and also degraded offline ratings (see Xiang, Grove, and Giannakidou (2013) for more experimental data on *only* as a NPI licensor).

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 among the NPI conditions are associated with position of the NPI or the structural complexity of the sentences containing it.

3.3. Summary of Experiment 2

Although that information obtained from the self-paced reading paradigm is less fine-grained than ERP results, the findings in Experiment 2 are informative in two important respects. First, they show that the basic findings in Experiment 1 are replicable in a different experimental paradigm. In the ERP experiment, we found that the four grammatical conditions showed similar N400 reduction in the early time window, but in the later P600 window, the emotive condition showed a larger P600 amplitude than the other three licensed conditions. In the self-paced reading experiment, we showed that the four grammatical conditions initially showed similar reading time, but that the emotive condition elicited a longer reading time at a later time point. The fact that the replication of the basic contrasts only appeared in the **NPI-present** conditions but not in the **Plain** conditions is crucial in ruling out the possibility that the observed effects among the NPI conditions might 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 difference should have been observed regardless of the presence of an NPI.

Admittedly, some caution needs to be taken in drawing parallels 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 differences, 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.

4. 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 of these licensors helped to trigger a reduced central-frontally distributed 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 later 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 emotive condition elicited a larger positivity in the posterior-parietal area, relative to conditions containing asserted negation (e.g., *no*, *few*, and *only*), resulting in quantitatively similar processing cost in the later time window between the emotive and 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.

4.1. Negative meaning from different sources is computed immediately

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 quickly in online comprehension. This result is in line with Nieuwland and Kuperberg (2008) and

Tian et al. (2010)—but the unique contribution from the current study is that we showed that both implicit and explicit negation can be computed quickly. The finding regarding the N400 effects in Experiment 1 provided strong evidence for this conclusion. 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 (Lau et al., 2008; Kuperberg, 2013; Kutas & Federmeier, 2011). 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 or 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 an emotive predicate (i.e., one of the four NPI licensors) contains negative meaning, explicitly or implicitly, therefore activating a [+Neg] feature prior to the appearance of the NPI word *ever*. This pre-activation led 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 morphological to pragmatic (Table 1). The fact that they all triggered a reduced N400 on the target NPI word suggests that the computation of negation is NOT delayed until a complete affirmative proposition has been comprehended, regardless of how the negative information is encoded. In particular, our results suggest that rapid incremental computation is carried out, not only for the obvious instances of negation like *no*, but 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).

4.2. Pragmatically driven processing cost with implicit negation

Although different types of negation can be rapidly processed by the early N400 time window, we still observed substantial differences between explicit/asserted negation (i.e., negation from *no*, *few*, and *only*) and implicit/non-asserted negation (i.e., negation from emotive predicates). In the posterior-parietal areas, the emotive condition elicited larger positivities than the other three licensed conditions in both an early (200–400 ms) and a late time window (450–750 ms).

We view the late positivities in the 450–750 ms time window as belonging to the group of standard P600 effects. Two slightly different P600 effects were observed in the current results. The first is the larger P600 amplitude on the emotive condition in the 450–750 ms window; and the other is the larger P600 amplitude on the ungrammatical condition in the 500–700 ms window. The two P600 effects have similar topographic distributions, with overlapping but slightly different time windows. It has been proposed that P600 broadly indexes combinatorial processing in response to conflict when different sources of linguistic information are integrated (e.g., Kuperberg, 2007). In the context of NPI licensing, a number of authors have suggested that processing during 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). But we also want to distinguish the two P600 effects we observed in the emotive-factive and unlicensed conditions, as they reflect two kinds of integration cost. The enhanced P600 on the ungrammatical condition is associated with the failure to construct a well-formed grammatical representation. The enhanced P600 on the grammatical emotive condition, on the other hand, reflects the pragmatic processing cost that integrates an NPI into the context with a non-asserted negation. NPI licensing in regular circumstances requires asserted negation (as in the conditions with *no*, *few*, and *only*); although pragmatic licensing as a secondary licensing mechanism is possible with non-asserted negation (see more discussion in the next section), it comes with some processing cost. Pragmatically driven P600 effects have been observed for other grammatical but pragmatically demanding situations, such as irony (Regel, Gunter, & Friederici, 2011; Spotorno, Cheylus, Van Der Henst, & Noveck, 2013), indirect requests (Coulson & Lovett, 2010), and metaphors (Grauwe, Swain, & Holcomb, 2010).

As for the larger positivity on the emotive condition in the early time window (200–400 ms), we speculate two possible interpretations. On the one hand, it is possible to reinterpret this positivity as a reduced negativity. Linear word position has been shown to impact the N400 amplitude. For example, Van Petten and Kutas (1990) observed a reduced N400 amplitude during the course of encountering a sentence. This effect has mainly been associated with strengthened anticipation of the upcoming word. In the current design, the critical NPI was not cued by any contextual expectations or constraints, making it less at risk of for word-position effect. Nevertheless, it is still possible that the critical word in the emotive condition elicited a reduced negativity compared to other grammatical conditions, due to the fact that the critical word is always in an embedded clause in the emotive condition and therefore in a later linear position in the sentence. The second possibility is that this early effect is indeed an enhanced positivity, and reflects pragmatic processes that started very early. In previous ERP studies, some pragmatic effects have been shown to elicit early positivities. For example, Regel, Meyer, and Gunter (2014) observed that sentences containing irony, compared to literal sentences, elicited P200 followed by a P600 component, suggesting that pragmatic processes can start very early when appropriate contextual information is provided.

4.3. Two distinct NPI licensing mechanisms

The contrast between the emotive and the other licensed conditions provide evidence supporting the existence of two distinct mechanisms for NPI licensing (e.g., Giannakidou (1998, 2006)'s "dual mode" of licensing). 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 have negative entailments, thus contributing negation at the level of the grammatical representation (semantics), which is immediately accessible from the syntax. The other, secondary mode of licensing sanctions the NPI by accessing a negative inference pragmatically.[11] This pragmatic mechanism is relevant for emotive licensors. The syntactic-semantic licensing mechanism is the primary licensing mechanism that the comprehension system consults. The comprehension system resorts to pragmatic inferences to license NPIs only in case there is no negation available in the syntax-semantics, and this happens, crucially, at an additional processing cost. If negation is not present either in the syntactic-semantic or the pragmatic levels of representation, as in the ungrammatical condition, the comprehension process fails to construct a coherent interpretation, leading to an enhanced P600 (in some situations, pragmatic inferences may lead to grammatical "illusions", see Xiang et al. (2013) for a discussion).

In the context of the clear difference between the emotive factives and the other grammatical conditions, it is again worth noting the similarities between these conditions in the early N400 window in the central-frontal regions. 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 target word: the more relevant the features that have been pre-activated in the previous context, the easier it is to process the target word, and hence the more reduction of the N400 amplitude. The N400 itself does not necessarily 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. As we discuss above, the N400 reduction on the target word *ever* appeared for all licensed conditions, suggesting that explicit and implicit negation have both been successfully computed quickly by the N400 time window. However, as the ERP results in the posterior-parietal areas suggest, it takes additional processing effort for the comprehension system to ultimately license NPIs with an implicit negation, whereas explicit negations can be recruited for NPI licensing directly.

The ERP evidence that suggests additional pragmatic processes on the emotive condition is corroborated by evidence from both the self-paced and the offline rating results. It took people a longer time to read the emotive condition in the spill-over region, compared to the other licensed conditions. In offline acceptability ratings, the emotive condition is also rated as less acceptable than at least the licensors *no* and *few*.

5. 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 have shown that negative information from different sources, e.g., from asserted and non-asserted content, is accessed equally rapidly in online processing. At the same time, however, we observe that asserted and explicit negation, i.e., negation in the syntactic-semantic representation, is applied immediately to grammatically license NPIs; pragmatically inferred negation, or implicit negation, on the other hand, is only adopted at a later processing stage as a secondary NPI-licensing mechanism, leading to additional pragmatic processing cost.

Acknowledgments

The authors thank Lelia Glass, Steven SanPietro, Genna Vegh and Katina Vradelis for their assistance in constructing the stimuli and collecting the data.

References

Atlas, J. D. (1993). The importance of being only: testing the neo-gricean versus neo-entailment paradigms. *Journal of Semantics, 10*, 301–318.
Baker, C. L. (1970). Double negatives. *Linguistic Inquiry, 1*, 169–186.
Beaver, D., & Clark, D. (2008). *Sense and Sensitivity: How focus determines meaning*. Oxford: Blackwell.
Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews, 59*, 55–73.
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.
Chierchia, C. (2006). Broaden your views. Implicatures of domain widening and the 'logicality' of language. *Linguistic Inquiry, 37*, 535–590.
Chow, W. Y., & Phillips, C. (2013). No semantic illusion in the "Semantic P600" phenomenon: ERP evidence from Mandarin Chinese. *Brain Research, 1506*, 76–93.
Clark, H. H. (1976). *Semantics and comprehension*. The Hague: Mouton.
Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, 3*, 472–517.
Coulson, S., & Lovett, C. (2010). Comprehension of non-conventional indirect requests: an event-related brain potential study. *Italian Journal of Linguistics, 22*, 107–124.
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.

[11] It should be noted that not all negative inferences license NPIs; see discussions in (Giannakidou, 2006; Horn, 2001; 2002; Linebarger, 1987). A negative presupposition, as is the case with emotives, is part of 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).

Drenhaus, H., Blaszczak, J., & Schutte, J. (2007). Some psycholinguistic comments on NPI licensing. In E. Puig-Waldmuller (Ed.), *Proceedings of Sinn und Bedeutung* (Vol. 11, pp. 180–193). Barcelona, Spain: Universitat Pompeu Fabra.
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.
von Fintel, K. (1999). NPI-licensing, strawson-entailment, and context-dependency. *Journal of Semantics, 16*, 97–148.
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., & Garrett, M. F. (1975). The psychological unreality of semantic representations. *Linguistic Inquiry, 6*, 515–531.
Friederici, A. D., & Weissenborn, J. (2007). Mapping sentence form onto meaning: the syntax-semantic interface". *Brain Research, 1146*, 50–58.
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 H. Klaus, C. Maienborn, & P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning de Gruyter* (pp. 1660–1712).
Giannakidou. The subjunctive as evaluationd nonveridicality, epistemic subjunctive, and emotive-as-expressive. In J. Blaszack, D. Klimek-Jankowska, K. Mygdalski, & A. Giannakidou (Eds.), *Tense, mood, and modality: New Perspectives on Old Question, (to appear)*, 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.
De Grauwe, S., Swain, A., & Holcomb, P. (2010). Electrophysiological insights into the processing of nominal metaphors. *Neuropsychologia, 48*, 1965–1984. http://dx.doi.org/10.1016/j.neuropsychologia.2010.03.017.
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, & 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)).
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.
Hoeksema, J. (2012). On the natural history of negative polarity items. *Linguistic Analysis, 38*(1/2), 3–33.
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* (2nd). Stanford, CA: Center for the Study of Language and Information. Originally published in 1989 (Chicago: University of Chicago Press).
Horn, L. R. (2002). Assertoric inertia and NPI licensing. In *Proceedings of the Annual Meeting of the Chicago Linguistic Society* (Vol. 38). University of Chicago.
Just, M. A., & 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.
Kadmon, N., & Landman, F. (1993). Any. *Linguistics and Philosophy, 16*, 353–422.
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.
Kiparsky, P., & Kiparsky, C. (1970). Fact. In M. Bierwisch, & K. Heidolph (Eds.), *Progress in Linguistics*. The Hague: Mouton [Reprinted in D. Steinberg & L. Jakobovits, eds. Semantics. an interdisciplinary reader, Cambridge: Cambridge University Press.].
Klima, E. S. (1964). Negation in English. In J. A. Fodor, & J. J. Katz (Eds.), *The structure of language* (pp. 246–323). Englewood Cliffs, NJ: Prentice-Hall.
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, G. R. (2007). Neural mechanisms of language comprehension: challenges to syntax. *Brain Research, 1146*, 23–49.
Kuperberg, G. R. (2013). The pro-active comprehender: what event-related potentials tell us about the dynamics of reading comprehension. In B. Miller, L. Cutting, & P. McCardle (Eds.), *Unraveling the behavioral, neurobiological, and genetic components of reading comprehension* (pp. 176–192). Baltimore: Paul Brookes Publishing.
Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621–647.
Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*, 161–163.
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, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: [de]constructing the N400. *Nature Reviews Neuroscience, 9*, 920–933.
Linebarger, M. (1980). *The grammar of negative polarity*. Ph.D. dissertation. MIT. Cambridge, MA.
Linebarger, M. (1987). Negative polarity and grammatical representation. *Linguistics and Philosophy, 10*, 325–387.
van de Meerendonk, N., Kolk, H. H. J., Chwilla, D. J., & Vissers, C. T. W. M. (2009). Monitoring in language perception. *Language & Linguistics Compass, 3*, 1211–1224.
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.
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., & 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.
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.
Regel, S., Gunter, T. C., & Friederici, A. D. (2011). Isn't it ironic? an electrophysiological exploration of figurative language processing. *Journal of Cognitive Neuroscience, 23*, 277–293. http://dx.doi.org/10.1162/jocn.2010.21411.
Regel, S., Meyer, L., & Gunter, T. (2014). Distinguishing neurocognitive processes reflected by P600 effects: evidence from ERPs and neural oscillations. *PLoS One, 9*(5). http://dx.doi.org/10.1371/journal.pone.0096840.
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.
Spotorno, N., Cheylus, A., Van Der Henst, J.- B., & Noveck, I. A. (2013). What's behind a P600? integration operations during irony processing. *PLoS One, 8*(6), e66839. http://dx.doi.org/10.1371/journal.pone.0066839.
Stalnaker. (1978). Assertion. In P. Cole (Ed.), *Syntax and semantics 9* (pp. 315–332). New York: New York Academic Press.
Steinhauer, S., Drury, J., Portner, P., Walenski, M., & 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., Beaver, D., Roberts, C., & Simons, M. (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., & 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 Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: benefits, costs, and ERP components. *International Journal of Psychophysiology, 83*, 176–190.
Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory and Cognition, 18*(4), 380–393.
Vasishth, S., & Broe, M. (2011). *The foundations of statistics: A simulation-based approach*. Berlin: Springer.
Venables, W. N., & Ripley, B. D. (1999). *Modern applied statistics with s-PLUS*. NewYork, NY: Springer. http://dx.doi.org/10.1007/978-1-4757-3121-7.
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, M., Dillon, B., & Phillips, C. (2009). Illusory licensing effects across dependency types: ERP evidence. *Brain and Language, 108*, 40–55.
Xiang, M., Grove, J., & Giannakidou, A. (2013). Dependency dependent interference: NPI interference, agreement attraction, and global pragmatic inferences. *Frontiers Psychology, 4*, 708. http://dx.doi.org/10.3389/fpsyg.2013.00708.
Yanilmaz, A., & 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 contexts. *Linguistic Analysis, 25*, 286–312.