Alexander Klapheke & Kathryn Davidson

2019

NPI Intervention: Crosslinguistic Data

NPI Intervention: Crosslinguistic Data

Alexander Klapheke & Kathryn Davidson

Harvard University

{klapheke,kathryndavidson}@fas.harvard.edu

1 Introduction

The question of where and how negative polarity items (NPIs) such as English *any* and *ever*, Italian *mai*, and French *le moindre* are licensed is a long-standing question in linguistics research. Complicating this is the presence of *interveners*—elements which unilaterally block NPIs in their scope. It is claimed that the class of factive predicates intervene in Italian (1a) and French (1b), but not in English (idiomatic glosses of (1)).

(1) a. **Rimpiango che questo sia mai accaduto.*
       I regret   that this   had ever happened

   'I regret that this ever happened.'   (Chierchia, 2015)

   b. **Jean ne sait   pas que Marie a   la   moindre chance de gagner.*
       John   know not that Mary has the least   chance of winning

   'John doesn't realize that Mary has any chance of winning.'   (Homer, 2008)

We investigate this variation in a judgment study in English ($n = 46$), Italian ($n = 29$), and French ($n = 34$) of sentences with NPIs under factives. We find that, consistent with previous claims, this effect seems to exist in Italian and not English, but that, against expectation, French does not pattern with either language.

2 Background: NPIs

2.1 NPI licensing

NPIs, such as *any*, *ever*, *yet*, *either*, *at all*, *in years*, and *bother to*, are famously licensed in negative environments (2a) but not positive ones (2b).

(2) a. Sue didn't ever eat cake.
    b. *Sue ever ate cake.

Accounts have variously been semantic, such as Baker's (1970) theory that NPIs are licensed in sentences with negative implicata, or syntactic, such as Linebarger's (1980) theory that they are licensed in the immediate scope of a negative operator. The most commonly-cited account is that of Ladusaw (1979), who postulates that NPIs are licensed just in downward-entailing (DE) environments which reverse the direction of entailment—that is, for p, q such that $p \Rightarrow q$, $\text{env}(q) \Rightarrow \text{env}(p)$—and not in upward-entailing (UE) environments which preserve it. Negation, for instance, is downward entailing: without it, entailment is from more specific cases to more general (*Sue ate chocolate cake* entails *Sue ate cake*; cf. (2b)), but with it, the reverse is true (*Sue didn't eat cake* entails *Sue didn't eat chocolate cake*; cf. (2a)). Chemla et al. (2011) provide experimental evidence for this account, showing that whether French speakers judge the scope of a quantifier to be downward or upward entailing closely predicts whether they judge the NPI *le moindre* 'the least' grammatical or not.

Evidence from neurolinguistics, however, less clearly supports one kind of hypothesis; in ERP studies of unlicensed NPIs, Shao & Neville (1998), find that they elicit anterior negativity between

300–500 ms, associated with morphosyntactic errors such as feature mismatch, Saddy et al. (2004), find an N400 response, associated with semantic incongruities, and Xiang et al. (2009) find a P600 response, associated with syntactic errors such as valence errors. Complicating matters further is the phenomenon of *illusory* licensing, by which an NPI *preceded* but not *licensed* by a DE operator, as in (3), is erroneously judged as grammatical.

(3) *The diplomats that no congressman could trust have ever supported a drone strike.

Drenhaus et al. (2005) find *both* an N400 and a P600 response for such sentences, and Parker & Phillips (2016), in speeded acceptability and self-paced reading tasks, show that illusory licensing is selective, occurring with *ever* but not with *any*, and that it can be mitigated by introducing grammatical distance or a time delay between 'licenser' and NPI, indicating that these are not general licensing failures, but the products of more complex underlying processes. It is these complex interactions which may give rise to phenomena such as factive intervention.

One flaw in the DE theory is pointed out by von Fintel (1999), who notes that it undergenerates: NPIs are licensed in certain environments (4) which turn out not to be downward-entailing (5).

(4) a. Only Sue ate any cake.
    b. Bill regrets ever buying a car.

(5) a. Only Sue ate cake. $\nRightarrow$ Only Sue ate chocolate cake.
    b. Bill regrets buying a car. $\nRightarrow$ Bill regrets buying a red car.

(6) a. Only Sue ate chocolate cake. $\gg$ Sue ate chocolate cake.
    b. Bill regrets buying a red car. $\gg$ Bill bought a red car.

Von Fintel (1999) notes that in each inference in (5), a presupposition is introduced in the conclusion which is not present in the premise (6), shifting the common ground mid-evaluation; if these are satisfied, the downward inference again holds. He proposes a weaker licensing condition, termed *Strawson downward-entailingness*, which disregards any presuppositions when computing entailment. Importantly, this means that the content of a presupposition should never interfere in NPI licensing.

2.2 Intervention

Linebarger (1980) notes that certain elements such as conjunctions and universal quantifiers, which appear between an NPI and a DE operator, block the latter from licensing the former.

(7) She didn't wear any earrings to every party.

    a. $\neg > \exists_{NPI} > \forall$ (No particular pair was always worn.)
    b. *$\neg > \forall > \exists_{NPI}$ (She wore no earrings to some parties.)

The only reading of (7) is (a), in which the universal element *every* does not come between the NPI *any* and its licensor, *didn't*. This is in contrast with the equivalent sentence with *some* replacing *any*, as in (8), in which both readings are possible.

(8) She wore some earrings to every party.

    a. $\exists > \forall$ (Some particular pair was always worn.)
    b. $\forall > \exists$ (She wore one pair or another to every party.)

The ungrammaticality of (7b) is therefore not due to a general constraint on the raising of the universal quantifier. Linebarger posits that it is due to an *immediate structure constraint* requiring NPIs to be in the *immediate* scope of negation. Other elements, however, such as disjunctions and existential quantifiers, do not intervene:

(9)   a.   The professor didn't give some students any homework.
   b. * The professor didn't give every student any homework.

(10)   a.   I don't care about John or give a damn about Mary.
   b. * I don't care about John and give a damn about Mary.

Chierchia (2004) theorizes that this is because, in a DE environment, such elements generate a positive implicature; e.g., *Sue didn't eat all the cookies* implicates that *Sue ate some of the cookies*. Thus, (7) will generate the implicature *She did wear any earrings to some party*. If this implicature is generated as a grammatical process *before* NPI licensing is evaluated, then the implicature itself will contain an unlicensed NPI, rendering the interpretation in (7b) unacceptable.

Denić et al. (2018) test this explanation experimentally, reasoning that if this account is true, speakers' judgments of sentences like (9) as ungrammatical should correlate with their propensity to compute implicatures with the same items. However, they find no such correlation.

A similar contrast is identified by Homer (2009): *too* but not *either* intervenes:

(11)   a. * *Context: Mary read an interesting book.* John$_F$ didn't read anything interesting too.
   b.   *Context: Mary didn't read anything interesting.* John$_F$ didn't read anything interesting either.

This seems to pose a problem for the Strawson-DE theory above, since the contributions of these particles is thought to be presuppositional (Krifka, 1995). Ahn (2016) proposes in response that *too* and *either* are not, in fact, presuppositional, but rather, respectively, a disguised conjunction and disjunction, thus reducing this contrast to the one in (10).

Other problematic contrasts exist, however. In particular, factive predicates, but not non-factive ones, are claimed to intervene in Italian by Chierchia (2015) and in French by (Homer, 2008). Factivity is generally thought to be a presupposition (Karttunen, 1971). Interestingly, factives do not seem to intervene in English, as shown in (1). In our paper we focus on this cross-linguistic contrast, gathering quantitative data across the three languages and different types of embedding predicates and NPIs, in order to determine the empirical status of this claim and how it should be considered in a theory of negative polarity item licensing. If it proves true, then either there is some cross-linguistic difference in factive verbs between English and Romance; or the Straswon-DE theory must be replaced, either by a more liberal version which incorporates presuppositions as Homer (2009) suggests, or by a difference licensing mechanism entirely;[1] or else it must be amended for certain environments such as the scope of factives. We discuss these possibilities in more detail in Section 7.

3   Experimental conditions

In determining the acceptability of NPIs under factives, it is important to examine internal variation in these categories. In this section, we compare cognitive with emotive factives, and weak with strong NPIs.

3.1   Types of factive

The class of factive predicates has been known to contain internal variation since Karttunen (1971). In particular, two broad classes of factives are often recognized: cognitive factives such as *discover* and emotive factives such as *regret*.

---

[1]Alternative theories of NPI licensing which do not rely on downward entailingness exist, such as Giannakidou's (1998) theory that they are licensed in the broader category of *nonveridical* environments, or Barker's (2018) theory of NPIs as existentials with obligatory narrow scope, a comparison of which is beyond the scope of this paper.

Abusch (2002, 2010) observes that cognitive factives are *soft* presupposition triggers which can be cancelled by placement in a DE environment (12a), while emotive factives are *hard* triggers whose presupposition projects in all environments (12b).

(12) a. If I later discover that I've lied, I'll confess. ≫̸ I lied.
  b. If I later regret that I've lied, I'll confess. ≫ I lied.

Abusch posits that soft triggers are not presuppositions at all, but implicatures, which are known to be defeasible in such environments.

In addition, Djärv et al. (to appear) note that the complements of cognitive factives—but not emotive factives—are *at-issue*, meaning that, even as embedded clauses, they can serve as answers to questions:

(13) Where was Harriet yesterday?

  a. — Henry discovered that she had an interview at Princeton.
  b. *— Henry regretted that she had an interview at Princeton.

Djärv et al. find that, when presented with responses in a discourse that contradict a prior factive complement, English speakers prefer responses that contradict *emotive* factives.

Finally, emotive factives, but not cognitive factives, license NPIs in English:[2]

(14) a. *Carol discovered that she had left any food in the fridge.
  b. Carol regretted that she had left any food in the fridge.

This final difference is especially significant for a study involving NPIs.

3.2 Types of NPI

Like factives, NPIs are also not monolithic, and come in at least two varieties: *weak* and *strong* (Zwarts, 1998). While weak NPIs can be found in the environments discussed in 2.1, strong NPIs are not licensed in certain environments, although they are Strawson DE, such as the complements of emotive factives (15) and non-neg-raising predicates (16).

(15) a. I regret ever meeting you.
  b. *I regret meeting you in years.

(16) a. I didn't claim I'd ever seen her.
  b. #I didn't claim I'd seen her in years. (for the reading in which *in years* modifies *see*)

NPIs with similar 'strong' distributions are found in other languages as well, such as Italian *affatto* 'at all', German *im Entferntesten* 'in the slightest', and Dutch *ook maar* 'anything'.

With these differences in mind, we test both emotive and cognitive factives, and, when possible, both weak and strong NPIs, in order to account for such differences.

4 Experiment 1

A judgment study of English sentences was conducted on Amazon Mechanical Turk to investigate the empirical status of the claim that factives do *not* intervene in English. In addition to studies such as Schnoebelen & Kuperman (2010), which generally validate the results of psycholinguistic

---

[2]A reviewer points out that not all factives pattern so neatly—in particular, *Henry realized that…* may be highly marked as an answer to (13), although it is a cognitive factive, and *be happy* does not license NPIs, although it is an emotive factive: **John is happy that he had ever started smoking*. This deserves to be examined in more detail, and may bear on the theoretical aspects of factive intervention; however, what is most relevant for us is that the emotive factive used is an NPI licenser in English, and the cognitive factive used is not.

studies done on Mechanical Turk against the same studies performed in the lab with undergraduates, Sprouse (2011) finds that, in particular, the results of acceptability judgment studies of English-speaking Mechanical Turk workers ('Turkers') track those of lab studies closely enough to serve as reliable substitutes for them, and so we take advantage of this method of collecting data quickly for our study.

4.1 Participants

48 participants were recruited on Amazon Mechanical Turk. Participants were restricted to the United States and Canada, and were paid 1 USD. Two participants were excluded for failing our screening task: they rated weak NPIs under *think* without negation—i.e., an environment with no possible licenser—*better* than corresponding sentences without the NPI.

4.2 Procedure

Sentences were presented along with a Likert scale from 1 (worst) to 5 (best), which was compiled using TurkTools (Erlewine & Kotek, 2016). This is shown in Figure 1. Participants were instructed to rate the 'naturalness' of the sentences, and examples of grammatical and ungrammatical sentences (due to island violations) were provided for reference.

---

1. 1: Shane is surprised that Rachel has at some point eaten carrots.

2: **No, he isn't surprised that she has ever eaten carrots.**

1	2	3	4	5
○	○	○	○	○

---

2. 1: Sadie thinks that Lester has at some point eaten olives.

2: **Yeah, she thinks that he has ever eaten olives.**

1	2	3	4	5
○	○	○	○	○

Figure 1. Part of questionnaire, as seen by participants

As discussed in Section 3.1, the two varieties of factive predicate, cognitive and emotive, have quite different properties. Thus, one of each type was tested. In addition, for non-factive predicates, both neg-raising and non-raising predicates were tested. A negative-raising, or *neg-raising* predicate is one in which an embedded negation is interpreted as wide-scope, strengthening the meaning. For instance, a sentence such as (17a) with the neg-raising predicate *think* admits the interpretation in (17b), but not so with the non-neg-raising predicate *claim* in (18).

(17) a. I don't think I've seen her.
    b. I think I haven't seen her.

(18) a. I didn't claim I'd seen her.
    b. I claimed I hadn't seen her.

It is plausible that the low interpretation of matrix negation would affect its ability to license an embedded NPI. In addition, neg-raising predicates have certain salient properties, such as admitting strong NPIs (19), and in fact Richter & Radó (2014) find German strong NPIs to be rated more highly than weak NPIs under neg-raising verbs.

(19) a. I don't think I've seen her in years.
    b. #I didn't claim I'd seen her in years.

Thus, four predicates in total were used in each experiment. The same predicates were used for each language, so they were chosen so as to have the same properties in each language—most importantly, they all easily embed 'that'-clauses. Finally, two types of NPI were tested (control sentences had no NPI). *Strong NPIs* such as *in years* or *either* are excluded from certain DE environments (including the complements of emotive factives; see Section 3.1) in which *weak NPIs* such as *ever* or *any* are licit.

Each participant was shown one of eight lists of 24 sentences, with each list containing, in a fully random order, all combinations of four matrix predicates (non-factive, non-neg-raising *claim*; non-factive, neg-raising *think*; cognitive factive *realize*; and emotive factive *be surprised*) × 2 matrix polarities (negation, no negation) × 3 kinds of embedded NPI (weak *ever*, strong *in years*, and control sentences with no NPI). Lists were presented in a pseudorandom order to participants such that each list was seen roughly the same number of times. 'Realize' was chosen over the more common 'know' because French (tested in Experiment 3 in Section 6) *savoir* 'know' does not easily embed 'that'-clauses. The sentences were presented in the context of conversations between people about dietary habits, samples of which are shown in Figure 1. The conversational contexts were meant to neutralize any pragmatic strangeness of having negated sentences, or NPIs with exhaustive interpretations—particularly Italian *affatto* 'at all'—out of the blue. Food names were normalized to count nouns, since one of the French NPIs used, *le moindre*, is incompatible with mass nouns. No filler items were used.
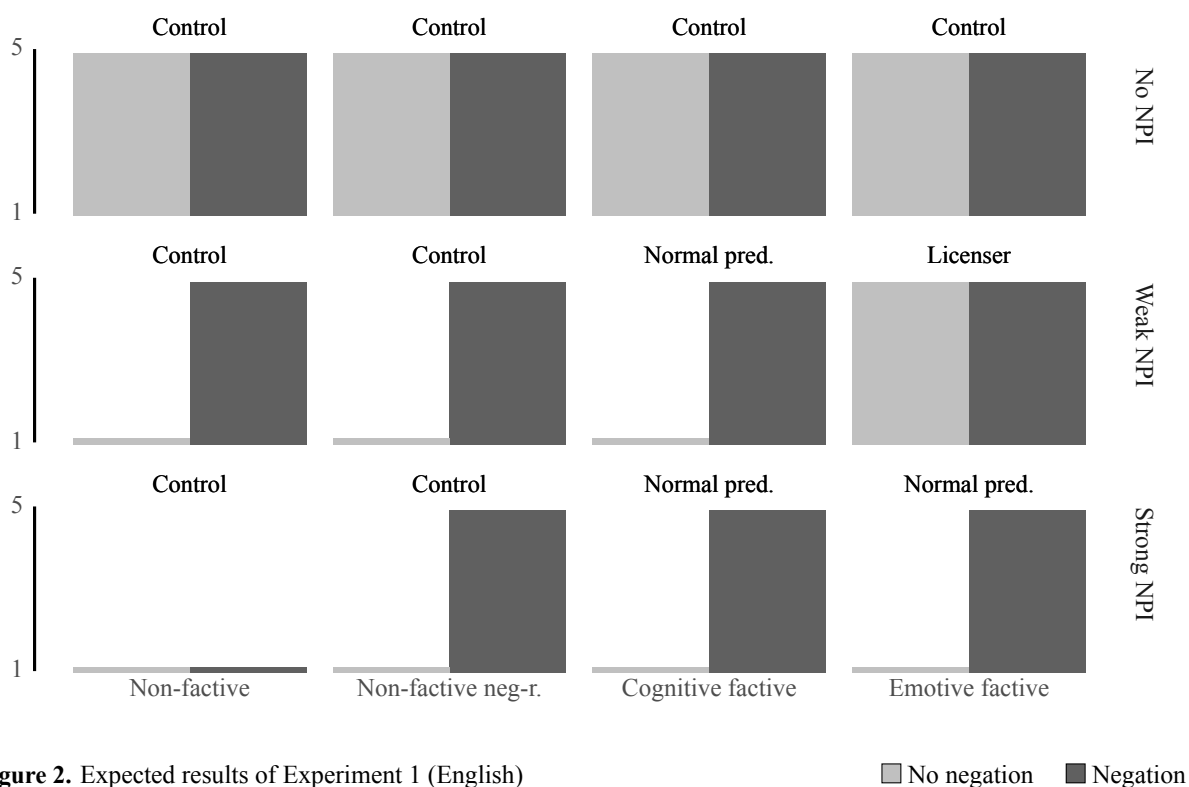
Figure 2. Expected results of Experiment 1 (English)   ☐ No negation   ■ Negation

The pattern of results expected is in Figure 2. These show the effect of adding negation to a particular sentence pattern—in particular, the 'rescuing' effect on NPIs which are otherwise not licensed. Control sentences with no NPI are at ceiling with or without negation (top row). Control sentences with non-factive predicates allow weak NPIs if there is a licensing negation (left two columns). Factive predicates, not being of a special category in English, behave like the control predicates, except that emotive factives license weak NPIs (even without negation).

4.3   Results and analysis

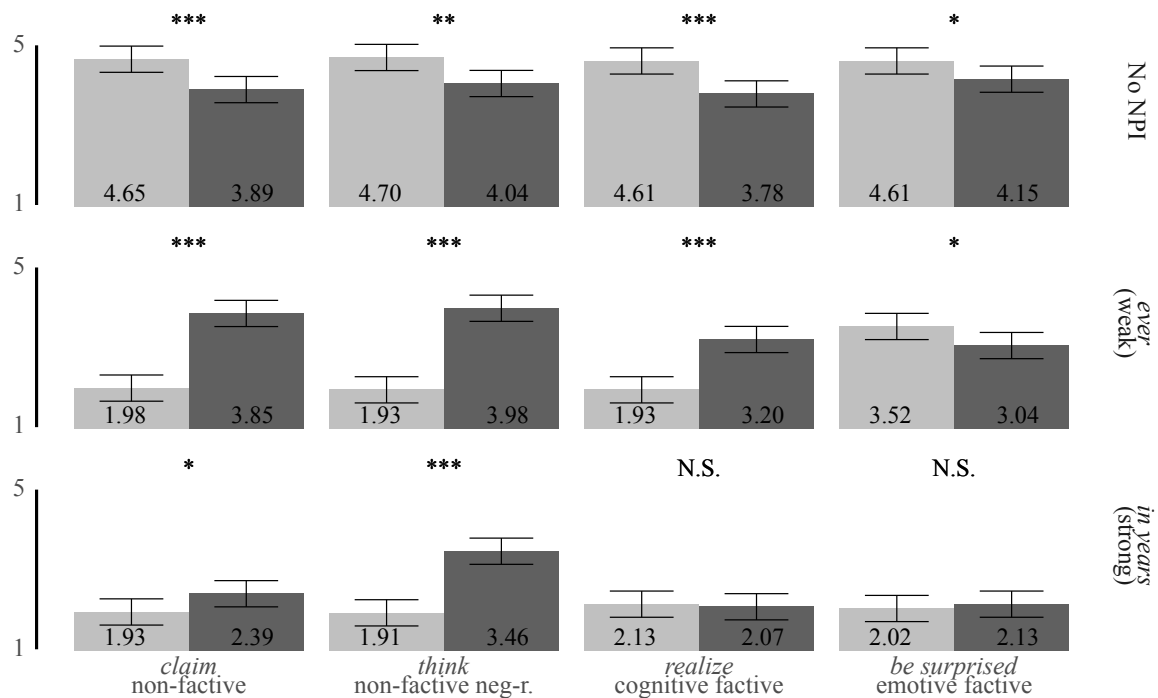The results of Experiment 1 are shown in Figure 3.

Figure 3. Results of Experiment 1 (English), $n = 46$. In all graphs, bars show 95% confidence limits, and significance levels are: *** $p < .001$, ** $p < .01$, * $p < .05$ □ No negation   ■ Negation

We analyze the data in a linear mixed effects model (LMEM)[3] using the lme4 v1.1.15 (Bates, Mächler, Bolker & Walker, 2015) package for R v3.4.4 (R Core Team, 2016), with the NPI, matrix verb, and matrix polarity as fixed effects, and Worker ID as random effect.[4] Since names and foods were counterbalanced in each list, they were not expected to significantly affect results when added to the model as random effects, and adding list number, food, and name as random effects does not improve model fit (ΔAIC $= 2.0$, ΔBIC $= 7.0$). List, name, and food are therefore omitted from the model.

The lsmeans v2.27-61 (Lenth, 2016) package was used to calculate least-squares means, which are shown in Figure 3. For each verb–NPI condition, a comparison between negation and no negation was made, in order to determine whether negation 'rescued' the NPI. In the control sentences with no NPI, there was a significant effect of negation for *claim* ($t(1035) = 3.576$, $p = .0004$), *think* ($t(1035) = 3.065$, $p = .002$), *realize* ($t(1035) = 3.883$, $p = .0001$), and *be surprised* ($t(1035) = 2.146$, $p = .03$). With the weak NPI *ever*, there was a significant effect of negation for *claim* ($t(1035) = -8.788$, $p = 6 \times 10^{-18}$), *think* ($t(1035) = -9.605$, $p = 6 \times 10^{-21}$), *realize* ($t(1035) = -5.926$, $p = 4 \times 10^{-09}$), and *be surprised* ($t(1035) = 2.248$, $p = .02$). With the strong NPI *in years*, there was a significant effect of negation for *claim* ($t(1035) = -2.146$, $p = .03$) and *think* ($t(1035) = -7.255$,

---

[3]While this data is technically categorical, it was collected in numeric form from participants; we therefore treat it as continuous here on the assumption that participants would find meaningful a rating of, say, 3.5, although such an option was not presented.

[4]Barr et al. (2013) recommend maximal models including as many random slopes as are justified by the experimental design in order to minimize the risk of a type I error—however, all models more complex than the one used failed to converge, a contingency noted by Bates, Kliegl, Vasishth & Baayen (2015). Therefore, our model includes only random intercepts.

$p = 8 \times 10^{-13}$), but not for *realize* ($t(1035) = 0.307, p = .8$) or *be surprised* ($t(1035) = -0.511, p = .6$).

To ensure that variation in the acceptability of NPIs under different predicates is not due to the acceptability of the predicates themselves, we also checked variation between the verbs themselves in the condition without NPIs. None of the predicates differed from the calculated mean either without negation: *claim* ($t(1035) = 0.083, p = 1$); *think* ($t(1035) = 0.417, p = 1$); *realize* ($t(1035) = -0.250, p = 1$); *be surprised* ($t(1035) = -0.250, p = 1$); or with negation: *claim* ($t(1035) = -0.584, p = 1$); *think* ($t(1035) = 0.584, p = 1$); *realize* ($t(1035) = -1.418, p = 0.6$); *be surprised* ($t(1035) = 1.418, p = 0.6$).

4.4 Discussion

Since the difference between the predicates was not significant in the baseline condition without negation and without NPI, any differences in NPI licensing can safely be said to be due to semantic properties of these predicates, such as factivity, and not to pragmatic idiosyncrasies of the predicates themselves.

In sentences with no NPI, there is, unexpectedly, a slight but significant dip in acceptability when negated. This may be a consequence of the well-known processing cost of negation (Wason, 1959, 1961; Gough, 1965; Trabasso et al., 1971; Just & Carpenter, 1971; Slobin, 1966), for which many explanations have been postulated (for an overview, see Horn, 1989, 186ff.). Wason (1965), for instance, explains the pragmatic effect of negation as a denial of some (possibly tacit) affirmative sentence, and finds lower processing costs for statements such as *a whale is not a fish*, which deny a more plausible affirmation, than for *a whale is not a bird*. This dip would be surprising, then, given that each sentence was a response—affirmation or denial—to some context sentence. Nevertheless, the effect is reported numerously across the literature for a wide range of sentences and contexts.

With the weak NPI *ever* there is, as expected, a major 'rescuing' effect of negation in the first three predicates. The fourth, being an emotive factive, is already an NPI licenser, and negation has no effect. Interestingly, for weak NPIs under negation, there is a small but significant intervention effect of factivity in English—the cognitive and emotive factives are together rated less highly than the non-factive predicates ($t = 3.955, p = 1.096 \times 10^{-4}$).

5 Experiment 2

Our next experiment turned to Italian to investigate whether factives intervene as claimed, in contrast to English.

5.1 Participants

34 participants were recruited via Amazon Mechanical Turk. Participants were restricted to Italy, and paid 1 USD. Two participants were excluded for rating weak NPIs under 'think' without negation *better* than corresponding sentences without the NPI.

5.2 Procedure

The procedure was identical to that of Experiment 1. Participants were shown a list in one of eight random orders. The lists paralleled those in Experiment 1 in the relative orders of verb, NPI, negation, and lexical items.

The verbs used were non-factive, non-neg-raising *sostenere* 'claim'; non-factive, neg-raising *pensare* 'think'; cognitive factive *rendersi conto* 'realize'; and emotive factive *essere sorpreso* 'be surprised' The NPIs used were weak *mai* 'ever' and strong *affatto* 'at all.'
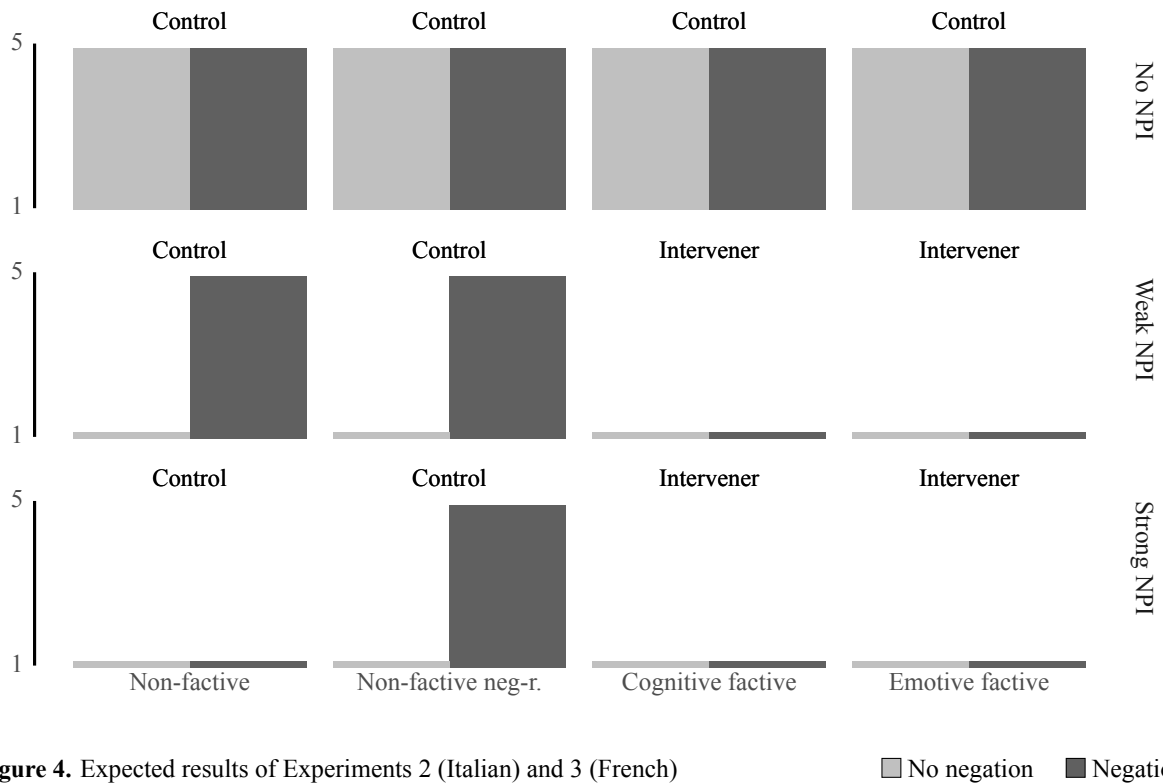
Figure 4. Expected results of Experiments 2 (Italian) and 3 (French)   ▢ No negation   ▪ Negation

The pattern of results expected is in Figure 4. Control sentences with no NPI are at ceiling with or without negation (top row). Control sentences with non-factive predicates allow weak NPIs if there is a licensing negation (left two columns). Finally, in Italian (and French) factive predicates are claimed to be interveners—they do not allow any NPI, with or without negation.

5.3 Results and analysis

The results of Experiment 2 are shown in Figure 5.

We analyze the data in a linear mixed effects model (LMEM) using lme4 for R, with the NPI, matrix verb, and matrix polarity as fixed effects, and Worker ID as random effect. Adding list number, food, and name as random effects does not improve model fit (ΔAIC $= 2.0$, ΔBIC $= 6.5$), and they are therefore omitted from the model.

The lsmeans package was used to calculate least-squares means, which are shown in Figure 5. For each verb–NPI condition, a comparison between negation and no negation was made, in order to determine whether negation 'rescued' the NPI. In the control sentences with no NPI, there was no significant effect of negation for *sostenere* ($t(644) = -0.143$, $p = 0.9$), *pensare* ($t(644) = 0.572$, $p = .6$), *rendersi conto* ($t(644) = 0.716$, $p = .5$), or *essere sorpreso* ($t(644) = 1.288$, $p = .2$). With the weak NPI *mai*, there was a significant effect of negation for *sostenere* ($t(644) = -7.584$, $p = 1 \times 10^{-13}$), *pensare* ($t(644) = -7.298$, $p = 9 \times 10^{-13}$), *rendersi conto* ($t(644) = -2.576$, $p = .01$), and *essere sorpreso* ($t(644) = -1.717$, $p = .09$). With the strong NPI *affatto*, there was a significant effect of negation for *sostenere* ($t(644) = -3.291$, $p = .001$) and *pensare* ($t(644) = -3.148$, $p = .002$), but not for *rendersi conto* ($t(644) = -0.859$, $p = .4$) or *essere sorpreso* ($t(644) = -1.288$, $p = .2$).

To ensure that variation in the acceptability of NPIs under different predicates is not due to the acceptability of the predicates themselves, we also checked variation between the verbs themselves in the condition without NPIs. None of the predicates differed from the calculated mean either

without negation: *sostenere* ($t(644) = -0.643$, $p = 0.9$); *pensare* ($t(644) = 1.227$, $p = .7$); *rendersi conto* ($t(644) = -1.344$, $p = .7$); *essere sorpreso* ($t(644) = 0.759$, $p = .9$); or with negation: *sostenere* ($t(644) = 0.584$, $p = 1$); *pensare* ($t(644) = 1.285$, $p = .6$); *rendersi conto* ($t(644) = -1.519$, $p = .5$); *essere sorpreso* ($t(644) = -0.351$, $p = 1$).
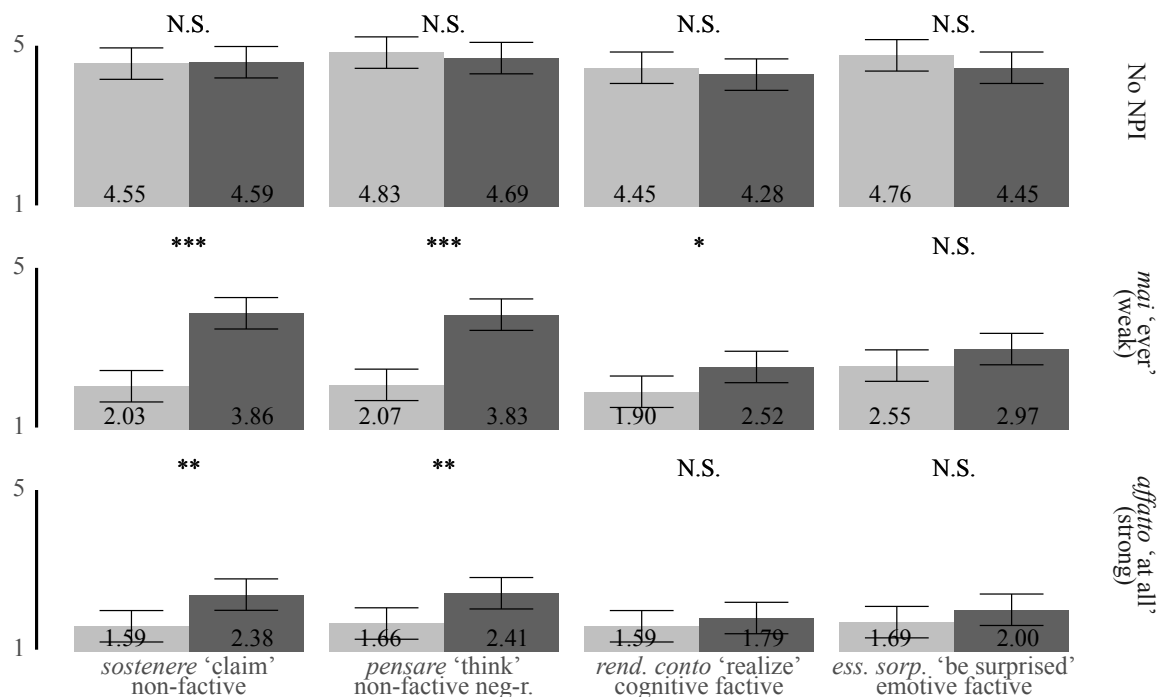

Figure 5. Results of Experiment 2 (Italian), $n = 28$ ☐ No negation   ■ Negation

5.4 Discussion

As expected, sentences with no NPI were rated at ceiling, although interestingly, there is no negation penalty in Italian as there is in English. With the weak NPI *mai*, negation rescues the NPI under non-factive *sostenere* and *pensare* and fails to under factive *essere sorpreso*, although the mean rating for the predicate 'be surprised' under negation was close to English (English: 3.04, Italian 2.97). With *rendersi conto*, there is a slight but significant rescuing effect of negation. While the predicates with a weak NPI seem to pattern more or less as in English, there is not nearly so much of a 'rescuing' effect in the factive predicates, which indicates that some kind of intervention effect may be at work.

With the strong NPI *affatto*, nearly all predicates were rated at floor, although rescued a small amount under the nonfactive predicates. Since the items were presented in the form of a dialogue, the NPI should not have seemed 'out of the blue'. Potentially, some other pragmatic infelicity prevented these sentences from being accepted.

6 Experiment 3

Our final experiment turned to a second Romance language, French, in which we conducted a study to test the robustness of the claim of factive intervention in Romance.

6.1 Participants

34 participants were recruited on Amazon Mechanical Turk.[5] Participants were restricted to Canada and France, and were paid 1 USD. No participants were excluded based on screening criteria (as for our English and Italian studies, rating sentences with NPIs and no negation *better* than those with negation).

6.2 Procedure

The procedure was identical to that of Experiment 1. Participants were shown a list in one of eight random orders. The lists paralleled those in Experiments 1 and 2 in the relative orders of verb, NPI, negation, and lexical items.

The verbs used were non-factive, non-neg-raising *déclarer* 'claim'; non-factive, neg-raising *penser* 'think'; cognitive factive *réaliser* 'realize'; and emotive factive *être surpris* 'be surprised' Due to the lack of a suitable French strong NPI (and the relative paucity of NPIs in French in general), two weak NPIs were used: *quelque* N *que ce soit* 'any N at all' and *le moindre* 'the least.' The latter was listed as the 'strong NPI' for the purpose of randomizing the lists; hence, it appears in the same respective position in the lists as *in years* and *affatto*. However, it is analyzed here as a weak NPI.

The pattern of results expected is identical to the previous experiment (see Figure 4).

6.3 Results and analysis

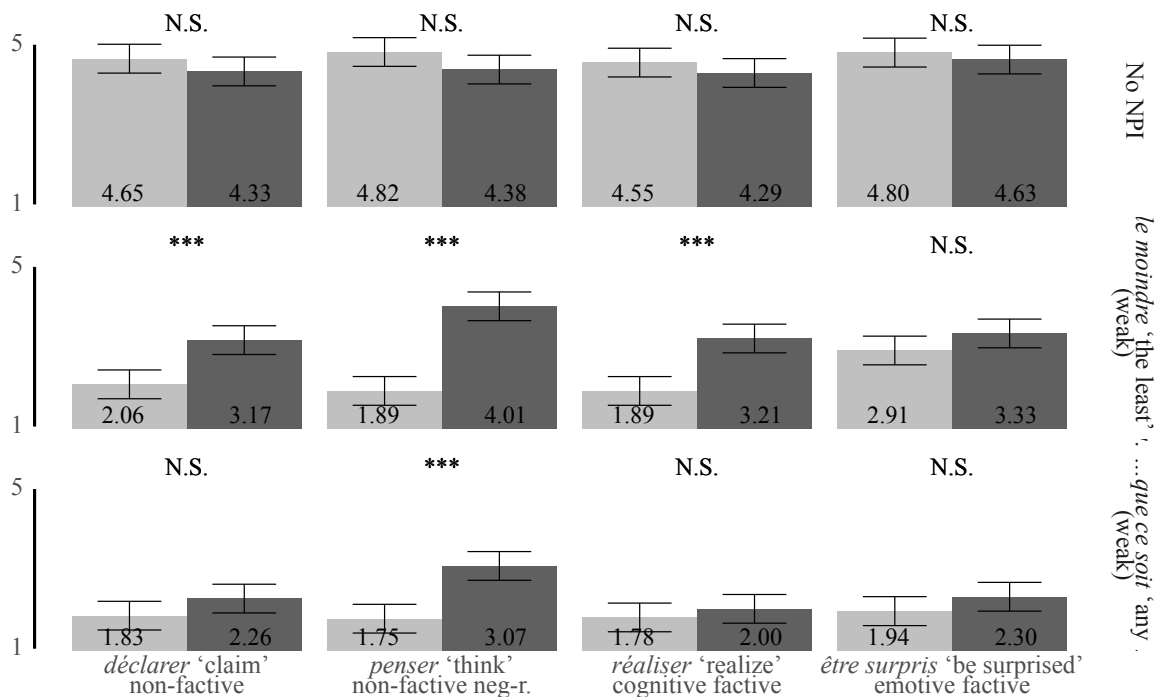The results of Experiment 3 are shown in Figure 6.

Figure 6. Results of Experiment 3 (French), $n = 34$ ☐ No negation   ■ Negation

We analyze the data in a linear mixed effects model (LMEM) using lme4 for R, with the NPI, matrix verb, and matrix polarity as fixed effects, and Worker ID as random effect. Adding list number, food,

---

[5]The lists were checked for dialectal differences by native speakers of French and Canadian French.

and name as random effects does not improve model fit (ΔAIC = 2.0, ΔBIC = 6.7), and they are therefore omitted from the model.

The lsmeans package was used to calculate least-squares means, which are shown in Figure 6. For each verb–NPI condition, a comparison between negation and no negation was made, in order to determine whether negation 'rescued' the NPI. In the control sentences with no NPI, there was no significant effect of negation for *penser* ($t(759) = 1.850$, $p = .06$), but not for *déclarer* ($t(759) = 1.335$, $p = .2$), *réaliser* ($t(759) = 1.096$, $p = .3$), or *être surpris* ($t(759) = 0.733$, $p = .5$). With the weak NPI *le moindre*, there was a significant effect of negation for *déclarer* ($t(759) = -4.626$, $p = 5 \times 10^{-6}$), *penser* ($t(759) = -8.959$, $p = 3 \times 10^{-18}$), *réaliser* ($t(759) = -5.515$, $p = 5 \times 10^{-8}$), and *être surpris* ($t(759) = -1.801$, $p = .07$). With the weak NPI *quelque ...que ce soit*, there was a significant effect of negation for *déclarer* ($t(759) = -1.797$, $p = .07$) and *penser* ($t(759) = -5.555$, $p = 4 \times 10^{-8}$), but not for *réaliser* ($t(759) = -0.904$, $p = .4$) or *être surpris* ($t(759) = -1.504$, $p = .1$).

To ensure that variation in the acceptability of NPIs under different predicates is not due to the acceptability of the predicates themselves, we also checked variation between the verbs themselves in the condition without NPIs. None of the predicates differed from the calculated mean either without negation (*déclarer* ($t(759) = -0.407$, $p = 1$); *penser* ($t(759) = 0.813$, $p = 1$); *réaliser* ($t(759) = -1.017$, $p = 1$); *être surpris* ($t(759) = 0.610$, $p = 1$); or with negation (*déclarer* ($t(759) = -0.610$, $p = 1$); *penser* ($t(759) = -0.203$, $p = 1$); *réaliser* ($t(759) = -0.813$, $p = 1$); *être surpris* ($t(759) = 1.626$, $p = .4$).

6.4 Discussion

The two weak NPIs tested in French yielded quite disparate results. *Le moindre* is a minimizer, which has been analyzed as decomposable into *even* and a superlative morpheme (Gonzalez & Moracchini, 2016), similar to the English minimizer *a single*. This item roughly followed the pattern of English *ever*: it displayed little, if any, effect of factive intervention.

The other NPI, *quelque ... que ce soit*, was rated poorly in all environments except under negated *penser*. If this due to intervention, it is apparently an effect *both* of factive and of non-neg-raising predicates. On the other hand, there may be pragmatic properties of this particular item that precluded its acceptance in the contexts given. For one, the item is relatively infrequent in spoken French, the register the sentences were presented as being in, which may have compromised their felicity, despite all sentences having been checked by native French speakers. Although this NPI is clearly somewhat felicitous under *penser*, being significantly 'rescued' by negation ($p = 4 \times 10^{-8}$), it is not fully so, reaching only a mean rating of 3.07. This suggests that it is the item itself, and not an intervention effect, effecting such low ratings for this item. However, these results make it difficult to identify French squarely as an intervention or a non-intervention language.

7 General discussion

Experiment 1 established, as a baseline for comparison, that there is no intervention effect of factivity in English for the weak NPI *ever*. In Italian, by contrast, there appears to be a small intervention effect for the weak NPI *mai*, at least for the two predicates tested in Experiment 2. The situation in French is more subtle, with a less pronounced intervention effect for *le moindre*. (A reviewer notes that this French's bipartite negation structure likely sets it apart from English and Italian.) Conclusions cannot be readily drawn about Italian *affatto* or French *quelque ... que ce soit*, which nearly failed to be licensed at all. As mentioned in Section 6.4, this may be due simply to the fact that *que ce soit* is not a common item in spoken French.

An interesting ancillary finding is that, in sentences with no NPI, there is a highly significant

penalty of negation in English, but not in Italian or French. As mentioned in Section 4.4, there is a well-known processing cost of negation, which may degrade acceptability judgments—however, if this is the cause, we should expect to see it roughly equally in all languages tested. Another possibility is that the *absence* of an NPI deprived the sentences of a certain definitiveness that the sentences with NPIs were perceived to have (cf. the well-known 'domain-widening' effect of NPIs noted by Kadmon & Landman (1993)), and this made the former pragmatically odd. The reason may simply be that the English negative sentences were phrased in a way that participants found slightly less pragmatically felicitous, perhaps because the way information was denied was thought inappropriate, or the repetition of information in the second sentence was awkward (though this should have arisen in the affirmative responses as well). (A reviewer points out that cross-linguistic or cross-cultural pragmatic differences concerning blunt denials may explain the presence of this effect only in English.) Nevertheless, negative sentences with NPIs were not so degraded that there was not a sufficiently significant rescuing effect of negation on the NPI to establish their acceptability under all predicates tested.

Assuming a presuppositional semantics for factivity, the existence of factive intervention argue in favor of Homer's (2008, 2009) account on which presuppositions can be NPI interveners. The effect of factivity could be further tested by cross-linguistically testing NPI licensing in sentences such as (14), in which the factive inference is cancelled. This would require testing both acceptability judgments and intuitions as to whether the inference had indeed been cancelled. However, this would not rule out a non-presuppositional analysis of factivity, such as Romoli's (2012), and would therefore not be a *coup de grace* for the Strawson-entailment theory, as Homer (2009) alleges.

While the presence of the effect has been established, its cause is still unknown. Chierchia (2013) postulates that Romance factives such as 'be surprised', like universals (see Section 2.2) are high scalar items, with alternatives that they entail (say, 'know'—if you are surprised that *p*, you must know that *p*). In this case, intervention again occurs because an UE implicature is generated. Chierchia later hypothesizes (2015) that the difference may lie in the semantics of the embedded clause: whereas an English factive asserts the proposition *p* in its complement, an Italian factive asserts that there *exists* such a proposition which is true—this creates an UE environment in which an NPI is never licensed. We know, however, of no independent data which would corroborate this account, and the difference between these languages may well lie in the syntax or semantics of the factive predicates, as even basic categories often display highly disparate behavior between languages (Haspelmath, 2010), or even in the NPIs which, as mentioned in Section 2.1, display considerable variation in licensing conditions even within languages.

There is clearly further work to be done: to determine conclusively if the dip in acceptability in Italian factives is robust; whether an intervention effect can ultimately be gleaned in French; whether there is intervention only for some NPIs, only for some factives, or only under certain conditions; and whence such an intervention effect would arise in the syntax/semantics of the sentence.

8   Conclusion

In this paper, we have produced quantitative evidence providing a first indication that factive predicates intervene in NPI licensing in Italian, and possibly also in French, though not in English. Much further work is left to be done, including testing different predicates, possibly in other syntactic configurations, and testing other kinds of NPI, which are known to behave quite differently even within a language. We leave unresolved the exact nature of factives in French, as well as the theoretical question of what precisely differs between languages to create this effect—if it is presuppositional, it would call into question the theoretical foundation of an important theory of NPIs.

References

Abusch, D. (2002). Lexical alternatives as a source of pragmatic presuppositions. In B. Jackson (Ed.), *Proceedings of SALT, Vol. 12* (pp. 1–19). Ithaca, NY: Cornell Linguistics Circle, CLC Publications.

Abusch, D. (2010). Presupposition triggering from alternatives. *Journal of Semantics*, *27*(1), 37–80.

Ahn, D. (2016). NPI intervention of *too*. In N. Bade, P. Berezovskaya & A. Schöller (Eds.), *Proceedings of Sinn und Bedeutung*, Vol. 20, (pp. 80–90). Tübingen: University of Tübingen.

Baker, C. L. (1970). Double negatives. *Linguistic Inquiry*, *1*(2), 169–186.

Barker, C. (2018). Negative polarity as scope marking. *Linguistics and Philosophy*, *41*(5), 483–510.

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

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint*, 255–278. arXiv:1506.04967.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Chemla, E., Homer, V., & Rothschild, D. (2011). Modularity and intuitions in formal semantics: The case of polarity items. *Linguistics and Philosophy*, *34*(6), 537–570.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and Beyond*, *The Cartography of Syntactic Structures*, Vol. 3, Chapter 2, (pp. 39–103). Oxford: Oxford University Press.

Chierchia, G. (2013). *Logic in Grammar: Polarity, Free Choice, and Intervention*. Oxford: Oxford University Press.

Chierchia, G. (2015). Factivity meets polarity. Linguistics colloquium talk, University of California, Berkeley, CA.

Denić, M., Chemla, E., & Tieu, L. (2018). Intervention effects in NPI licensing: A quantitative assessment of the scalar implicature explanation. *Glossa*, *3*(1), 1–27.

Djärv, K., Zehr, J., & Schwarz, F. (to appear). Cognitive vs. emotive factives: An experimental differentiation. In R. Truswell, C. Cummins, C. Heycock, B. Rabern & H. Rohde (Eds.), *Proceedings of Sinn und Bedeutung*, Vol. 21, Edinburgh.

Drenhaus, H., Frisch, S., & Saddy, D. (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. 145–163). Berlin: Mouton de Gruyter.

Erlewine, M. Y., & Kotek, H. (2016). A streamlined approach to online linguistic surveys. *Natural Language & Linguistic Theory*, *34*(2), 481–495.

Giannakidou, A. (Ed.). (1998). *Polarity Sensitivity as (Non)Veridical Dependency*. Amsterdam: John Benjamins.

Gonzalez, A., & Moracchini, S. (2016). A morpho-semantic decomposition of French *le moindre* into *even* + superlative. Paper presented at *Linguistics Symposium on Romance Languages* 46.

Gough, P. B. (1965). Grammatical transformations and speed of understanding. *Journal of Verbal Learning and Verbal Behavior*, *4*(2), 107–111.

Haspelmath, M. (2010). Comparative concepts and descriptive categories in crosslinguistic studies. *Language*, *86*(3), 663–687.

Homer, V. (2008). Disruption of NPI licensing: The case of presuppositions. In T. Friedman & S. Ito (Eds.), *Proceedings of SALT*, (pp. 429–446). Ithaca, NY: Cornell Linguistics Circle.

Homer, V. (2009). Presuppositions can be disruptors too: A case against Strawson-entailment. In N. Abner & J. Bishop (Eds.), *Proceedings of the 27th West Coast Conference on Formal Linguistics*, (pp. 220–228). Somerville, MA. Cascadilla Proceedings Project.

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

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

Kadmon, N., & Landman, F. (1993). *Any. Linguistics and Philosophy*, *16*(4), 353–422.

Karttunen, L. (1971). Some observations on factivity. *Papers in Linguistics*, *4*(1), 55–69.

Krifka, M. (1995). The semantics and pragmatics of polarity items. *Linguistic Analysis*, *25*(3), 209–257.

Ladusaw, W. A. (1979). *Polarity Sensitivity as Inherent Scope Relations*. PhD thesis, University of Texas, Austin, TX.

Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software*, *69*(1), 1–33.

Linebarger, M. (1980). *The Grammar of Negative Polarity*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Parker, D., & Phillips, C. (2016). Negative polarity illusions and the format of hierarchical encodings in memory. *Cognition*, *157*, 321–339.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Richter, F., & Radó, J. (2014). Negative polarity in German: Some experimental results. *Journal of Semantics*, *31*(1), 43–65.

Romoli, J. (2012). *Soft but Strong: Neg-Raising, Soft Triggers, and Exhaustification*. PhD thesis, Harvard University, Cambridge, MA.

Saddy, D., Drenhaus, H., & Frisch, S. (2004). Processing polarity items: Contrastive licensing costs. *Brain and Language*, *90*(1–3), 495–502.

Schnoebelen, T., & Kuperman, V. (2010). Using Amazon Mechanical Turk for linguistic research. *Psihologija*, *43*(4), 441–464.

Shao, J., & Neville, H. (1998). Analyzing semantic processing using event-related brain potentials. In *The Newsletter of the Center for Research in Language, Vol. 11*. La Jolla, CA: Center for Research in Language.

Slobin, D. I. (1966). Grammatical transformations and sentence comprehension in childhood and adulthood. *Journal of Verbal Learning and Verbal Behavior*, *5*(3), 219–227.

Sprouse, J. (2011). A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods*, *43*(1), 155–167.

Trabasso, T., Rollins, H., & Shaughnessy, E. (1971). Storage and verification stages in processing concepts. *Cognitive Psychology*, *2*(3), 239–289.

von Fintel, K. (1999). NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics*, *16*(2), 97–148.

Wason, P. C. (1959). The processing of positive and negative information. *Quarterly Journal of Experimental Psychology*, *11*(2), 92–107.

Wason, P. C. (1961). Response to affirmative and negative binary statements. *British Journal of Psychology*, *52*(2), 133–142.

Wason, P. C. (1965). The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior*, *4*(1), 7–11.

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

Zwarts, F. (1998). Three types of polarity. In F. Hamm & E. Hinrichs (Eds.), *Plurality and Quantification*, (pp. 177–238). Dordrecht: Springer.