AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: Learning, Memory, and Cognition

ISSN: 0278-7393

2022, Vol. 48, No. 1, 122–141
https://doi.org/10.1037/xlm0001059

Negative Sentences Exhibit a Sustained Effect in Delayed Verification Tasks

Galit Agmon[1], Yonatan Loewenstein[1, 2, 3, 4], and Yosef Grodzinsky[1, 5]

[1] The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem
[2] Department of Neurobiology, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem
[3] Department of Cognitive Sciences, The Hebrew University of Jerusalem
[4] The Federmann Center for the Study of Rationality, The Hebrew University of Jerusalem
[5] Forschungszentrum Jülich, Institute of Neuroscience and Medicine (INM-1)

Negated sentences are known to be more cognitively taxing than positive ones (i.e., *polarity effect*). We present evidence that two factors contribute to the polarity effect in verification tasks: processing the sentence and verifying its truth value. To quantify the relative contribution of each, we used a delayed verification task. The results show that even when participants are given a considerable amount of time for processing the sentence prior to verification, the polarity effect is not entirely eliminated. We suggest that this sustained effect stems from a retained negation-containing representation in working memory.

Keywords: sentence processing, sentence representation, negation, quantifiers, verification

Consider hearing the sentence "The square is blue." Your task is to press a button to determine the sentence's truth value against a picture that follows it (see Figure 1a). The time it takes you to perform this task reflects three factors: (a) the transformation from a sound representation into a meaning representation (*processing cost*), (b) the comparison of the meaning against the picture (*verification cost*), and (c) the motor execution of the decision (*motor cost*). Consider now the same task, but with the negative sentence "The square is not blue." You would probably find this negative sentence more difficult to verify, and indeed, it is well-known that sentences containing a negation evoke longer reaction times (RTs) in verification tasks compared with sentences that do not (Carpenter & Just, 1975; Clark & Chase, 1972, 1974; Gough, 1965; Just & Carpenter, 1971; Wason, 1961). We term this phenomenon the *polarity effect* ($\Delta RT = RT_{negative} - RT_{positive} > 0$). Where does the polarity effect stem from? If we assume a constant RT of the motor cost, then the polarity effect must stem either from differences in the processing cost or from differences in the verification cost.

Early approaches attribute the polarity effect to the verification cost, namely, the additional effort that negation imposes on the comparison process (Carpenter & Just, 1975; Clark & Chase, 1972). On the other hand, more recent approaches, based on lexical decision, recognition and naming paradigms, attribute the polarity effect mainly to the processing cost, namely, the time it takes to integrate the negation into the meaning of the sentence (Hasson & Glucksberg, 2006; Kaup et al., 2006, 2005; Kaup & Zwaan, 2003). One way to test which of these two costs is dominant is by using an important observation from the processing camp, who find that processing usually takes up to 1500 ms. Thus, if we perform an experiment that includes 1,500 ms for processing before the probing task, and if we still find a polarity effect, then we could attribute that effect to verification costs alone. Therefore, in this paper we design a protocol with a delay to separate the verification and processing costs in the polarity effect.

Few published studies measured the polarity effect in a delay paradigm. Whereas Gough (1966) found a significant effect with a delay, Tversky (1975) found no delayed polarity effect, though this may be attributable to methodological differences (see Experiment 1, Discussion section). The discrepancy between these earlier studies and some limitations that we discuss subsequently, led us to reopen the question.

In the current study, we provide evidence that the processing cost and verification cost both contribute to the polarity effect and assess the relative contribution of each. We use a delayed verification task and manipulate the time between the sentence and the picture to test for two things: a decrease of the polarity effect with longer delays and a delayed polarity effect. By introducing a delay (of at least 1,500 ms) between the sentence presentation and the picture verification task, we give the participant enough time to process the sentence. Thus, if the polarity effect decreases when adding a delay, it is evidence that processing has occurred. The

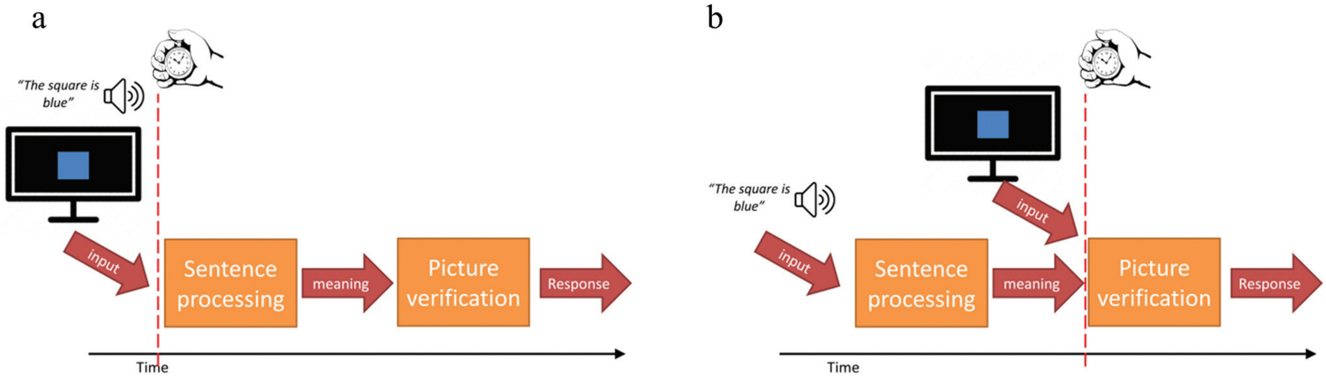
Galit Agmon https://orcid.org/0000-0001-9588-5953
Yonatan Loewenstein https://orcid.org/0000-0003-2577-2317
Yosef Grodzinsky https://orcid.org/0000-0002-9912-7196

This work was supported by grants from the Israel Science Foundation (ISF 2093/16 and ISF 757/16) and funding provided by the Gatsby Charitable Foundation.

Data and code are available on the Open Science Framework (https://osf.io/t56bx/files/).

Correspondence concerning this article should be addressed to Galit Agmon, who is now at Penn Frontotemporal Degeneration Center, Department of Neurology, University of Pennsylvania, West Gates, 3400 Spruce Street, Philadelphia, PA 19104, United States. Email: galit.agmon@mail.huji.ac.il

Figure 1
Illustration of a Delayed Verification Task



Note. The delayed verification task can separate verification cost from processing cost. Response time can be measured only when there is a picture to verify. Hence, by manipulating the time of the picture input (a blue square on a computer screen) relative to the auditory input ("the square is blue"), we can control the starting time of the response (dashed red line). (a) When the picture appears immediately after the auditory input, RT reflects the sum of processing and verification. (b) When the picture appears enough time after auditory input, the processing sequence presumably has been concluded, and RT should reflect only verification. See the online article for the color version of this figure.

magnitude of the decrease would then reflect the portion of the polarity effect attributed to processing costs. Additionally, if we find a polarity effect even after a sufficiently long delay, then we know that it is due to verification costs and not processing costs. The magnitude of the delayed polarity effect relative to the one without a delay would reflect the portion of the polarity effect attributed to verification costs. Compared with most past experiments, our design uses aurally presented sentences, as reading a sentence might interfere with its visual verification compared with listening to the same sentence, and artificially increase the verification cost (Brooks, 1967, 1970; Claus & Kelter, 2009; Eddy & Glass, 1981). Importantly, our design controls for two key issues which seldom have been addressed in past experiments. First, we equalize positive and negative sentences for number of words. Second, we control for informativity, which is the number of true scenarios that match each sentence type (see the Method section of Experiment 1). Moreover, we extend the scope of research on the polarity effect to test negative quantifiers as well as sentential negation.[1]

Although our delayed verification paradigm allows us to explore the polarity effect specifically, on a general level it also allows us to research sentence representation in working memory (WM). An increased verification cost of negative sentences (i.e., a polarity effect which is not eliminated by additional time for processing) is an indicator of retained differences between positive and negative sentences. To understand the relationship between a verification cost and WM representation, let us consider what participants do once they have acquired the sentence's meaning. Participants are required to keep the meaning in WM for when the picture appears, and then to compare the represented meaning with the picture to decide whether or not they match. Either of these steps could be more costly in the presence of negation and thus increase verification RTs of negative sentences. Difficulty in accessing this WM representation could cause a delayed onset of the verification procedure. Alternatively, the negative representation itself might cause a longer verification procedure of the picture against the sentence. Either way, the increased verification cost of negative sentences would be a result of a more complex representation of negative sentences in WM. Hence, our rich set of results allows us to explore interesting questions on how sentences are represented after processing and prior to verification (see the General Discussion).

The Processing Cost of Negation

First, let us review some of the main theories explaining the polarity effect, focusing on the findings relevant to our purpose and their predictions for the proposed delayed verification task. The polarity effect is originally based on the observation that sentential negation adds a considerable processing difficulty to the sentence. This processing difficulty, in part, has been attributed to the pragmatic licensing of sentential negation. Sentential negation serves to deny a statement that is part of the common ground of the speakers and is salient in the discourse (Clark, 1976; Givón, 1978; Glenberg et al., 1999; Horn, 1989; Roberts, 2012; Tian et al., 2010, 2016). Therefore, sentential negation must be pragmatically licensed in the sense that the context must support the uttering of a negative sentence, which is usually underinformative compared with a positive sentence. For example, the negative sentence "The square is not blue" provides little information of the actual color of the square, and it is odd to say such a sentence with no context;

[1] A different approach to decompose the polarity effect into a processing component and a verification component is found in Macleod et al. (1978) and in Mathews et al. (1980; see also Trabasso et al., 1971 for a similar protocol). These studies used a self-paced verification paradigm, in which RTs are measured separately for reading the sentence and for verifying it against a picture, based on the participants' explicit report via two button presses. These studies found a polarity effect in both RT components, but that of the verification component disappears when examining only a subgroup of participants who seemed to have formed a pictorial mental representation of the sentence's meaning. Whereas these results are consistent with ours, they suffer from the same limitations outlined above. We discuss the results of these studies in light of ours in the Discussion to Experiment 1 and in the General Discussion.

it is sensible to say such a sentence only if the statement "The square is blue" had been considered, thus addressing some contextual question under discussion. Therefore, it has been suggested that the longer RTs associated with negative sentences, as found in lab settings, are (at least partially) a result of using negation out of context, with no salient question under discussion that licenses it (Dale & Duran, 2011; Glenberg et al., 1999; Orenes et al., 2016; Tian & Breheny, 2016; Wason, 1965).

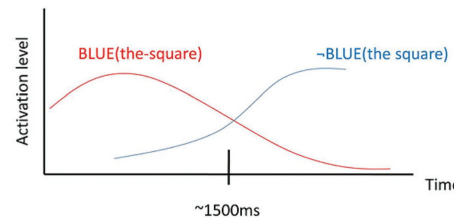
Context, however, cannot be the only relevant parameter in explaining why negative sentences take longer to process than positive ones. Context cannot explain the well-known interaction of polarity with truth value: The polarity effect is smaller for false sentences (i.e., mismatch between sentence and picture) than it is for true sentences. This highly replicable finding suggests that truth value also is an important factor (Agmon et al., 2019; Carpenter & Just, 1975; Just & Carpenter, 1971; Krueger, 1972; MacDonald et al., 1992; Mayo et al., 2004; Tucker et al., 2018).

One way to reconcile context and truth value in their impact on the polarity effect is through the concept of mental models. A mental model (often referred to as *mental simulation* or *representation*) is a small-scale representation of the real world that is created in one's mind in order to reason and to interact with the real world. According to these theories, a mental model of a sentence's meaning activates similar cognitive circuits to those involved in the actual experience described by the sentence (Barsalou, 2003; Johnson-Laird, 1980; Kaup et al., 2007; Stanfield & Zwaan, 2001). Thus, a mental model of the sentence "The square is blue" simply would be one mental image of a blue square. In the processing of negative sentences, however, two mental models are required, one of the negated state and another of the actual state. For example, in processing "The square is not blue," a mental model of a blue square is constructed in addition to a mental model of a nonblue square. Reaction to negative sentences takes longer, since the simulation of the actual state is obtained only later in the processing sequence, as shown both in priming and in even-related potential (ERP) studies (Dudschig & Kaup, 2018; Fischler et al., 1983; Hasson & Glucksberg, 2006; Lüdtke et al., 2008). Such studies estimate that integrating the negation into the meaning of the sentence can take up to 1,500 ms. In our example, the representation of a blue square lasts for around 1,500 ms before it is inhibited by the nonblue square's representation (see Figure 2). Importantly, the mental models approach also explains the role of context. A supportive context for negation means that there is a prominent question under discussion that primes the model of a nonblue square, which in turn facilitates the simulation of the actual state (Kaup et al., 2006, 2007; Tian & Breheny, 2016; Tian et al., 2016).

How, then, do mental models explain the interaction of polarity with truth value? Because mental models involve similar cognitive processes as actual imagery, a mental model can (in principle) prime a picture. Thus, if a negative sentence first triggers a mental model of the positive sentence, then the response to pictures that match this positive representation is facilitated (notice that positive representations are false for negative sentences.) Hence, verification of false pictures is faster for negative sentences, which explains the well-known interaction of the polarity effect with truth value (Kaup et al., 2005). By this logic, after the participant waits the required time to comprehend the sentence and inhibit the positive representation, we expect that this interaction should disappear.

Figure 2
Prediction of the Mental Models Theory

"The square is not blue"



Note. A schema of the predicted temporal dynamics in processing negative sentences. After the sentence "The square is not blue" is presented, there is initially an activation of the false mental model, in this case of a blue square. The integration of negation brings about an activation of the true mental model of a nonblue square at 1,500 ms, whereas the false mental model is inhibited. This schema predict no polarity effect 1,500 ms after sentence offset. This schema also predicts that effects stemming from the priming of the false model, such as the Polarity × Truth Value interaction, will diminish if measured 1,500 ms after sentence offset. See the online article for the color version of this figure.

Based on the preceding theoretical considerations, after processing negation and reaching the final representation there should be no difference between RTs of sentences which are identical in meaning but differ in polarity (if one assumes that the two meanings are equivalent, which is not a necessary assumption; see the General Discussion). For example, in a context of only two colors, blue and yellow, "The square is not blue" simply means that "The square is yellow." Finding a delayed effect therefore would be evidence of representations which are structure-dependent, as the representation in WM retains the difference between positive and negative sentences.

The polarity effect can be found not only overtly in sentential negation, but also covertly in negative quantifiers (as are *few* and *less than half* relative to their positive counterparts *many* and *more than half*). Negative quantifiers have been suggested to contain an implicit negation in their representation. For example, *few* is understood implicitly as *not many* (Hackl, 2000, p. 126; Heim, 2006; Jacobs, 1991; Penka, 2011; Rullmann, 1995). This hidden negation can be detected by various linguistic tests (Chierchia, 2013; Klima, 1964; Penka, 2011). Indeed, as with sentential negation, negative quantifiers have been shown to evoke cognitive effects relative to their positive counterparts: prolonged response times in verification tasks (Agmon et al., 2019; Deschamps et al., 2015; Just & Carpenter, 1971) and more complex reference patterns in discourse (Moxey & Sanford, 1986; Sanford et al., 2007). The negative quantifier *few* has also been shown to evoke ERP patterns similar to those of sentential negation (Xiang et al., 2016). Therefore, we expand our investigation to include not only sentential negation (which was the focus of most previous studies) but also negative quantifiers.

In the current study, we tested whether a polarity effect can be found when delaying the verification task for sentential negation, as in Experiment 1, and for quantificational negation, as in Experiment 2. Our main results are that (a) the polarity effect does decrease when introducing a delay, which is evidence for an increased processing cost of negation; (b) the polarity effect is not eliminated when introducing a delay, which is evidence for an increased verification cost of negation; and (c) the increased verification cost of negation exists even when the meaning of a negative sentence is identical to the meaning of a positive sentence, which is evidence that the verification process is a function of the linguistic structure. We discuss the repercussions of these results in the General Discussion.

Experiment 1

Method

Our primary goal in this experiment was to probe picture verification in negated and non-negated sentences by manipulating the interstimulus interval (ISI) to test whether an increased delay between a sentence and the picture eliminates the polarity effect, or at least decreases it. The rationale was that a decay of the polarity effect with longer ISI is indicative of complexity that stems from the processing sequence, and that a polarity effect found after delaying the verification task is indicative of complexity of the verification process. We thus introduced a verification paradigm with the polarity factor (positive and negative sentences), where the ISI factor regarded the appearance of the picture against which the sentence should be verified, which was either 100 ms after sentence offset or delayed by additional 1,500 ms (see Table 1).

Table 1
The Polarity, ISI and Truth Value Factors

	ISI	
Polarity	100 ms	1,600 ms
Positive	True/False	True/False
Negative	True/False	True/False

Note. ISI = interstimulus interval (time between sentence offset and picture onset). Persistence of the polarity effect implies a simple effect of polarity at the 1,600 ms ISI. Diminishment of the polarity effect means a Polarity × ISI interaction.

A second goal of this experiment was to control for informativity. For example, "The square is not blue" can describe multiple scenarios (e.g., "The square is yellow," "The square is red"), hence it is less informative than "The square is blue," which describes only one. This might affect the processing difficulty of negative sentences (Khemlani et al., 2012). In addition, when negative sentences describe only one scenario, participants could use a recoding strategy; that is, they could rephrase "The square is not blue" as "The square is red," either verbally or pictorially, thus predicting a smaller polarity effect given additional processing time (Carpenter & Just, 1975, p. 66; cf. Clark & Chase, 1972, p. 500; Tversky, 1975). We therefore manipulated the number of possible scenarios that can match a negative sentence via a number of colors factor (a similar manipulation was done by Trabasso et al., 1971): one block had two possible colors (blue, red), and a second block had three (blue, red, yellow). Thus, there is only one alternative for negative sentences in the two-color block (if a square is not blue, then it is red) and two alternatives in the three-color block (if a square is not blue, then it is either red or yellow). If the number of alternative scenarios to be considered is relevant, then a larger polarity effect is predicted in the three-color block than in the two-color block (i.e., a Polarity × Number of Colors interaction). Participants were instructed of the possible colors.

Materials

We used a novel control factor to resolve the well-known confound that negating a sentence requires the addition of a word, whose processing may incur morphophonological, lexical, or syntactic cost. Some, but not all, past experiments have added a short word to the positive sentence as a morphophonological/lexical control, or normalized RTs based on sentence length. The choice of the control (that varies across experiments) is obviously critical due to its potential semantic import, and yet despite its importance, this issue is seldom discussed. To control for the additional word in negative sentences, we harness the Hebrew language.

Hebrew nominal sentences can come with a copula, which can be either negative or positive. The use of a copula in nominal sentence does not affect its meaning. Hence, comparing positive and negative copular sentences in Hebrew enables us to balance the sentences with respect to the number of words in each sentence without any arbitrary change in meaning. We used nominal sentences in which the subject denoted shape (square, circle, star, triangle) and the predicate denoted color (blue, red, yellow). As Table 2 shows, *lo* negates a sentence without a copula, whereas *'eyno* does that for a copular sentence.[2] This creates a 2 × 2 matrix, with polarity and copula as two-level factors. This enables a copula factor for testing whether an unequal number of words is a confounding factor (despite a minimal difference in the number of syllables: the negative copula *'eyno* has one more syllable compared with the positive copula *hu*). Although we are not aware of any difference in meaning caused by adding a copula, it is important to note that any such difference is orthogonal to the other factors in our design.

The second control factor was truth value—sentences were either true or false (equal number of tokens each), so that the task is carried out within a balanced setup. To conclude, our design was a 2 × 2 × 2 × 2 × 2 design with the following factors: polarity (negative/positive), ISI (100 ms/1,600 ms), number of colors (two-color/three-color), copula (copula/null), and Truth Value (true/false). In the two-color block, participants had to respond to 256 trials: 2 Polarity × 2 ISI × 2 Copula × 2 Truth Value × 16 Repetitions (repetitions were counterbalanced for the two colors and

[2] *hu* is an optional copula (masculine agreement) in nominal sentences besides in equative sentences where it is obligatory (Doron, 1983). Negating a nominal sentence can be done by adding sentential negation (*lo*) or by using the negative copula *'eyno*, which is composed of the sentential negation '*eyn* and an agreeing clitic *-o*. Following Greenberg (2008) who studied the distribution of copular *hu*, one can be convinced that the distribution of '*eyno* is the same, and therefore that *'eyno* is the negation of *hu* (equivalent to *hu lo*): in equative sentence where *hu* is obligatory, '*eyno* (or *hu lo*) is obligatory as the negation (the teacher is Danny: *ha-more* *NULL/hu Danny*; the teacher isn't Danny:* ha-more *lo/'eyno Danny). Whenever *hu* is optional, '*eyno* (or *hu lo*) is optional as the negation (Danny is a teacher: *Danny NULL/hu more*; Danny isn't a teacher: *Danny lo/'eyno more*). Whenever *hu* is ungrammatical, *'eyno* (or *hu lo*) is ungrammatical as the negation (what Danny was is dangerous to himself: *ma še-Danny haya* *hu/*'eyno mesukan le-'acmo).

Table 2
The Copula and Polarity Factors

Copula	Polarity: Positive (*"The square is blue"*)	Negative (*"The square is not blue"*)
Without a copula	*ha-ribu'a kaxol* the-square blue	*ha-ribu'a lo kaxol* the-square NEG blue
With a copula	*ha-ribu'a hu kaxol* the-square COP blue	*ha-ribu'a 'eyno kaxol* the-square COP.NEG blue

Note. The first line in each cell shows the transliteration of the Hebrew example sentences – "The square is blue" (Positive column) and "The square is not blue" (Negative column). The second line in each cell shows the corresponding word-by-word translation and gloss. Notice that with a copula, negation does not add an extra word to the sentence (second row). Gloss abbreviations: NEG = negation, COP = copula, COP.NEG = negative copula.

four shapes). In the three-color block, participants had to respond to 288 trials: 2 Polarity × 2 ISI × 2 Copula × 2 Truth Value × 18 Repetitions (repetitions were counterbalanced for the three colors and four shapes).

All sentences were recorded in Hebrew, by a female native Hebrew speaker, and later processed in Audacity (Version 2.0.5; Audacity Team, 2015) software to equalize them in term of their average pitch, average amplitude and duration. See Appendix A for the full list of sentences.

Procedure

Experiment was run using Presentation software (Version 17.0, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com). In each trial, participants had to decide whether a sentence they heard on earphones correctly described a picture that later appeared on the screen. The sentence always indicated that some shape was or was not in a certain color (see Appendix A). ISI was manipulated—in one half of the trials there was an interval of 100 ms between the sentence offset and the picture onset, and in the other half the appearance of the picture was delayed by additional 1,500 ms (i.e., ISI = 1,600 ms). The procedure and timing of a single trial are summarized in Figure 3.

Each trial started with a fixation cross on a black screen. After 300 ms, the participant heard a sentence while the fixation cross was still on the screen. Sentence duration was 1,700 ms. The fixation cross disappeared 50 ms before the sentence offset, and a picture appeared either after 100 ms or after 1,600 ms. The picture stayed on the screen until the participant decided whether the sentence correctly described the picture (true/false) by pressing one of two possible buttons on the keyboard (the *Z* key or the "/" key, counterbalanced between participants). Participants were encouraged to respond as fast and as accurately as they could. Once verification was made, participants saw on the screen a smiling face if they answered correctly or a sad face if they answered incorrectly. The face stayed on the screen for 500 ms, then the next trial started.

Each individual participated in both a two-color block and a three-color block, with a short practice session preceding each block. The order of blocks was randomized between participants. Participants were informed prior to each block on the possible colors for that block. The two-color block consisted of four runs (64 trials each) and the three-color block consisted of three runs (96 trials each). Participants could take a short break between runs if they wished to. Each block was counterbalanced for Polarity (positive/negative), Truth value (true/false) and the color and shapes mentioned in the sentence or depicted in the picture. Participants were informed that the picture could either appear immediately or after a short moment, to prevent them from stressing out about possible technical problems in the experiment.

Participants

We decided to recruit 30 to 40 participants for this study, which is above the common number used in past experiments in which a polarity effect was found. 38 undergraduate students from the Hebrew University, native Hebrew speakers, participated in the experiment for either payment or credit, after signing an informed consent approved by the Hebrew University Research Ethics Committee. Mean age of participants was 24 ($SD = \pm 2$), 34 were right-handed, and 32 were women.

Analysis

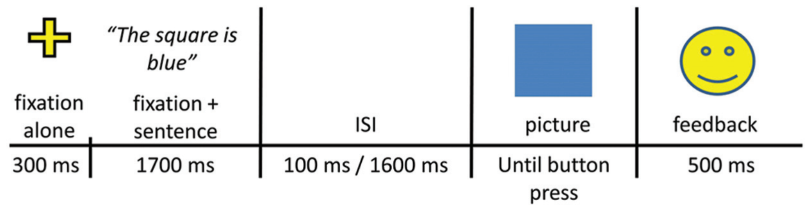
We fitted a linear mixed effects regression, using R *lme4* package (Bates et al., 2015). The dependent variable was the RT of correct responses, log-transformed in order to make its skewed distribution closer to normal (Whelan, 2008). As fixed effects, we entered Polarity, ISI, Truth Value and all their interactions, Copula, Polarity × Copula, Number of Colors, Polarity × Number of Colors, the order of blocks (Order) and its interaction with polarity (Polarity × Order). All factors were sum coded. Fixed effects were chosen based on our research questions and theoretical framework, which we repeat here for convenience: Polarity, ISI and their interaction were included to test our main hypothesis regarding the decay of the polarity effect with additional time allotted for processing.[3] The three-way interaction of Polarity × Truth Value × ISI was included as a fixed factor due to the relevance of the time dynamics of the Polarity × Truth Value interaction: Mental model theories predict this interaction to disappear when delaying the verification task. Polarity × Copula (and the main effects that constitute it) was included to test to what degree an unequal number of words is a confounding factor. Polarity × Number of Colors (and the main effects that constitute it) were included to control for the sentence informativity. Order of blocks was included to control for potential effects that relate to the specific order in which participants performed each block. The effects involving number of colors were not expected to interact with the time into the block (e.g., due to the time it takes to realize the number of possible colors), as participants were informed prior the experiment on the number of possible scenarios.[4]

The random structure was chosen based on stepwise model comparison, using the step function from the R *lmerTest* package

[3] To verify the soundness of our results, we ran the model again with the full factorial design and found the results did not change. Some minor differences between the full model and our model are reported in Appendix B.

[4] We nevertheless tested the effect of time on the Number of Colors in various ways, and our results did not change with regard to the effects of interest. For the full analysis, see the data and code on the Open Science Framework (https://osf.io/t56bx/?view_only=ba9c8f6648d44f299b796f4e62a1d8e5).

Figure 3
Trial Structure of Experiment 1



Note. ISI = interstimulus interval. See the online article for the color version of this figure.

(Kuznetsova et al., 2017) and by removing variables with low variability in order for the model to converge. This procedure resulted in the inclusion of random intercepts and random slopes for Polarity, ISI, Truth Value, Number of Colors and the interactions of Polarity × ISI and Polarity × Truth Value, adjusted by participants. To validate our results, p values of the fixed effects were obtained in two different ways: the first, using the Satterthwaite (1946) approximation of degrees of freedom, which is implemented in R *lmerTest* package (Kuznetsova et al., 2017). The second was using LRTs of the full model with the effect in question against the model without the effect in question, using R *afex* package (Henrik et al., 2019). The two methods resulted in similar p values. For ease of reading, we report here only t statistics and p values from the Satterthwaite approximation. For a comparison with the LRT, see Table B1 in Appendix B.

Results

Accuracy was overall high across participants (average was 96%, with a standard deviation of 3%), and it significantly improved with the longer delay (from 95.5% to 97%; $z = 2.1$, $p = .03$ in a generalized mixed effects model with ISI and Polarity as independent variables). For the RT analysis, only trials with correct responses were considered (thus removing 3.8% of the data).

Our multifactorial mixed effects regression resulted in a significant main effect of Polarity ($\beta = .1$, $t = 11$, $p < .0001$), ISI ($\beta = –.06$, $t = –14.4$, $p < .0001$) and crucially—a significant Polarity × ISI interaction ($\beta = –.02$, $t = –6.5$, $p < .0001$), indicating that the polarity effect is affected by the ISI and is significantly smaller when delaying the verification task. Indeed, in the short 100 ISI conditions, mean difference between negative and positive sentences was 265 ± 28 ms ($M \pm SE$ of within-participant overall polarity effects), and this difference decreased in the 1,600 ISI to 191 ms ± 26 ms. Crucially, the polarity effect did not vanish after an ISI of 1,600 ms. A significant polarity effect was found in all conditions, besides in the three-color copula false condition, as revealed by an analysis of simple effects ($p < .05$, linear mixed effects regressions, with log-RT as the dependent variable, Polarity as the predictor, random slopes and random intercepts adjusted by participants; Bonferroni corrected for multiple comparisons). Full summary of the simple effects is provided in Table B2 in Appendix B.

The multifactorial mixed-effects regression also resulted in a main effect of Number of Colors ($\beta = .03$, $t = 2.7$, $p = .01$), as responses in the three-color block were longer by 42 ms ±19 ms on average relative to the two-color block ($RT_{three\text{-}colors} = 835 \pm 42$ ms, $RT_{two\text{-}colors} = 794$ ms ±36 ms). However, there was no significant difference in this effect between the negative and the positive condition (no significant Polarity × Number of Colors interaction: $\beta = .004$, $t = 1.6$, $p = .1$). A main effect of Copula was also found ($\beta = .007$, $t = 2.9$, $p = .004$), as the presence of a copula added on average about 15 ms ±6 ms to RT. Examining the means, this effect of copula resulted mainly from the difference in the positive condition ($RT_{copula} = 714$ ms ±32 ms, $RT_{null} = 687$ ms ±28 ms), where adding a copula also adds a word to the sentence, whereas in the negative condition the reaction times seemed to be similar whether a copula was present or not ($RT_{copula} = 930$ ms ±49 ms, $RT_{null} = 927$ ms ±50 ms), as also indicated by a two-way Polarity × Copula interaction ($\beta = –.007$, $t = –2.8$, $p = .005$). An effect of Order was found ($\beta = –48$, $t = –3.7$, $p = .0007$); participants' responses were faster in the second block compared with the first block. Participants also improved their processing of negative sentences, as the polarity effect got smaller in the second block of the experiment (Polarity × Order: $\beta = –.01$, $t = –5.1$, $p < .0001$).

The effect of Truth Value was significant ($\beta = .05$, $t = 6.6$, $p < .0001$), as false sentences took longer to respond to. We found a highly significant interaction of Polarity × Truth Value ($\beta = –.06$, $t = –11$, $p < .0001$), indicating an overall smaller polarity effect for false sentences. We examined the temporal dynamics of this interaction: For negative sentences, $RT_{false} < RT_{true}$ with an ISI of 100 ms, and this difference disappears with an ISI of 1,600 ms. For positive sentences, however, $RT_{false} > RT_{true}$ with both ISIs. This was reflected in a significant three-way interaction Polarity × ISI × Truth Value ($\beta = .01$, $t = 3.8$, $p = .0002$). This finding suits the prediction of a two-step mental model processing approach (Kaup et al., 2005): First, the positive is represented, so at short ISIs false negatives are expected to be faster than true negatives. Then, when the negation has already been incorporated, false sentences take longer than true sentences no matter their polarity (long ISI). However, a simple Polarity × Truth Value interaction at the longest ISI level did turn out to be significant in a mixed linear effect regression, with by-participant random intercepts and slopes for Polarity and Truth Value ($\beta = –.05$, $t = –12.8$, $p < .0001$; fixed effects were similar to those used in the full multifactorial model, besides the ISI which is not a variable when testing for this simple effect). This interaction indicates that the polarity effect is smaller for false sentences even beyond the 1,600-ms time window. This interaction stems from the fact that whereas truth value affects positive sentences (true sentences are facilitated), is seems to have almost no impact on negative sentences.

Other results that do not have direct relevance to the discussion appear in Appendix B. See Table B1 for a full list of the effects and statistical details.

Discussion

Our results from Experiment 1 can be summarized in the following points: (a) a main effect of Polarity, replicating previous results; (b) only a small Polarity × Copula interaction, fortifying the argument that the effect is driven mostly by the semantic import of negation and not by an addition of a word; (c) a Polarity × ISI interaction, due to the decay of the polarity effect with longer ISIs; and (d) a simple effect of Polarity for the longest level of ISI.

According to our hypothesis, the Polarity × ISI interaction is indicative of the processing component of the polarity effect (which is delay-sensitive). The time dynamics of the interaction of polarity with truth value supports a processing sequence in which the positive meaning is represented at early stages of processing (the representation of the positive facilitates false negatives at short ISIs, and less so at long ISIs). Additionally, the simple effect of Polarity at ISI = 1,600 ms is indicative of the verification component of the polarity effect (which is delay-independent). This verification component was not shown to be due to the level of informativity (a polarity effect in the two-color condition, and no Polarity × Number of Colors interaction).[5] Note that the additional 1,500-ms delay reduces the polarity effect by 28% (74 ms out of the initial 265 ms difference). Most of the effect persists, suggesting that most of the polarity effect stems from verification costs rather than processing costs.

An immediate concern from Experiment 1 is whether participants are slower than expected in integrating the negation, and that therefore the delayed polarity effect might actually reflect a delayed processing cost instead of a verification cost. In that case, more time should further diminish or even eliminate the polarity effect. To assess the plausibility of this possibility, we can look at RTs at ISI = 100 ms. As shown in Figure 4, processing negative sentences at ISI = 100 ms took less than 1,100 ms. If the delayed effect is due to slow processing, a delay of 1,500 ms should suffice to eliminate processing differences of that magnitude. It is therefore unlikely that the delayed effect reflects a processing cost. However, to address this concern, in Experiment 2 we use longer ISIs.

A second concern is what happens during the delay. Are sentences with equivalent meanings represented in a similar way, detached from their linguistic structure? Some early studies suggested that participants could adopt a strategy of pictorial encoding (Macleod et al., 1978; Mathews et al., 1980; Tversky, 1975).[6] However, such a strategy can be applied only when it is clear from the negative sentence what the corresponding picture is. Moreover, in natural settings, it is rare that a sentence would provide an exact and full description of a picture with all its attributes specified. Therefore, for generalizability, in Experiment 2 we control for the possibility of a pictorial encoding by providing several picture tokens to match each sentence, positive or negative. We also expand our research to explore the processing and verification costs of other types of negation.

Our results match most of the literature on the delayed polarity effect. One exception is Tversky (1975), who found no delayed effect. However, upon careful examination of her methods, we suggest that several factors can explain this discrepancy. Tversky used sentences similar to those of Clark and Chase (1972): She used written sentences of the form "The plus is (*not*) above the star," depictable by two possible pictures (a star below a plus sign, a star above a plus sign), which are easily encoded pictorially, as explained in the preceding text (the same sentences were used by Mathews et al., 1980 and by Macleod et al., 1978). The duration of the delay is also debatable: A short and simple four- to five-word written sentence was displayed for 5 s followed by a 5-s delay, followed by the picture, which remained in view until the participant responded. This extremely long ISI was not constrained in any way. Thus, in the absence of eye-tracking information (or even standard deviations or standard errors of the mean), it is difficult to compare this experiment with its contemporaries, which have used self-paced paradigms or shorter delays, or to compare it with our study, which used aural stimuli.

Experiment 2

In Experiment 1, we found evidence for two sources contributing to the polarity effect: one that decreases, which we interpret as due to the longer processing sequence of the negative sentence, and another that persists, which we interpret as due to the more complex verification of the negative sentence. The question arises whether other types of negation show a delayed polarity effect, indicative of an increased verification cost of negation. Therefore, in Experiment 2, we tested the negative quantifier *fewer* and existential sentential negation. We probed three ISIs to get two time windows in which we could assess the decrease of the polarity effect. To determine how late after the sentence offset we could still measure the polarity effect, we used longer ISIs—up to 6.8 s—expecting that at this point participants would have enough processing time to overcome the relative difficulty imposed by negative sentences, unless this complexity is due in part to the verification of the meaning against the picture.

Method

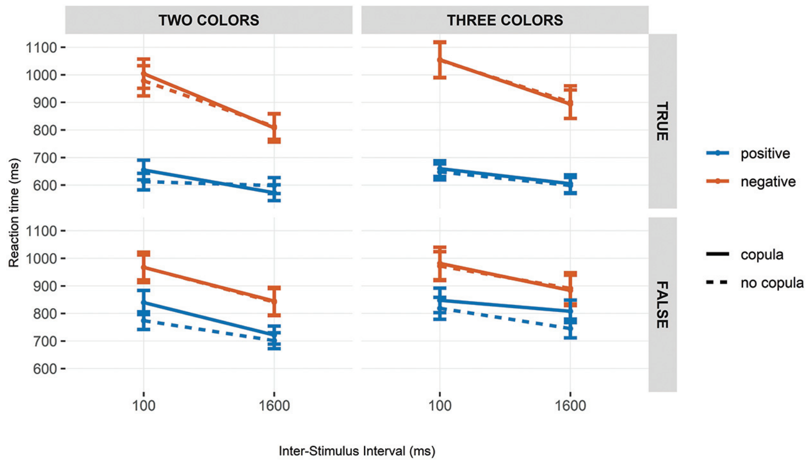
In Experiment 2, we once again harnessed the Hebrew language to make sentential negation equal in number of syllables and morphemes to its positive counterpart. We did this by using the existential negation *'eyn* (=¬∃), which is the only way in Hebrew to negate existential sentences with *yeš* (=∃) (an explicit negation with an existential – **lo yeš* – is ungrammatical). In addition to the polarity of the sentence which was expressed with *yeš* or *'eyn* (Sentential Polarity; Polarity-Sen), we added a factor of quantifier

[5] We also examined this interaction separately for ISI = 100 and ISI = 1,600 and found it to be non-significant in both levels of ISI (t = .06, p = .95 and t = 1.5, p =.15 respectively) in two separate linear mixed effects regressions, with by-participant intercepts and slopes for Polarity, Number of Colors and their interaction. Fixed effects included polarity, truth value, number of colors, copula, Order and the interaction of each with polarity. In addition, we added an analysis to trim the data such that we removed responses that exceeded two standard deviations within each participant (3.9% of data). With such trimming, all results were similar under the log-likelihood ratio test (LRT) approach. However, under the Satterthwaite approach, the Polarity × Number of Colors interaction was indeed significant (p = .04), whereas all other results were the same as before.

[6] Notice that our results do not support such a strategy: We found a polarity effect in the two-color condition and no Polarity × Number of Colors interaction.

Figure 4

Mean RTs for Negative and Positive Sentences in Each Condition



Note. Solid lines represent sentences with a copula; dashed lines represent sentences without a copula. Blue lines represent positive sentences; red lines represent negative sentences. Error bars represent one standard error of the mean. RT = reaction time. See the online article for the color version of this figure.

polarity (Quantificational Polarity; Polarity-Qua) that can be positive or negative. Negative quantifiers are often analyzed as an instance of implicit negation (Clark, 1976; Heim, 2006; Klima, 1964; Penka, 2011) and evoke a polarity effect in verification tasks relative to their positive counterparts (Deschamps et al., 2015; Just & Carpenter, 1971; Tucker et al., 2018). Specifically in this experiment, we used the comparative quantifiers *yoter* (*more*) and *paxot* (*fewer*, a one-morpheme word).[7] Another advantage in using negative quantifiers is in making sentences identical in their logical meaning despite their contrast in polarity, thus removing informativity from being an explanatory factor (i.e., *there are more Xs than Ys* is logically equivalent to *there are fewer Ys than Xs*).

Materials

Items again comprised sentences and pictures. Sentences indicated a proportion between blue circles and red circles (see Appendix C), and pictures depicted blue and red circles that either matched or did not match the descriptions in the sentences. All sentences used existential comparatives in a 2 × 2 paradigm (see Table 3).

The positive–positive condition is used as baseline for both negative conditions (negative–positive and positive–negative). A Polarity-Sen × Polarity-Qua interaction is not of interest, and the negative-negative condition is used merely to counterbalance the design for truth value but is not part of the data analyzed (see Appendix D for the results of the double negation condition in comparison with the other conditions).

All sentences were recorded in Hebrew, by a male native Hebrew speaker, and later processed in Audacity (Version 2.0.5; Audacity Team, 2015) to equalize them in term of their average pitch, duration and average amplitude. Pictures were created in Matlab (Version 8.5.0, 2015, the MathWorks Inc., Natick, Massachusetts, USA). The picture was always an array of 5 circles × 5 circles, divided into two clustered groups of 10 and 15 circles each. Half of the images had 15 reds and 10 blues, and the other half had 10 reds and 15 blues. Only two ratios were used in the pictures (i.e., 10:15 and 15:10) to keep the visual task easy for the participants. The separation between the two groups of circles varied: vertical, horizontal or diagonal border. Notice that in this design, there were six picture tokens that matched each negative sentence (not just one or two as in Experiment 1). Thus, neither positive nor negative sentences could be encoded as a fully specified image.

The design was a 2 × 2 × 3 × 2 design with the following factors: Polarity-Sen (sentential positive/negative), Polarity-Qua (quantificational positive/negative), ISI (800 ms/3,800 ms/6,800 ms) and Truth Value (true/false). Each condition was repeated 6 times (counterbalanced for the color mentioned in the sentence and the ratio between circles depicted in the picture). Thus, participants had to respond to 144 trials in total, divided into three runs.

Procedure

The procedure was similar to that in Experiment 1, with only the timing being a bit different. The experiment was run using Presentation software (Version 17.0, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com). In each trial, participants had to decide whether a sentence they heard on earphones correctly described a picture that appeared on the screen. The sentence always described some ratio between blue and red circles (e.g., there are more blue circles than red circles). In one third of the trials ISI was 800 ms, in one third of the trials it was 3,800 ms, and in the remaining trials ISI was 6,800 ms.

Each trial started with a fixation cross on the screen (see Figure 5). After 400 ms, the participant heard a sentence while the

[7] Extra complications that may arise with comparative sentences are ignored in this presentation (cf. Schwarzschild, 2008 and Grodzinsky et al., 2018). Past experiments found a polarity effect for these pairs in both the error and the time domains (Deschamps et al., 2015; Just & Carpenter, 1971).

Table 3
The Two Types of Polarity Factors: Sentential and Quantificational

	Sentential polarity	
Quantificational polarity	Positive	Negative
Positive	*yeš yoter igulim adumim me-kxulim* EXIST.POS more circles red than-blue 'There are more red circles than blue'	*'eyn yoter igulim adumim me-kxulim* EXIST.NEG more circles red than-blue 'There are not more red circles than blue'
Negative	*yeš paxot igulim adumin me-kxulim* EXIST.POS less circles red than-blue 'There are fewer red circles than blue'	*'eyn paxot igulim adumin me-kxulim* EXIST.NEG less circles red than-blue 'There aren't fewer red circles than blue'

Note. In each cell, first line is the transliteration of the Hebrew example sentence, second line is its word-by-word translation and gloss, third line is the English translation of the sentence. Notice that the positive-positive condition is used as baseline in the statistical analysis, and the two types of negatives are two levels in the same Polarity factor. The negative-negative condition is used to counterbalance the design, and as we have no predictions regarding the processing of double-negations, it is not taking part in the statistical analysis. NEG = negation, EXIST.POS = positive existential, EXIST.NEG = negative existential.

fixation cross was still on the screen. The duration of each sentence was 2,200 ms. The fixation cross disappeared 100 ms before sentence offset, and a picture appeared after an ISI of 800 ms; 3,800 ms; or 6,800 ms. The picture stayed on the screen until the participant decided whether the sentence correctly described the picture (true/false) by pressing one of two possible buttons on the keyboard (the "←" key or the "→" key, counterbalanced for coding between participants). Participants were encouraged to respond as fast and as accurately as possible. Once decision was made, a smiling face appeared on the screen if a correct answer was given, and a sad face if not. The face stayed on the screen for 500 ms, and then the next trial started.

Each individual participated in three runs (48 trials each) and could take a short break between runs if needed. Each run was counterbalanced for Polarity-Sen (positive/negative), Polarity-Qua (positive/negative), Truth Value (true/false), color mentioned in the sentence (red/blue), and ratio depicted in the picture.

Participants

For this experiment, we recruited individuals who had not participated in Experiment 1. Thirty-one native Hebrew speakers from the Hebrew University participated in the 30-min experiment for either payment or credit, after signing an informed consent approved by the Hebrew University Research Ethics Committee. Mean age was 24 ± 4 ($M \pm SD$), 27 were right-handed, and 19 were women.

Analysis

We fitted a linear mixed effects model, with a three-level Polarity factor—the positive baseline, sentential negation, and quantificational negation. This factor was dummy coded, such that each type of negation was compared against the positive baseline. The model also included an ISI factor (800 ms/3,800 ms/6,800 ms) with two contrasts: ISI-1, which compared the RT at ISI = 3,800 ms with the RT at ISI = 800 ms; and ISI-2, which compared the RT at ISI = 6,800 ms with the RT at ISI = 3,800 ms. Also included in the model were the Truth Value factor (true/false), and all possible interactions. Truth Value was sum coded. We also included random intercepts and random slopes for Polarity, adjusted by participants. This random structure was chosen because any model with a larger number of random variables did not converge, and the Polarity random variable had the largest variance. As in experiment 1, p values were obtained in two ways for validation purposes: Satterthwaite approximation and LRT. We report here only the p values of the Satterthwaite approximation (see Table E1 to compare with results from LRT, which were generally similar).

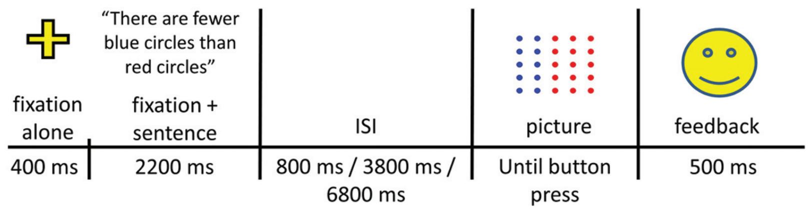
Results

Out of the 31 participants, three were removed due to poor performance (less than 65% correct responses), although all the other participants were correct in at least 80% of the trials (M = 92%, SD = 4%). Two additional participants were removed from the analysis due to extremely slow responses: Both the mean and the median of their RT were above 2,000 ms, whereas the mean RT of the rest of the participants was 1,126 ms (SE = ±129 ms). Thus, analysis was carried out with 26 participants. The accuracy rate of these 26 participants was high (M = 92%, SD = ± 4%), and it improved with the ISI: 90% in 800-ISI conditions; 93% in 3,800-ISI conditions; and 94% in 6,800-ISI conditions. A generalized mixed effects model (using Rs *glmer* function), with ISI, Polarity-Sen and Polarity-Qua as independent variables and a by-participant random slope for ISI, revealed a significant improvement between 800-ISI and 3,800-ISI (z = 2.8, p = .005), and a non-significant improvement between 3,800-ISI and 6,800-ISI (z = 1, p = .3).

For the RT analysis, only correct responses were considered (thus removing 8% of the trials). Some participants self-reported difficulties to concentrate throughout the whole experiment, and indeed some RTs were very long, probably reflecting hesitation or lack of concentration rather than mere processing time (for 2.1% of the data RT > 3,000 ms). To remove slow responses while respecting participants' individual response patterns, data was trimmed such that for each participant only responses that were within two standard deviations from the participant's average were used (thus removing 4% of the remaining trials; after trimming only .5% of the data was RT > 3,000 ms). In total, 12% of the trials were excluded due to accuracy or trimming. As before, all analyses were performed on the logarithmic transformation of the RTs.

The model resulted in a significant effect of Sentential Polarity (β = .1, t = 7.2, $p < .0001$). As depicted in Figure 6, the sentential polarity effect appears to decrease with longer ISIs (a difference of

Figure 5
Trial Structure of Experiment 2



Note. ISI = interstimulus interval. See the online article for the color version of this figure.

350 ms ± 55 ms at the 800-ISI; 226 ms ± 56 ms at the 3,800-ISI; 188 ms ± 36 ms at the 6,800-ISI). This resulted in a significant interaction of Sentential Polarity with the ISI only at the first time window (Polarity-Sen × ISI-1: $\beta = -.1$, $t = -2.4$, $p = .02$), but not the second (Polarity-Sen × ISI-2: $\beta = -.02$, $t = -.5$, $p = .6$). Between the 800-ISI and the 3,800-ISI, the effect loses 35% of its size. We also found a significant effect of Quantificational Polarity ($\beta = .2$, $t = 5.8$, $p < .0001$). As depicted in Figure 6, the quantificational polarity effect also seems to decrease with longer delays (144 ms ± 34 ms at the 800-ISI; 99 ms ±32 ms at the 3,800-ISI; 86 ms ± 26 ms at the 6,800-ISI), but this was not significant at either time window (Polarity-Qua × ISI-1: $\beta = -.03$, $t = -.9$, $p = .4$; Polarity-Qua × ISI-2: $\beta = -.008$, $t = -.2$, $p = .8$). As for the interaction with truth value: It was significant for Sentential Polarity ($\beta = -.05$, $t = -3.5$, $p = .0005$), but not for Quantificational Polarity, although it was close to significance threshold ($\beta = -.02$, $t = -1.9$, $p = .06$). Unlike our results from Experiment 1, this interaction did not change with the ISI (Polarity-Sen × ISI-1 × Truth Value: $\beta = -.003$, $t = -.8$, $p = .4$; Polarity-Sen × ISI-2 × Truth Value: $\beta = .02$, $t = .5$, $p = .6$). See Table E1 in Appendix E for the full list of results. To make sure our results were solid, we also ran two separate models, one for Sentential Polarity and one for Quantificational Polarity—and got the same results; that is, only Sentential Polarity significantly diminished with ISI and only at the first time-window.

Crucially, for both types of negative polarity, most of the effect persisted: For sentential negation, the polarity effect at the 6,800-ISI was on average 54% of the size of the initial effect; for quantificational negation, the polarity effect at the 6,800-ISI was on average 60% of the size of the initial effect. Looking at simple effects at ISI = 6,800 ms, besides for the false quantificational condition, negative was more costly than positive for all conditions ($p < .005$, linear mixed effects regression, with log-RT as the dependent variable, Polarity as the predictor, random slopes and random intercepts adjusted by participants; Bonferroni corrected for multiple comparisons). For the full list of simple effects, see Table E2 in Appendix E.

Discussion

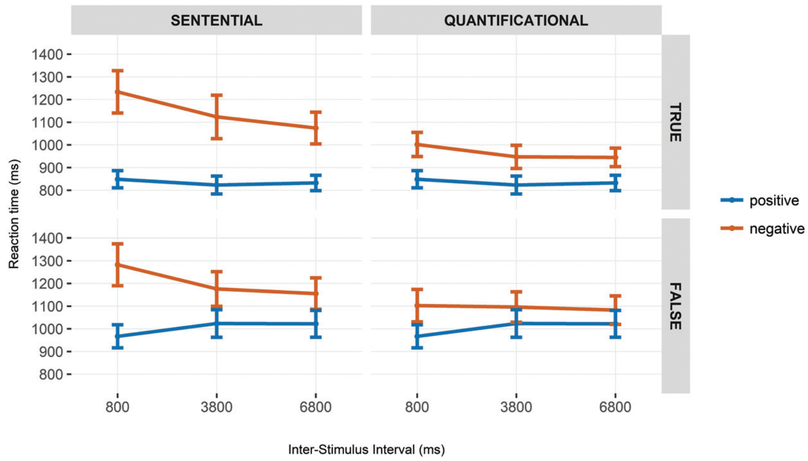
In Experiment 2, we examined the polarity effect for sentential negation and for quantificational negation. For sentential negation, we report the following: a decrease of the polarity effect at the first time window (800 ms to 3,800 ms); no decrease of the polarity effect at the second time window (3,800 ms to 6,800 ms); a polarity effect measured 6,800 ms after sentence offset. For quantificational negation, we report the following: no decrease of the polarity effect, in either time window and a polarity effect measured 6800 ms after sentence offset.

In Experiment 2, we probed the verification sequence using three levels of ISI, thus measuring the amount of decay of the polarity effect in two time windows. As in Experiment 1, we replicated the delay sensitivity of the polarity effect at early stages of processing. As predicted by the mental models theory (e.g., Kaup et al., 2007), most of the processing difficulty induced by sentential negation is eliminated at early stages of processing, somewhere between the 800-ISI and 3,800-ISI (indeed, we found a significant Polarity-Sen × ISI-1 interaction). After the 3,800-ISI, the effect retains most of its size (and indeed, no significant Polarity-Sen × ISI-2 effect). It follows that a considerable part of the effect is not due to the processing sequence.

Negative quantifiers have been theorized to have a similar pragmatic role as does sentential negation, which is to deny some contextual question under discussion. For example, "Few people came to the meeting" would be used to deny the statement "Many people came to the meeting," which entails a mental representation of both negative and positive statements (Moxey et al., 2001; Sanford & Moxey, 2004). Such a dual representation presumably should result in a processing pattern similar to that of negation, which is to say a polarity effect that diminishes given additional processing time. However, for quantificational negation, we did not find a diminishment of the polarity effect. This null result might be because the polarity effect for quantifiers is already smaller than for sentential negation, which makes it more difficult to find a significant decay. Alternatively, if indeed the polarity effect for quantifiers is not sensitive to additional processing time, it could suggest that negative quantifiers are not processed through an initial representation of the positive, as arguably happens with negation. To support this possibility, we compared the two types of overall negation and found the sentential polarity effect to be much larger than the quantificational polarity effect ($p < .001$ in a similar LMER comparing directly the two types of negations; see Grodzinsky et al., 2021 for a similar finding). This could suggest that the two types of negation induce different verification strategies. Indeed, if *few* does not contain an implicit negation, as argued by some theoreticians (De Swart, 2000; Geurts, 1996), we should not expect it to be processed as sentential negation. It is important to note that for negative quantifiers, evidence for an increased verification cost was found, that is, there was a significant polarity effect at long ISIs. We further discuss this verification component in the General Discussion.

Figure 6

Mean RTs for Negative and Positive Sentences in Each Condition



Note. For Sentential Polarity, positive = "there are" (EXIST.POS), negative = "there are not" (EXIST.NEG). For Quantificational Polarity, positive = *more*, negative = *fewer*. Error bars represent one standard error of the mean. RT = reaction time. See the online article for the color version of this figure.

General Discussion

The polarity effect is a well-known phenomenon in psycholinguistics. The early literature on transformational grammar analyzed the polarity effect solely as an outcome of the time it takes to apply the negative transformation (Gough, 1966; Katz & Postal, 1964). Later research continued with this line of thought, and attributed the polarity effect either to an additional step in the construction of the mental model (e.g., Kaup et al., 2007) or to the accommodation of infelicitous context (e.g., Nieuwland & Kuperberg, 2008). Either way, these approaches assume that the polarity effect derives from the longer time it takes to reach the meaning of the negative sentence, which we call the processing cost. According to these approaches, additional time allotted for processing should decrease and even eliminate the measured processing cost of negative sentences. By probing verification after several ISIs, we were able to assess the existence of a polarity effect up to 6.8 s after sentence offset, long after the processing of the sentence is thought to have ended. Therefore, a polarity effect that is measured after the processing sequence has been concluded most likely stems not from the processing cost but rather from the process of comparing the sentence against the picture (i.e., the verification cost). Our results demonstrate that a considerable part of the polarity effect remains after a delay (50% to 70% of the effect without a delay), and therefore should be attributed to the verification cost.

The verification cost associated with negation can be explained by two sources: a later onset of the verification procedure, or a longer duration of the verification procedure. Negation can affect the onset of the verification procedure if accessing the meaning of a negative sentence is more difficult. Negation also can affect the duration of the verification procedure if it takes longer to check the picture against a negative sentence. For the verification procedure to be sensitive to negation, negation-related features must be preserved in WM during the delay. Therefore, regardless of whether the delayed effect stems from accessing the representation or verifying it, the WM representation of negative sentences is likely to be systematically different from that of positive sentences. Thus, from the broader question of whether negation affects verification times after sentence processing, we can narrow down the question to: How is negation represented in WM?

Broadly speaking, the representation of meaning in WM can be linguistic or nonlinguistic. By "linguistic" we mean symbols or features related to the structure of the sentence, for example, by logical, propositional or syntactic trees. We use the term *nonlinguistic* to represent some abstract form of meaning that is detached from the particular linguistic structure, such as mental models. There is no consensus in the literature as to whether WM maintains language-specific representations (cf. Caplan & Waters, 1999; Fedorenko et al., 2006). However, old experiments with delayed verification did find a sustained effect for linguistic factors, such as morphological complexity and syntactic complexity, whereas extralinguistic factors such as the frequency of the items were eliminated in longer delays (Holyoak et al., 1976; Seymour, 1974). This suggests that WM does not preserve any kind of complexity, but that linguistic structure is somehow part of the WM representation. These two possibilities, though, are not necessarily mutually exclusive: It is possible that nonlinguistic factors could have an impact in addition to that of linguistic factors (see also Dale & Duran, 2011). However, these two types of representations are fundamentally different with regard to the kinds of features that they maintain in WM. Next, we discuss these two types of representations and their predictions regarding a sustained polarity effect.

Nonlinguistic Representations

Mental models provide a theoretical, nonlinguistic framework for language comprehension. In this framework, a sentence such as "The square is blue" is comprehended through the activation of the same cognitive processes that are activated upon seeing a blue square, namely, visualizing its properties in one's mind. Such pictorial encoding already had been suggested prior to the emergence of theories of mental models. In this framework, the time it takes to translate the sentence into a mental image contributes to the polarity effect, and therefore a diminishment of the polarity effect is predicted if measured after a delay (Macleod et al., 1978; Mathews et al., 1980; Tversky, 1975). However, such a strategy can be used only under particular conditions when there is a clear one-to-one translation from sentence to picture. This is not the case in many experimental settings, including ours (the three-color condition of Experiment 1 and in Experiment 2 where each sentence could match any of six pictures). This also is not the case in most real-life examples, not only for negative sentences but also for positive ones. Outside the experiment room, blue squares can come in various hues, sizes and locations, such that even a simple sentence like "The square is blue" does not match any one particular picture unequivocally (Clark & Chase, 1972 raised a similar point). It is therefore unlikely that simple visual mental imagery serves as the sole representation of meaning in general, and of negative sentences in particular.

As opposed to the simple strategy of pictorial encoding, contemporary views on mental representations of sentence meaning do not assume a fully specified, picture-like image. A mental representation, according to this view, is much less restricted than an actual picture, and can leave many attributes unspecified. Thus, the representation of the sentence "The square is not blue" would contain a square with its color unspecified (Kaup et al., 2007). An immediate question that arises from such a theory is how to account for the difference in meaning between two sentences that potentially could be described by the same mental image (e.g., "There is a square" and "The square is not blue"). Both should evoke a mental model of a square with an unspecified color. One possible answer would be that for negative sentences, the actual state is not the only one that is encoded. Kaup et al. (2007) suggested that negative sentences, due to their pragmatic role as denying a presupposed statement, encode the deviation between two mental models—that of the positive and that of the negative—and that the meaning of a negative sentence arises from the comparison between the two (see also Giora et al. (2005; 2007) for a similar view). Thus, "The square is not blue" gives rise to two representations (one of a blue square and one of an unspecified square), whereas "There is a square" gives rise only to one (an unspecified square).

This means that the positive sentence is part of the sentence meaning, and has to be represented throughout the comprehension process, until the verification process is initiated. If the positive representation is kept with the negative one, this might be costly at any delay. Our finding of a Polarity × Truth Value interaction after 1,600 ms (Experiment 1) supports this kind of explanation. A representation of the positive should facilitate the response to false negatives, bringing about a Polarity × Truth Value interaction (see The Processing Cost of Negation section in the Introduction). Therefore, this interaction after 1,600 ms suggests that a positive representation is still active at later processing stages (although weaker, as evident by the Polarity × ISI × Truth Value interaction). However, retention of the positive representation is inconsistent with findings in the literature that support its full suppression after ~1,500 ms (Hasson & Glucksberg, 2006; Kaup et al., 2006, 2005; Kaup & Zwaan, 2003). Those studies show a priming effect of negative sentences to positive-related concepts only within a short time window after sentence offset. No such priming is found when delaying the task by 1,000 ms to 1,500 ms, which is evidence of suppression of the positive representation. The findings in the literature, though, were obtained not in a verification paradigm but in experiments whose focus was manipulating only the lexical material of the sentence (e.g., naming paradigms). It is likely that there are substantial differences in the cognitive processes that are required when probing the meaning of the whole sentence, such as in verification paradigms, as opposed to when probing single words in a sentence. Thus, it may be that the suppression of the positive representation works differently for single words than for full sentences. This hypothesis, however, requires additional investigation.

A second issue that arises from the theory of mental models is determining how unspecified a representation can be. A sentence like "This is not a blue square" should have a dual representation of a blue square alongside something which is not blue, or which is not a square, or both. If all its attributes are left unspecified, what exactly would contain the mental model of the negative representation? To resolve the issue of unspecified attributes, several relevant scenarios might be considered for a negative sentence. For example, the sentence "The square is not blue" evokes simulations of squares of various specified colors, such as red or yellow. The complexity of a negative sentence might depend on the number of representations, that is, it grows with the number of alternative mental models that the listener must consider (Khemlani et al., 2012; Macbeth et al., 2017). In Experiment 1, we saw that this was not the case, as we found no difference in the polarity effects of the two-color condition and the three-color condition. To be cautious of overinterpreting null results, it is important to note that the polarity effect exists even when the positive and negative sentences are equivalent in meaning and have the same number of alternatives (Experiment 2: There are more X than Y ≈ there are fewer Y than X). Hence, the polarity effect is not simply an outcome of a different number of considered scenarios.

Finally, a third question that can arise from mental models is whether a dual representation theory can apply to the polarity effect in quantifiers. In Experiment 2, we found that the polarity effect of quantifiers did not diminish with longer ISIs. This may be due to negative quantifiers not including an implicit negation in their mental representation, and hence having no processing time dedicated to the "integration of negation." That might mean that there is no dual representation for negative quantifiers. We leave this question open.

Linguistic Representations

Linguistic representations are abstract symbols, possibly in the form of propositional logic. Regarding the processing of negation, Clark and Chase (1972) and Carpenter and Just (1975) suggested that both the sentence and the picture in a verification task are represented as logical propositions, and the two propositions are compared against each other. For example, [NEG(BLUE, SQUARE)] would be the representation of the negative sentence "The square is not blue" and [AFF(BLUE, SQUARE)] would be the representation of the positive sentence "The square is blue" (AFF = affirmative; a positive polarity marker). Pictures always are represented with a positive polarity marker, which results in a mismatch in representations between negative sentences and pictures. According to this theory, then, the polarity effect stems from an additional step in the comparison procedure between the negative proposition of the sentence and the positive proposition of the picture. This should result in a polarity effect at any delay.

As for negative quantifiers, some studies have suggested that, rather than comparing two propositions, the verification itself is determined by the quantifier's logical form. According to the Interface Transparency Thesis (Lidz et al., 2011), verification strategies are affected by the logical structure of the sentence, even in cases of logical equivalence. For example, the quantifiers *most* and *more than half* have the same truth conditions (i.e., they are true and false under the same scenarios) but evoke different verification strategies due to their different logical representations (Hackl, 2009; Hunter et al., 2017; Pietroski et al., 2009). The polarity effect of quantifiers, therefore, can stem from differences between the logical representation of negative and positive quantifiers, even in cases where they share the same truth conditions. Barwise and Cooper (1981); in their seminal work on the logical properties of quantifiers, hypothesized that negative quantifiers should evoke longer RTs due to their "[higher] complexity of the checking procedure" (Barwise & Cooper, 1981, p. 192). Their prediction was that a sentence with *few* (such as "Few of the dots are red") triggers a search algorithm that goes over all subsets of red dots, whereas a sentence with *many* (such as "Many of the dots are red") triggers a shorter search algorithm that goes over only some of the red dots. This is because for a sentence like "Few of the dots are red" to be true, there should be no subset of many red dots, so all subsets of red dots must be checked. On the other hand, for the sentence "Many of the dots are red" to be true, it is sufficient to encounter one subset of many red circles to stop the search algorithm. Hence, when comparing true sentences, the search algorithm takes more time for *few* than for *many* (see Deschamps et al., 2015 for criticism).

Overall, there seems to be a gap between the available explanations of a delayed polarity effect for sentential negation and for negative quantifiers. A unified account that explains both the delayed effect of sentential negation and of negative quantifiers would be more parsimonious and hence desirable.[8] Currently, there is one such suggestion for a unified account: Both negations share the logical property of reversing entailment patterns, which adds to their representational complexity (Agmon et al., 2019; Deschamps et al., 2015). In the following examples, adding either sentential negation (Sentence 1b) or a negative quantifier (Sentence 1c) changes the direction of the logical entailment presented in Sentence 1a:

1a. There are more red circles than blue circles.

⇒ There are more red circles than dark blue circles.

1b. There *are not* more red circles than blue circles.

⇐ There *are not* more red circles than dark blue circles.

1c. There are *fewer* red circles than blue circles.

⇐ There are *fewer* red circles than dark blue circles.

A reversed direction of entailment arguably is pertinent to linguistic behavior (Chierchia, 2013; Fauconnier, 1975; Ladusaw, 1980). For example, words like *any* or *ever* (also called *negative polarity items*) can appear only when the direction of entailment is reversed, such as in Sentences 2b and 3b in the following examples (asterisks denote unacceptability):

2a. *John read any book.

2b. John did *not* read any book.

3a. *Many students read any book.

3b. *Few* students read any book.

Native speakers of English implicitly know how to use negative polarity items, without being able to articulate it explicitly. This sensitivity in language usage possibly signals a similar sensitivity in cognitive processes. Indeed, the reverse direction of entailment was shown to have a cognitive impact on RT and on brain signals (Agmon et al., 2019, 2021). Maintaining an opposite entailment pattern might add to the WM load, for example if part of the representation includes a feature that signals a nondefault direction of entailment. To speculate, if part of sentence comprehension is knowing what the sentence entails and what is entailed by it, then it is crucial to have such a signaling feature as part of the representation. This feature in turn might contribute to the complex representation of both negative quantifiers and sentential negation.

Representation Revisited?

Finally, it may be that the polarity effect observed after a delay is not due to representation at all, but actually reflects a delayed processing cost. This could be the case if the sentence is reanalyzed when the verification task is presented. It also could be the case if participants simply memorize the linguistic stimulus, especially during long ISIs. In other words, in this framework the delayed polarity effect would not stem from the representation (linguistic or nonlinguistic), but rather it reflects a processing sequence *ab initio*, that is a shadow of the same cognitive processes that have already occurred. Similarly to this perspective, Pattamadilok et al. (2016) showed an fMRI effect of syntactic processing only upon the probing task. They hypothesize that upon task demand, the sentence is reactivated and undergoes full analysis. On the face of it, this hypothesis suggests a cognitive process that wastes cognitive resources in the reanalysis of the sentence, rather than keeping the representation in WM (which is intuitively what participants do when asked to keep a meaning in memory). Importantly, even if our cognitive system does prefer reanalysis instead of utilizing the WM capacity, the fact that the polarity effect decreases when adding more processing time is evidence for at least some processing. Indeed, it could be that the sentence is reanalyzed only partially, which could explain why the polarity effect decays with longer ISIs, but then the question is: What parts of the sentence are reanalyzed and what parts are preserved? This issue calls for further investigation.

[8] A correlation between the sentential polarity effect and the quantifier polarity effect would support the idea of a shared resource underlying the two. Indeed, at ISI = 800 ms, the correlation between the two polarity effects was .47 (p = .02). However, the correlation weakened at later ISIs (3,800 ms: r = .38, p = .05; 6,800 ms: r = .32, p = .12), possibly due to noise added by variation in WM functions.

Conclusions

To conclude, we showed that RT effects reflect not only a longer processing sequence, but also a component that is measured after processing is concluded and is likely due to representation. Consideration of both processing and verification costs in RT is important for understanding both language comprehension, in general, and negation, in particular. Only for sentential negation did we find a diminishment of the polarity effect with the delay, possibly due to two-stage processing. For both sentential negation and quantificational negation, we found a sustained effect, possibly due to a similar representational complexity. Further research is required to precisely characterize the sources of the sustained component and the interplay between them.

References

Agmon, G., Bain, J. S., & Deschamps, I. (2021). Negative polarity in quantifiers evokes greater activation in language-related regions compared with negative polarity in adjectives. *Experimental Brain Research*, *239*, 1427–1438. https://doi.org/10.1007/s00221-021-06067-y

Agmon, G., Loewenstein, Y., & Grodzinsky, Y. (2019). Measuring the cognitive cost of downward monotonicity by controlling for negative polarity. *Glossa: A Journal of General Linguistics*, *4*(1), 36. https://doi.org/10.5334/gjgl.770

Audacity Team. (2015). Audacity(R): Free audio editor and recorder (Version 2.0.5) [Computer application]. https://audacityteam.org/

Barsalou, L. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes*, *18*(5–6), 513–562. https://doi.org/10.1080/01690960344000026

Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, *4*(2), 159–219. https://doi.org/10.1007/BF00350139

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Brooks, L. R. (1967). The suppression of visualization by reading. *The Quarterly Journal of Experimental Psychology*, *19*(4), 289–299. https://doi.org/10.1080/14640746708400105

Brooks, L. R. (1970). An extension of the conflict between visualization and reading. *The Quarterly Journal of Experimental Psychology*, *22*(2), 91–96. https://doi.org/10.1080/00335557043000014

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*(1), 77–94. https://doi.org/10.1017/S0140525X99001788

Carpenter, P. A., & Just, M. A. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, *82*(1), 45–73. https://doi.org/10.1037/h0076248

Chierchia, G. (2013). *Logic in grammar: Polarity, free choice, and intervention*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199697977.001.0001

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

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

Clark, H. H., & Chase, W. G. (1974). Perceptual coding strategies in the formation and verification of descriptions. *Memory & Cognition*, *2*(1), 101–111. https://doi.org/10.3758/BF03197499

Claus, B., & Kelter, S. (2009). Embodied language comprehension: The processing of spatial information during reading and listening. In A. M. Columbus (Ed.), *Advances in Psychology Research*, (pp. 1–44). Nova Science Publishers.

Dale, R., & Duran, N. D. (2011). The cognitive dynamics of negated sentence verification. *Cognitive Science*, *35*(5), 983–996. https://doi.org/10.1111/j.1551-6709.2010.01164.x

De Swart, H. (2000). Scope ambiguities with negative quantifiers. In K. Von Heusinger & U. Egli (Eds.), *Reference and anaphoric relations* (Vol. 72, pp. 109–132). Springer. https://doi.org/10.1007/978-94-011-3947-2_6

Deschamps, I., Agmon, G., Loewenstein, Y., & Grodzinsky, Y. (2015). The processing of polar quantifiers, and numerosity perception. *Cognition*, *143*, 115–128. https://doi.org/10.1016/j.cognition.2015.06.006

Doron, E. (1983). *Verbless predicates in Hebrew*. University of Texas at Austin.

Dudschig, C., & Kaup, B. (2018). How does "not left" become "right"? Electrophysiological evidence for a dynamic conflict-bound negation processing account. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(5), 716–728. https://doi.org/10.1037/xhp0000481

Eddy, J. K., & Glass, A. L. (1981). Reading and listening to high and low imagery sentences. *Journal of Verbal Learning and Verbal Behavior*, *20*(3), 333–345. https://doi.org/10.1016/S0022-5371(81)90483-7

Fauconnier, G. (1975). Pragmatic scales and logical structure. *Linguistic Inquiry*, *6*(3), 353–375.

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of Memory and Language*, *54*(4), 541–553. https://doi.org/10.1016/j.jml.2005.12.006

Fischler, I., Bloom, P. A., Childers, D. G., Roucos, S. E., & Perry, N. W., Jr. (1983). Brain potentials related to stages of sentence verification. *Psychophysiology*, *20*(4), 400–409. https://doi.org/10.1111/j.1469-8986.1983.tb00920.x

Geurts, B. (1996). On "no." *Journal of Semantics*, *13*(1), 67–86. https://doi.org/10.1093/jos/13.1.67

Giora, R., Balaban, N., Fein, O., & Alkabets, I. (2005). Negation as positivity in disguise. In H. L. Colston & A. Katz (Eds.), *Figurative language comprehension* (pp. 233–258). Lawrence Erlbaum. https://doi.org/10.4324/9781410611628

Giora, R., Fein, O., Aschkenazi, K., & Alkabets-zlozover, I. (2007). Negation in context: A functional approach to suppression. *Discourse Processes*, *43*(2), 153–172. https://doi.org/10.1080/01638530709336896

Givón, T. (1978). Negation in language: Pragmatics, function, ontology. In P. Cole (Ed.), *Syntax and semantics* (Vol. 9, pp. 69–112). Academic Press.

Glenberg, A. M., Robertson, D. A., Jansen, J. L., & Johnson-Glenberg, M. C. (1999). Not propositions. *Cognitive Systems Research*, *1*(1), 19–33. https://doi.org/10.1016/S1389-0417(99)00004-2

Gough, P. B. (1965). Grammatical transformations and speed of understanding. *Journal of Verbal Learning and Verbal Behavior*, *4*(2), 107–111. https://doi.org/10.1016/S0022-5371(65)80093-7

Gough, P. B. (1966). The verification of sentences: The effects of delay of evidence and sentence length. *Journal of Verbal Learning and Verbal Behavior*, *5*(5), 492–496. https://doi.org/10.1016/S0022-5371(66)80067-1

Greenberg, Y. (2008). Predication and equation in Hebrew (nonpseudocleft) copular sentences. In S. Armon-Lotem, G. Danon, & S. D. Rothstein (Eds.), *Current issues in generative Hebrew linguistics* (pp. 161–196). John Benjamins. https://doi.org/10.1075/la.134.07pre

Grodzinsky, Y., Agmon, G., Snir, K., Deschamps, I., & Loewenstein, Y. (2018). The processing cost of downward entailingness: The representation and verification of comparative constructions. In U. Sauerland & S. Solt (Eds.), *Proceedings of Sinn Und Bedeutung* (pp. 435–451). https://doi.org/10.21248/zaspil.60.2018.475

Grodzinsky, Y., Behrent, K., Agmon, G., Bittner, N., Jockwitz, C., Caspers, S., Amunts, K., & Heim, S. (2021). A linguistic complexity pattern that defies aging: The processing of multiple negations. *Journal of Neurolinguistics*, *58*, 100982. https://doi.org/10.1016/j.jneuroling.2020.100982

Hackl, M. (2000). *Comparative quantifiers*. MIT Press.

Hackl, M. (2009). On the grammar and processing of proportional quantifiers: Most versus more than half. *Natural Language Semantics*, *17*(1), 63–98. https://doi.org/10.1007/s11050-008-9039-x

Hasson, U., & Glucksberg, S. (2006). Does understanding negation entail affirmation? An examination of negated metaphors. *Journal of Pragmatics*, *38*(7), 1015–1032. https://doi.org/10.1016/j.pragma.2005.12.005

Heim, I. (2006). Little. In M. Gibson & J. Howell (Eds.), *Proceedings of Semantics and Linguistic Theory* (pp. 35–58). Cornell University. https://semanticsarchive.net/Archive/zIyMzJkY/Little%20SALT%2016.pdf

Henrik, S., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2019). afex: Analysis of factorial experiments (R package, Version 0.25-1) [Computer software]. https://CRAN.R-project.org/package=afex

Holyoak, K. J., Glass, A. L., & Mah, W. A. (1976). Morphological structure and semantic retrieval. *Journal of Verbal Learning and Verbal Behavior*, *15*(2), 235–247. https://doi.org/10.1016/0022-5371(76)90022-0

Horn, L. R. (1989). *A natural history of negation*. University of Chicago Press.

Hunter, T., Lidz, J., Odic, D., & Wellwood, A. (2017). On how verification tasks are related to verification procedures: A reply to Kotek et al *Natural Language Semantics*, *25*(2), 91–107. https://doi.org/10.1007/s11050-016-9130-7

Jacobs, J. (1991). Negation. In A. von Stechow & D. Wunderlich (Eds.), *Semantics: An international handbook* (pp. 560–596). de Gruyter.

Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science*, *4*(1), 71–115. https://doi.org/10.1016/S0364-0213(81)80005-5

Just, M. A., & Carpenter, P. A. (1971). Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior*, *10*(3), 244–253. https://doi.org/10.1016/S0022-5371(71)80051-8

Katz, J. J., & Postal, P. M. (1964). *An integrated theory of linguistic descriptions*. MIT Press.

Kaup, B., & Zwaan, R. A. (2003). Effects of negation and situational presence on the accessibility of text information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(3), 439–446. https://doi.org/10.1037/0278-7393.29.3.439

Kaup, B., Lüdtke, J., & Zwaan, R. A. (2005). Effects of negation, truth value, and delay on picture recognition after reading affirmative and negative sentences. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 27, pp. 1114–1119). Lawrence Erlbaum.

Kaup, B., Lüdtke, J., & Zwaan, R. A. (2006). Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed? *Journal of Pragmatics*, *38*(7), 1033–1050. https://doi.org/10.1016/j.pragma.2005.09.012

Kaup, B., Yaxley, R. H., Madden, C. J., Zwaan, R. A., & Lüdtke, J. (2007). Experiential simulations of negated text information. *Quarterly Journal of Experimental Psychology*, *60*(7), 976–990. https://doi.org/10.1080/17470210600823512

Kaup, B., Zwaan, R. A., & Lüdtke, J. (2007). The experiential view of language comprehension: How is negation represented? In F. Schmalhofer & C. A. Perfetti (Eds.), *Higher level language processes in the brain: Inference and comprehension processes* (pp. 255–288). Lawrence Erlbaum.

Khemlani, S., Orenes, I., & Johnson-Laird, P. N. (2012). Negation: A theory of its meaning, representation, and use. *Journal of Cognitive Psychology*, *24*(5), 541–559. https://doi.org/10.1080/20445911.2012.660913

Klima, E. (1964). Negation in English. In J. A. Fodor & J. J. Katz (Eds.), *The structure of language* (pp. 246–323). Prentice-Hall.

Krueger, L. E. (1972). Sentence-picture comparison: A test of additivity of processing time for feature matching and negation coding. *Journal of Experimental Psychology*, *95*(2), 275–284. https://doi.org/10.1037/h0033650

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Ladusaw, W. A. (1980). On the notion of affective in the analysis of negative-polarity items. *Journal of Linguistics Research*, *1*, 1–16.

Lidz, J., Pietroski, P., Halberda, J., & Hunter, T. (2011). Interface transparency and the psychosemantics of most. *Natural Language Semantics*, *19*(3), 227–256. https://doi.org/10.1007/s11050-010-9062-6

Lüdtke, J., Friedrich, C. K., De Filippis, M., & Kaup, B. (2008). Event-related potential correlates of negation in a sentence-picture verification paradigm. *Journal of Cognitive Neuroscience*, *20*(8), 1355–1370. https://doi.org/10.1162/jocn.2008.20093

Macbeth, G. E., Crivello, M., del, C., Fioramonti, M. B., & Razumiejczyk, E. (2017). Chronometrical evidence supports the model theory of negation. *SAGE Open*, *7*(2), 215824401771621. https://doi.org/10.1177/2158244017716216

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, *24*(1), 56–98. https://doi.org/10.1016/0010-0285(92)90003-K

Macleod, C. M., Hunt, E. B., & Mathews, N. N. (1978). Individual differences in the verification of sentence-picture relationships. *Journal of Verbal Learning and Verbal Behavior*, *17*(5), 493–507. https://doi.org/10.1016/S0022-5371(78)90293-1

Mathews, N. N., Hunt, E. B., & MacLeod, C. M. (1980). Strategy choice and strategy training in sentence—Picture verification. *Journal of Verbal Learning and Verbal Behavior*, *19*(5), 531–548. https://doi.org/10.1016/S0022-5371(80)90595-2

Mayo, R., Schul, Y., & Burnstein, E. (2004). I am not guilty" vs "I am innocent": Successful negation may depend on the schema used for its encoding. *Journal of Experimental Social Psychology*, *40*(4), 433–449. https://doi.org/10.1016/j.jesp.2003.07.008

Moxey, L. M., & Sanford, A. J. (1986). Quantifiers and focus. *Journal of Semantics*, *5*(3), 189–206. https://doi.org/10.1093/jos/5.3.189

Moxey, L. M., Sanford, A. J., & Dawydiak, E. J. (2001). Denials as controllers of negative quantifier focus. *Journal of Memory and Language*, *44*(3), 427–442. https://doi.org/10.1006/jmla.2000.2736

Nieuwland, M. S., & Kuperberg, G. R. (2008). When the truth is not too hard to handle: An event-related potential study on the pragmatics of negation. *Psychological Science*, *19*(12), 1213–1218. https://doi.org/10.1111/j.1467-9280.2008.02226.x

Orenes, I., Moxey, L., Scheepers, C., & Santamaría, C. (2016). Negation in context: Evidence from the visual world paradigm. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *69*(6), 1082–1092. https://doi.org/10.1080/17470218.2015.1063675

Pattamadilok, C., Dehaene, S., & Pallier, C. (2016). A role for left inferior frontal and posterior superior temporal cortex in extracting a syntactic tree from a sentence. *Cortex*, *75*, 44–55. https://doi.org/10.1016/j.cortex.2015.11.012

Penka, D. (2011). *Negative indefinites*. Oxford University Press.

Pietroski, P., Lidz, J., Hunter, T., & Halberda, J. (2009). The meaning of "most": Semantics, numerosity, and psychology. *Mind & Language*, *24*(5), 554–585. https://doi.org/10.1111/j.1468-0017.2009.01374.x

Roberts, C. (2012). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, *5*(1), 1–69. https://doi.org/10.3765/sp.5.6

Rullmann, H. (1995). The ambiguity of comparatives with "less". In J. M. Fuller, H. Ho, & D. Parkinson (Eds.), *Proceedings of the Eleventh Eastern States Conference on Linguistics* (pp. 258–269). Cornell University.

Sanford, A. J., & Moxey, L. M. (2004). Exploring quantifiers: pragmatics meets the psychology of comprehension. In I. A. Noveck & D. Sperber (Eds.), *Experimental pragmatics* (pp. 116–137). Palgrave Macmillan. https://doi.org/10.1057/9780230524125_6

Sanford, A. J., Dawydiak, E. J., & Moxey, L. M. (2007). A unified account of quantifer perspective effects in discourse. *Discourse Processes*, *44*(1), 1–32. https://doi.org/10.1080/01638530701285556

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin*, *2*(6), 110–114. https://doi.org/10.2307/3002019

Schwarzschild, R. (2008). The semantics of comparatives and other degree constructions. *Language and Linguistics Compass*, *2*(2), 308–331. https://doi.org/10.1111/j.1749-818X.2007.00049.x

Seymour, P. H. K. (1974). Generation of a pictorial code. *Memory & Cognition*, *2*(2), 224–232. https://doi.org/10.3758/BF03208987

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, *12*(2), 153–156. https://doi.org/10.1111/1467-9280.00326

Tian, Y., & Breheny, R. (2016). Dynamic pragmatic view of negation processing. In P. Larrivee & C. Lee (Ed.), *Negation and polarity: Experimental perspectives. language, cognition, and mind* (pp. 21–43, vol 1). Springer. https://doi.org/10.1007/978-3-319-17464-8_2

Tian, Y., Breheny, R., & Ferguson, H. J. (2010). Why we simulate negated information: A dynamic pragmatic account. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *63*(12), 2305–2312. https://doi.org/10.1080/17470218.2010.525712

Tian, Y., Ferguson, H., & Breheny, R. (2016). Processing negation without context—Why and when we represent the positive argument. *Language, Cognition and Neuroscience*, *31*(5), 683–698. https://doi.org/10.1080/23273798.2016.1140214

Trabasso, T., Rollins, H., & Shaughnessy, E. (1971). Storage and verification stages in processing concepts. *Cognitive Psychology*, *2*(3), 239–289. https://doi.org/10.1016/0010-0285(71)90014-4

Tucker, D., Tomaszewicz, B., & Wellwood, A. (2018). Decomposition and processing of negative adjectival comparatives. In E. Castroviejo, L. McNally, & G. W. Sassoon (Eds.), *The semantics of gradability, vagueness, and scale structure: Experimental perspectives* (pp. 243–273). Springer. https://doi.org/10.1007/978-3-319-77791-7_10

Tversky, B. (1975). Pictorial encoding of sentences in sentence-picture comparison. *The Quarterly Journal of Experimental Psychology*, *27*(3), 405–410. https://doi.org/10.1080/14640747508400500

Wason, P. C. (1961). Response to affirmative and negative binary statements. *British Journal of Psychology*, *52*(2), 133–142. https://doi.org/10.1111/j.2044-8295.1961.tb00775.x

Wason, P. C. (1965). The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior*, *4*(1), 7–11. https://doi.org/10.1016/S0022-5371(65)80060-3

Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record*, *58*(3), 475–482. https://doi.org/10.1007/BF03395630

Xiang, M., Grove, J., & Giannakidou, A. (2016). Semantic and pragmatic processes in the comprehension of negation: An event related potential study of negative polarity sensitivity. *Journal of Neurolinguistics*, *38*, 71–88. https://doi.org/10.1016/j.jneuroling.2015.11.001

(*Appendices follow*)

Appendix A

Sentences Used in Experiment 1

Table A1

English translation	Hebrew original
Without copula:	
1. The square is blue	הריבוע כחול
2. The square is red	הריבוע אדום
3. The square is yellow	הריבוע צהוב
4. The triangle is blue	המשולש כחול
5. The triangle is red	המשולש אדום
6. The triangle is yellow	המשולש צהוב
7. The circle is blue	העיגול כחול
8. The circle is red	העיגול אדום
9. The circle is yellow	העיגול צהוב
10. The star is blue	הכוכב כחול
11. The star is red	הכוכב אדום
12. The star is yellow	הכוכב צהוב
13. The square is not blue	הריבוע לא כחול
14. The square is not red	הריבוע לא אדום
15. The square is not yellow	הריבוע לא צהוב
16. The triangle is not blue	המשולש לא כחול
17. The triangle is not red	המשולש לא אדום
18. The triangle is not yellow	המשולש לא צהוב
19. The circle is not blue	העיגול לא כחול
20. The circle is not red	העיגול לא אדום
21. The circle is not yellow	העיגול לא צהוב
22. The star is not blue	הכוכב לא כחול
23. The star is not red	הכוכב לא אדום
24. The star is not yellow	הכוכב לא צהוב
With copula:	
1. The square is blue	הריבוע הוא כחול
2. The square is red	הריבוע הוא אדום
3. The square is yellow	הריבוע הוא צהוב
4. The triangle is blue	המשולש הוא כחול
5. The triangle is red	המשולש הוא אדום
6. The triangle is yellow	המשולש הוא צהוב
7. The circle is blue	העיגול הוא כחול
8. The circle is red	העיגול הוא אדום
9. The circle is yellow	העיגול הוא צהוב
10. The star is blue	הכוכב הוא כחול
11. The star is red	הכוכב הוא אדום
12. The star is yellow	הכוכב הוא צהוב
13. The square is not blue	הריבוע אינו כחול
14. The square is not red	הריבוע אינו אדום
15. The square is not yellow	הריבוע אינו צהוב
16. The triangle is not blue	המשולש אינו כחול
17. The triangle is not red	המשולש אינו אדום
18. The triangle is not yellow	המשולש אינו צהוב
19. The circle is not blue	העיגול אינו כחול
20. The circle is not red	העיגול אינו אדום
21. The circle is not yellow	העיגול אינו צהוב
22. The star is not blue	הכוכב אינו כחול
23. The star is not red	הכוכב אינו אדום
24. The star is not yellow	הכוכב אינו צהוב

(*Appendices continue*)

Appendix B

Full Statistical Results of Experiment 1

Table B1

Summary of the Effects of the Model Fitted for Experiment 1

log_RT ~ Polarity + ISI + Truth_value + Copula + Number_of_Colors + Order + Polarity:ISI + Polarity:Truth_value + Polarity:Copula + Polarity:Number_of_Colors + Polarity:Order + ISI:Truth_value + Polarity:ISI:Truth_value + (1 + Polarity + ISI + Truth_value + Number_of_Colors + Polarity:ISI + Polarity:Truth_value | Participant)

Effect(difference)	Average difference ± SEM (ms)	β±SEM (log ms)	*t*-value	Chi-square	*p*-value
Polarity (negative-positive)	228±27	0.1±0.01	11	52.6	<.0001
ISI (1600ms-100ms)	−100±9	−0.06±0.005	−14.4	69	<.0001
Truth-Value (false-true)	72±16	0.05±0.008	6.6	28	<.0001
Copula (copula-null)	15±6	0.007±0.003	2.9	6.7	=.004 (0.1)
Number-of-Colors (three-two)	42±19	0.03±0.009	2.7	4.9	=.01 (0.3)
Order (second-first)	−48±18	−0.03±0.009	−3.7	7.8	=.0007 (0.005)
Polarity×ISI	−75±13	−0.02±0.003	−6.5	26.7	<.0001
Polarity×Truth-Value	−183±21	−0.06±0.005	−11	52.7	<.0001
Polarity×Copula	−24±12	−0.007±0.003	−2.8	6	=.005 (0.01)
Polarity×Number-of-Colors	19±21	0.004±0.003	1.6	0.7	=.1 (0.4)
Polarity×Order	−42±20	−0.01±0.003	−5.1	24.3	<.0001
ISI×Truth-Value	18±11	0.005±0.003	2	2.5	=.04 (0.12)
Polarity×ISI×Truth-Value	88±21	0.01±0.003	3.8	12.5	=.0002 (0.0004)

Note. This table shows the summary of the effects of the model given above (written in R's syntax). Differences are calculated within subjects and then averaged across subjects. For interactions, we calculated the difference between differences, e.g. RT of negative polarity is longer by 228±27 on average than RT of positive polarity, and the Polarity Effect for false sentences is smaller by 183±21 ms on average compared to true sentences. For the statistical analyses, all contrasts were sum coded, such that the β is the estimated difference from the mean. Chi-square is the statistic of the log-likelihood ratio test model comparison (LRT). *P*-values reported are from the Satterthwaite approximation of degrees of freedom (Satterthwaite, 1946) implemented in R's lmerTest package (Kuznetsova et al., 2017). When the *p*-values from the LRT were different from the *p*-value from the Satterthwaite approximation, we added the LRT *p*-value in parentheses. When comparing to a model with the full factorial design, results are very similar (based only on Satterthwaite approximation, for ease of computation), besides the following few differences, which are peripheral to our research questions: interaction of ISI × Truth-Value became marginally significant (p = .05 instead of p = .04), Order became more significant (p = .0001 instead of .0007). T-statistics didn't change though. The full factorial model included 50 more predictors than in this model, out of which the following 8 were the only ones significant ($p < .05$). For the sake of brevity, we report here only these additional 8 predictors: ISI × Copula (t = −2.6, p = .01), ISI × Number-of-Colors (t = 2.6, p = .009), Truth-Value × Number-of-Colors (t = −2.9, p = .004), Polarity × Truth-Value × Copula (t = −2.1, p = .03), Polarity × Truth-Value × Number-of-Colors (t = −4.5, $p < .0001$), ISI × Copula × Number-of-Colors (t = 2.8, p = .005), Polarity × ISI × Truth-Value × Copula × Order (t = 2.5, p = .01), Polarity × Truth-Value × Copula × Number-of-Colors × Order (t = 2.6, p = .008). ISI = interstimulus interval.

Table B2

Simple Polarity Effects for Experiment 1

ISI	Number of colors	Copula	Truth value	Average difference ± SEM (ms)	β±SEM (log ms)	*t*-value	Chi-square	*p*-value
ISI = 1600 ms	2-Color Block	Copula	TRUE	235±65	0.13±0.02	6.3	26.9	<.0001*
			FALSE	123±56	0.06±0.02	3.3	9.6	=.002*
		No copula	TRUE	214±51	0.1±0.02	6.9	30.9	<.0001*
			FALSE	140±52	0.07±0.02	4.2	14.4	=.0002*
	3-Color Block	Copula	TRUE	289±50	0.2±0.01	11.6	57.6	<.0001*
			FALSE	76±44	0.02±0.01	1.8	3.1	=.08
		No copula	TRUE	303±60	0.2±0.02	10.7	52.9	<.0001*
			FALSE	146±56	0.06±0.02	3.6	11.3	=.0008*
ISI = 100 ms	2-Color Block	Copula	TRUE	349±66	0.2±0.02	10.7	52.7	<.0001*
			FALSE	127±46	0.07±0.02	4.6	16.9	<.0001*
		No copula	TRUE	366±54	0.2±0.02	13.3	66.0	<.0001*
			FALSE	193±63	0.1±0.02	5.7	23.4	<.0001*
	3-Color Block	Copula	TRUE	395±76	0.2±0.02	12.6	62.7	<.0001*
			FALSE	134±49	0.06±0.02	4.1	13.8	=.0002*
		No copula	TRUE	405±72	0.2±0.02	14.2	70.0	<.0001*
			FALSE	153±47	0.08±0.01	6.5	28.3	<.0001*

Note. This table shows the summary of the simple effects of Polarity within each of the other conditions. For example, first line in this table reports the result of the Polarity Effect for only true sentences that include a copula, within the 2-color block condition, at the longest level of ISI. At the longest level of ISI, an effect is indicative of an increased verification cost of negation. All models were fitted with the Polarity factor sum coded (negative: +1; positive: −1), and with random intercepts and Polarity slopes adjusted by participants. The *t*-values are the statistics of the β coefficient for Polarity in each simple model. The Chi-squares are the statistics of log-likelihood ratio test model comparison (LRT), comparing each model with a nested model that excludes the Polarity fixed effect. Nested models had the same random structure as the full models. *P*-values were obtained in two different methods, to validate the robustness of our results: (i) the Satterthwaite approximation of degrees of freedom (Satterthwaite, 1946) implemented in R's lmerTest package (Kuznetsova et al., 2017); (ii) LRT model comparison. The two methods resulted in the same *p*-values. All effects with $p < .05$ remain significant after a Bonferroni correction for multiple comparisons (* $p < .05$ corrected for multiple comparisons). ISI = interstimulus interval.

Appendix C

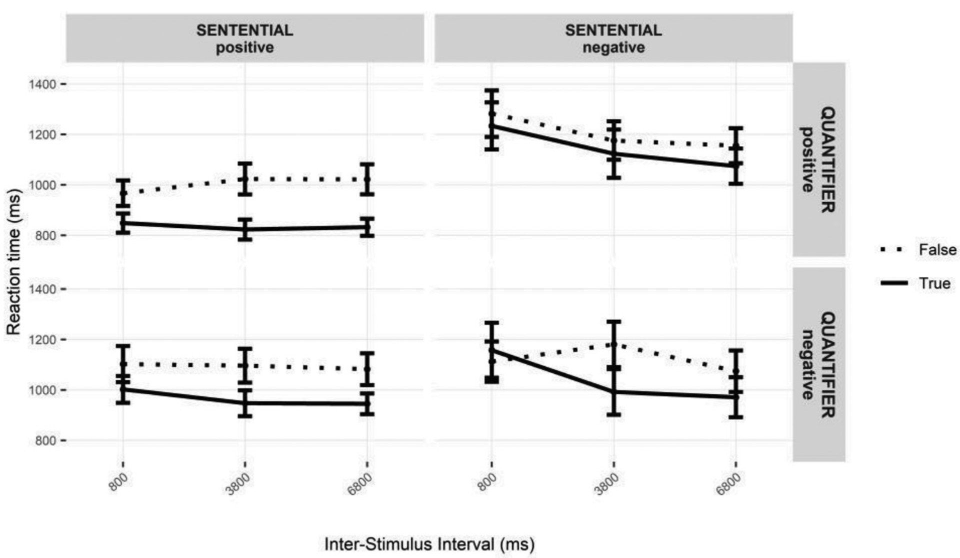
Sentences Used in Experiment 2

English translation	Hebrew original
1. There are more blue circles than red circles	יש יותר עיגולים כחולים מאדומים
2. There are fewer blue circles than red circles	יש פחות עיגולים כחולים מאדומים
3. There are more red circles than blue circles	יש יותר עיגולים אדומים מכחולים
4. There are fewer red circles than blue circles	יש פחות עיגולים אדומים מכחולים
5. There aren't more blue circles than red circles	אין יותר עיגולים כחולים מאדומים
6. There aren't fewer blue circles than red circles	אין פחות עיגולים כחולים מאדומים
7. There aren't more red circles than blue circles	אין יותר עיגולים אדומים מכחולים
8. There aren't fewer red circles than blue circles	אין פחות עיגולים אדומים מכחולים

Appendix D

The Negative-Negative Condition Relative to the Other Conditions

Figure D
Full RT Results for Experiment 2



Note. This table shows the full results – including the negative-negative condition (as opposed to Figure 6 which shows only the individual polarity effects).

(*Appendices continue*)

Appendix E

Full Statistical Results of Experiment 2

Table E1

Summary of the Effects of the Model Fitted for Experiment 2

log_RT ∼ Polarity*ISI*Truth_Value + (1 + Polarity | Participant)					
Effect (difference)	Average difference ± SEM (ms)	β±SEM (log ms)	*t*-value	Chi-square	*p*-value
Polarity-Sen (no-positive)	255±42	0.2±0.03	7.2	29.06	<.0001
Polarity-Qua (less-positive)	110±23	0.1±0.02	5.8	22.33	<.0001
ISI-1 (3800ms-800ms)	−41±17	0.007±0.02	0.3	0.11	=.8 (0.7)
ISI-2 (6800ms-3800ms)	−13±17	0.009±0.02	0.4	0.14	=.7
Truth-Value (false-true)	120±21	0.08±0.009	9	78.89	<.0001
Polarity-Sen×ISI-1	−124±42	−0.1±0.04	−2.4	8.92	=.02 (0.003)
Polarity-Sen×ISI-2	−39±46	−0.02±0.04	−0.5	0.3	=.6
Polarity-Qua×ISI-1	−46±33	−0.03±0.04	−0.9	1.08	=.4 (0.3)
Polarity-Qua×ISI-2	−12±36	−0.008±0.04	−0.2	0.06	=.8
Polarity-Sen×Truth-Value	−109±35	−0.05±0.01	−3.5	12.07	=.0005
Polarity-Qua×Truth-Value	−41±26	−0.02±0.01	−1.9	3.56	=.06
ISI-1×Truth-Value	45±41	0.04±0.02	1.8	3.16	=.07 (0.08)
ISI-2×Truth-Value	2±35	−0.01±0.02	−0.5	0.26	=.6
Polarity-Sen×ISI-1×Truth-Value	−78±77	−0.003±0.03	−0.8	0.54	=.4 (0.46)
Polarity-Sen×ISI-2×Truth-Value	39±94	0.02±0.03	0.5	0.15	=.6 (0.7)
Polarity-Qua×ISI-1×Truth-Value	−33±74	−0.02±0.03	−0.5	0.21	=.65
Polarity-Qua×ISI-2×Truth-Value	−1±85	−0.002±0.03	−0.07	0.01	=.9

Note. This table shows the summary of effects of the model given above (written in R's syntax). Differences are calculated within subjects and then averaged across subjects. For interactions, we calculated difference between differences, e.g. the sentential polarity effect is overall 255±42, and it gets smaller between the 800-ISI and 3800-ISI by 124±42. For the statistical analysis, the Polarity factor was dummy coded (i.e. the β of Polarity-Sen is the estimate of the difference between "there aren't" sentences and the positive baseline, and the β of Polarity-Qua is the estimate of the difference between "fewer" sentences and the positive baseline); the ISI factor was difference coded (i.e. the β of ISI-1 is the estimate of the difference between ISI = 3800ms and the ISI = 800ms, and the β of ISI-2 is the estimate of the difference between 6800ms and 3800ms); the Truth-Value factor was sum coded (i.e. the β reflects the difference between false sentences and the grand mean). Chi-square is the statistic of the log-likelihood ratio test (LRT) model comparison, provided by R's mixed function in the afex package (Henrik et al., 2019). *P*-values are based on the Satterthwaite approximation of degrees of freedom (Satterthwaite, 1946) implemented in R's lmerTest package. Where the p-values from the LRT were different from the *p*-values from the Satterthwaite approximation, we added the LRT *p*-value in parentheses. ISI = interstimulus interval.

Table E2

Simple Polarity Effects for Experiment 2

ISI	Polarity type	Truth value	Average difference ± SEM (ms)	β±SEM (log ms)	*t*-value	Chi-square	*p*-value
ISI = 6800 ms	Sentential Polarity	TRUE	242±51	0.1±0.02[a]	6.1	35.0	<.0001*
		FALSE	133±42	0.06±0.02	3.7	10.8	=.001*
	Quantificational Polarity	TRUE	122±27	0.06±0.02[a]	3.9	14.9	=.0001*
		FALSE	60±42	0.03±0.01	1.6	2.5	=.1
ISI = 3800 ms	Sentential Polarity	TRUE	300±70	0.1±0.03	4.7	15.9	<.0001*
		FALSE	152±59	0.07±0.02	3.1	8.2	=.005 (0.004)
	Quantificational Polarity	TRUE	124±33	0.06±0.02[a]	3.9	14.7	=.0001*
		FALSE	73±51	0.04±0.02	1.8	3.1	=.08
ISI = 800 ms	Sentential Polarity	TRUE	385±68	0.2±0.02[a]	8.6	64.1	<.0001*
		FALSE	315±52	0.1±0.02[a]	7.5	50.7	<.0001*
	Quantificational Polarity	TRUE	153±33	0.08±0.02[a]	4.5	19.1	<.0001*
		FALSE	135±48	0.06±0.02[a]	3.7	13.2	=.0003*

Note. This table shows the summary of the simple effects of Polarity within each of the other conditions. For example, first line this this table reports the result of the polarity effect for only true sentences of the sentential contrast ("there aren't" vs. "there are"), at the longest level of ISI. At the longest level of ISI, an effect of Polarity is indicative of an increased verification cost of negation. All models were fitted with the Polarity factor sum coded (negative: +1; positive: −1), and with random intercepts and Polarity slopes adjusted by participants. Some models had a singular fit with a Polarity random effect, and therefore their random structure included only a random intercept (marked with [a] on their β coefficient). The *t*-values are the statistics of the β coefficient for Polarity in each simple model. The Chi-squares are the statistics of a log-likelihood ratio test model comparison (LRT), comparing each model with a nested model that excludes the Polarity fixed effect. Nested models had the same random structure as the full models. P-values were obtained in two different methods, to validate the robustness of our results: (i) the Satterthwaite approximation of degrees of freedom (Satterthwaite, 1946) implemented in R's lmerTest package (Kuznetsova et al., 2017); (ii) LRT model comparison. The two methods resulted in the same p-values, besides in one model where they differed only slightly. In that case (ISI = 3800ms, sentential negation, false), Satterthwaite approximated p-value is reported and LRT *p*-value is added in parentheses. This was also the only simple effect that did not survive a Bonferroni correction for multiple comparisons (* $p < .05$ corrected for multiple comparisons). ISI = interstimulus interval.

Received July 28, 2020
Revision received May 10, 2021
Accepted May 11, 2021 ■