A model for pragmatic strengthening: Evidence from a new priming paradigm

Nadine Bade* WooJin Chung† Léo Picat‡ Rachel Dudley§
Salvador Mascarenhas¶

June 2023

Abstract

The paper explores the role of alternatives in pragmatic reasoning using a novel priming-with-feedback paradigm. In two experiments we investigate the question whether pragmatic strengthening involving contextual alternatives primes pragmatic strengthening with lexical alternatives, and the other way around. We found that, while training people with lexical cases (disjunction) primes them to derive strong readings of other cases involving lexical scales, it does not prime them to derive strong readings of sentences involving contextual alternatives. However, the converse holds: priming them for strong readings of sentences with contextual alternatives increases the rate of strengthened meanings of sentences with lexical alternatives. Based on our findings, we offer a model of pragmatic strengthening which distinguishes between thresholds for different alternative types.

Keywords: semantic priming, alternatives, implicature

1 Introduction

In conversation, listeners often derive meanings of sentences that go beyond their literal interpretation. This *utterance* or *pragmatic* meaning can be the result of the listener considering what sentences the speaker could have used but did not, i.e. *alternative utterances*. For example, given the context in (1), the listener might consider the alternative in (1b) when hearing (1a). Given that this alternative was not used, the pragmatic

*Department of Linguistics, University of Potsdam
†Department of Linguistics, Seoul National University
‡Université de Paris, UFR de médecine, F-75005 Paris, France
§Cognitive Development Center, Central European University
¶Institut Jean Nicod, Département d'études cognitives, ENS, EHESS, CNRS, PSL University

interpretation the listener might arrive at is (1c).

(1) CONTEXT: Yolanda had spinach and eggs for lunch.
   a. UTTERANCE: She liked the spinach.
   b. ALTERNATIVE: She liked the spinach and the eggs.
   c. ENRICHED MEANING: She liked the spinach but not the eggs.

Traditionally, it is not taken to be part of the *literal* meaning of (1a) that Yolanda did not like the eggs.[1] This is easy to see when considering a minimally different context where it is common knowledge that Yolanda only had spinach for lunch. In that context, an inference that she did not like the eggs does not arise for (1a). However, given the context in (1) and assuming that the speaker is as informative as possible (following the *cooperative principle* and *maxim of quantity*, Grice, 1989), the listener can infer (1c) from (1a): since (1b) would have been more informative but was not used by the speaker, they must have meant to convey that (1b) is not true. The result of this reasoning process is the stronger enriched meaning in (1c), which consists of the literal meaning in (1a) plus the inference that (1b) is false. This inferred non-literal meaning component is a type of *conversational implicature*.

There is an ongoing debate in the theoretical and experimental literature regarding the exact nature of the strengthening mechanism behind quantity-based *conversational implicatures* (referred to as quantity implicatures in what follows) (see in particular Breheny et al., 2006; Magri, 2009; Huang and Snedeker, 2009; Geurts, 2010; Grodner et al., 2010; Franke, 2011; Chierchia et al., 2012; Chierchia, 2013; Spector, 2016).[2] A central issue in this debate is whether different kinds of alternatives influence the pragmatic enrichment mechanism differently, and how (Horn, 1972; Levinson, 2000; Fox and Katzir, 2011; Breheny et al., 2013; Bott and Chemla, 2016; Rees and Bott, 2018; Waldon and Degen, 2020).

Two sources of alternatives for quantity implicatures have been identified in the literature: the *lexicon* and the *context*. Example (1) above illustrates a case where alternatives are provided by the context of utterance. But lexically determined alternatives also appear to play an important role in the derivation of quantity implicatures, and were in fact at the origin of theoretical and experimental investigations into the phenomenon. For example, since the disjunction 'or' forms a lexical scale with the logically stronger 'and', (2b) is a salient alternative to (2a), even in the absence of a rich context (Horn, 1972). Consequently, the disjunctive sentence in (2a) *competes* with the lexically derived alternative in (2b). The strengthening in (2c) emerges, often referred to as the exclusive reading of disjunction.

[1] On grammatical approaches to scalar implicatures of this sort, the view is on the contrary that the enriched meaning *is* part of the literal meaning of the utterance (Chierchia et al., 2012). In this introduction, we will continue assuming that the implicatures of interest are properly pragmatic, for ease of exposition. As we will explain in section 2, the question whether these processes are grammatical or properly pragmatic isn't directly relevant to the question of interest in this article.

[2] The term *conversational implicature* more generally refers to inferences resulting from all maxims of conversation introduced by Grice (1989): quantity, quality, relevance, and manner. We are exclusively concerned with inferences based on quantity in this article. We also refer to them simply as *implicatures* on occasion, when no confusion could arise.

(2) a. UTTERANCE: She liked the spinach or the eggs. DISJUNCTION
b. ALTERNATIVE: She liked the spinach and the eggs. CONJUNCTION
c. ENRICHED MEANING: She liked the spinach or the eggs but not both.

Quantity implicatures involving lexical scales have been shown to be far less context dependent. That is, (2c) is a salient reading of (2a) even without specifying a context of utterance as we did for (1). Yet, on the surface the alternatives involved in both cases are exactly the same.

The lexical (2) and contextual (1) implicatures just reviewed share a high-level description in terms of reasoning about a speaker's communicative intentions as a function of statements they conspicuously did not make. Yet, the underlying *mechanisms* for accessing and excluding *alternatives*, that is the relevant sentences that the speaker did not utter, are assumed to be distinct. Specifically, whereas one requires lexical access (what alternatives is the sentence lexically associated with?), the other requires a context search (what alternatives are salient in the context?). This article addresses the question whether these two cases can or should be treated uniformly. To do so, we report on two studies testing whether the enrichment processes involved in lexical and contextual cases prime each other. To look more into the nature of the alternatives involved, we included more complex cases involving conjunction within a complex disjunction as in (3a). The sentence has two readings beyond its literal meaning: a relatively weak one in (3b), and a much stronger reading in (3c).

(3) CONTEXT: Yolanda had spinach, eggs, and potatoes for lunch.
a. She liked the spinach and the eggs, or she liked the potatoes.
COMPLEX DISJUNCTION
b. She did not like all three at once (the spinach and the eggs and the potatoes).
c. Either she liked the spinach and the eggs but not the potatoes, or she liked the potatoes but not the spinach and not the eggs.

The simple exclusive inference in (3b) is well-known, and can be derived by any theory that derives exclusive interpretations for simple disjunctions, including a naive theory that simply states that natural-language 'or' is ambiguous between an inclusive interpretation and an exclusive one, the latter amounting to the logician's Xor ('either or'). The implicature in (3c) seems perhaps more exotic, but it has been observed in the literature on the basis of introspective judgments (Spector, 2007), and studied in some detail in the context of reasoning problems (Mascarenhas, 2014; Koralus and Mascarenhas, 2018; Picat, 2019).

Importantly, unlike the cases in (2c) or in (3b), the strengthening in (3c) *cannot* be derived by simply taking it that English 'or' sometimes behaves like the logician's *exclusive 'or'*. Take a schematic representation of (3a), for ease of exposition: [*a* and *b*] or *c*. A simple Xor analysis of 'or' in this schema yields [*a* and *b* and not *c*] or [*c* and not [*a* and *b*]]. But notice that (3c) is much stronger than this, it corresponds to the schema [*a* and *b* and not *c*] or [*c* and not *a* and not *b*].

Thus, the example in (3) works as a rather refined test case of what strategies humans use when deriving implicatures of this sort. Our experimental investigation

into the two processes illustrated in (1) and (2) involved a paradigm of priming with feedback, where we trained participants on sentences with 'or' and manipulated the contexts of utterance. We gave them feedback meant to push them toward literal or strengthened interpretations of those sentences. Testing participants with sentences as the one given in (3c) then allowed us to check precisely what participants were primed for: a particular strategy for dealing with the word 'or', which could not plausibly be generalized to derive the strong meaning in (3c), or a far more abstract mechanism as has been proposed in the theoretical literature, which would be able to generate the strong interpretation in (3c).

Our main finding is that priming across alternative types (contextual versus lexical) is only one-directional: priming with lexical cases affects contextual cases but not the converse. Based on this finding, we suggest that the process of strengthening is a one-step process, where activation of the alternative is sufficient to trigger the mechanism of exclusion (Rees and Bott, 2018). However we also offer a modification of the existing model which distinguishes between different types of alternatives. Specifically, we propose different activation thresholds for the derivation of lexical and contextual alternatives.

2 Background

Two dominating questions in the theoretical and experimental literature on quantity implicatures concern the exact nature of the strengthening mechanism and the role of alternatives therein. These issues are intricately linked, both theoretically and experimentally. On the theoretical side, different alternative-generating mechanisms have been shown to produce diverging results for strengthening (Franke, 2011; Spector, 2016). On the experimental side, alternatives have been shown to determine when these processes are activated (Degen and Tanenhaus, 2015; Rees and Bott, 2018; Waldon and Degen, 2020).

Theories of pragmatic strengthening can be categorized across two dimensions of interest for present purposes: whether they assume uniformity of alternatives, and whether they assume uniformity of mechanism.

(4) **Uniformity of alternatives**: Alternatives are derived in a unified manner for contextual and lexical quantity implicatures.

(5) **Uniformity of mechanism/operator**: There is a single mechanism of strengthening behind contextual and lexical quantity implicatures.

The aim of our investigation in this article is to study these two dimensions in combination, rather than in isolation, thereby informing models of quantity implicature competence and performance alike.

2.1 Theories of quantity implicature competence

Most competence theories of the phenomenon hold that the same exclusion mechanism for alternatives is at play in the derivation of all quantity implicatures. They differ as

to whether the alternatives themselves are uniform in the lexical and contextual cases.

According to the classical view founded by Grice (1975) the mechanism behind strengthening is abductive reasoning over more informative things the speaker could have said. What these more informative things — the alternatives — are is taken to be highly context dependent and not formally determined. However, differences between contextual and lexical cases had been observed early on (Horn, 1972; Levinson, 2000). Concretely, scales provided by the lexicon itself (e.g. 'all' $>$ 'most' $>$ 'some') seem to be involved in scalar implicatures with logical words, such as quantifiers ('some', 'most') and propositional connectives ('or', 'if…then') (Horn, 1972). These *lexical* cases of implicature have since been considered to be only minimally influenced by the alternatives provided by the context of utterance.

Under a more recent alternative view (the *grammatical view*) a covert syntactic operator is the (uniform) source of "pragmatic" strengthening.[3] Simplifying somewhat, this operator takes a set of formally defined alternatives *Alt* and a proposition *p*, and excludes those propositions in *Alt* that are not entailed by *p*. Its operation is very similar to that of the exclusive particle 'only' (*modulo* presuppositions). It can occur locally, arbitrarily embedded in the syntax of the utterance, and it can be obligatory (Chierchia et al., 2012; Chierchia, 2013; Magri, 2009). The idea behind postulating such an operator is that speakers, when describing a state of affairs, are as *exhaustive* as possible, a variation of the original Gricean observation that people are usually trying to be maximally informative. There exist two notable versions of this operator, one working with *minimal models* (Groenendijk and Stokhof, 1984; Van Rooij and Schulz, 2004; Schulz and Van Rooij, 2006; Spector, 2007) and one working with *innocent exclusion* (Fox, 2007). These two approaches differ in what alternatives are needed in order to derive observed implicatures in certain cases. These particulars are not operative in the sentences our experimental study is about, so we refer the reader to the literature for a detailed comparison between these two variants of the grammatical approach (Franke, 2011; Spector, 2016).

Another view is offered by Geurts (2010), who proposes a Gricean reasoning mechanism overall in quantity implicatures, with disjunction as a relatively special case. Against the traditional Gricean account, Geurts argues that, since '*a* or *b*' communicates independently that 'the speaker does not know *a* and the speaker does not know *b*', it makes little sense to think about the speaker uttering the *conjunction* of claims *a* and *b*, a necessary step in the lexical account reviewed above. The resulting inference, in this view, is not an implicature but an inference based on the probability of both disjuncts being true in a given context. Reasoning about communicative intentions is not required.

[3] Scare quotes around "pragmatic" are in order, for in this view the strengthening observed in scalar implicatures is not directly a matter of pragmatic enrichment via considerations about speaker communicative intentions, but rather an in-principle ambiguity as to what is the *literal* (semantic) meaning of the utterance. Pragmatics will of course still be involved in this process, but it will be in its much more general function of ambiguity resolution: is there good independent reason to assume that the speaker intended her utterance to be interpreted with the covert operator in question?

2.2 Experimental background and models of performance

The general importance and relevance of alternatives in generating quantity implicatures has previously been established with experimental methods (Breheny et al., 2013; Chemla and Bott, 2014; Degen and Tanenhaus, 2015; Bott and Chemla, 2016; Gotzner et al., 2016; Van Tiel and Schaeken, 2017; Rees and Bott, 2018; Waldon and Degen, 2020).

Degen and Tanenhaus (2015) report three different experiments using sentence verification tasks on the quantity implicature associated with 'some' ('some but not all'). Their results show that both the size of the domain and the presence of numerals as salient alternatives influence the rate of implicatures associated with 'some'. They propose a constraint-based model which predicts that the process of deriving the implicature can sometimes be delayed and sometimes immediate, depending on whether the right contextual conditions are met. They identify the presence and relevance of the right alternatives — such as evoked by a contextual question — as one such condition. They suggest that the difference between contextual and lexical cases need not be rooted in a difference in the mechanisms involved, but rather in different constraints on activation. This model reflects both the more complex theoretical issue of activation of the mechanism and its interaction with alternatives outlined above, as well as the conflicting empirical findings on delays.

Chemla and Bott (2014) present results from reaction-time studies with sentence-verification tasks showing that quantity implicatures associated with 'some' and free choice inferences associated with disjunction ('you may have cake or ice-cream' implies 'you may have cake and you may have ice-cream') display different signatures. They argue that this does not speak against assuming the same underlying mechanism for both inference types, but that the difference could instead be rooted in the fact that the generation of alternatives works differently for the two cases. To derive the implicature of sentences with 'some', the lexical alternative 'all' has to be accessed. However, in the case of disjunction the alternatives are found within the utterance itself: the two disjuncts. In a priming experiment, Bott and Chemla (2016) show that there is priming across lexical and contextual domains, suggesting that they share at least an important part of the enrichment mechanism. They consider different explanations for this: one based on the search for proper alternatives being shared (or not), the other being that the exclusion mechanism itself is primed (or not).

Van Tiel and Schaeken (2017) find differences between lexical ('some') and contextual cases in a picture-verification task with abstract shapes. They argue that this supports a *lexical access* view of scalar implicatures, where it is the accessing of the lexicalized scalar alternative that causes delay in decision times.

Rees and Bott (2018) used the same priming paradigm as Bott and Chemla (2016) to test the role of alternatives in implicature computation for existential constructions involving contextual alternatives ('there is a star' when there is also a heart) versus lexical cases involving 'some' and numerals ('some of the hearts are red' when all of them are). They found that the presence of the alternative itself primed exclusion mechanisms. That is, even when priming did not force participants to assume the stronger meaning but just made them aware of its presence they derived implicatures to a higher degree in the probing phase. Based on these findings, they argue for a *salience model*

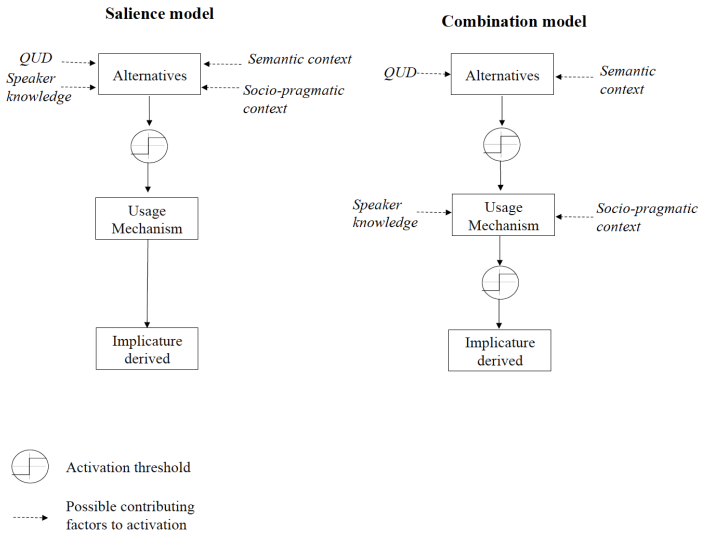


Figure 1: Two models for a strengthening mechanism.

of pragmatic strengthening, in contrast to a *combined model*. The latter is a two-step model which proposes that activating alternatives and activating a mechanism for their exclusion are discrete steps. Both are triggered independently and a certain activation threshold has to be met for each. The salience model is a simpler, one-step model which takes the activation of the alternative to be the threshold for activating exclusion. As soon as the threshold for activating the alternative is met, the strengthening mechanism itself (exclusion of the alternative) will be triggered. The two models are depicted in Figure 1 (on page 7).

In view of their results, Rees and Bott (2018) revisit the theoretical options discussed by Bott and Chemla (2016) and argue that the priming observed across domains was probably due to a search for alternatives activated in both processes. They find that contextual alternatives differ slightly from lexical alternatives, and suggest that this may be due to higher activation thresholds. This contrasts in part with the processing results reported by Van Tiel and Schaeken (2017), which suggest that contextual cases lead to lower reaction times than the quantity implicature associated with 'some'. More recently, Waldon and Degen (2020) partially replicated the findings of Bott and Chemla (2016) and Rees and Bott (2018). Employing the same priming paradigm, they again find evidence for priming of quantity implicatures across different expressions (numerals, existential expressions, and 'some'). Like Rees and Bott (2018), they find that exposure to their respective 'canonical' alternatives ('and' for contextual cases and 'all' for 'some') modulates inferences. However, unlike Rees and Bott (2018), they do find differences between priming with strong readings and priming only with alternatives.

2.3 Summary

Given the theoretical positions and extant experimental evidence laid out above, there are three possible hypotheses regarding the alternatives and strengthening mechanism behind quantity implicatures. Under what we'll call H_1, contextual and lexical cases are completely parallel in the derivation of alternatives as well as the mechanism for excluding them. According to our H_2, these two cases are completely distinct. H_3 assumes partial overlap between the cases. Previous experimental findings on lexical cases involving 'some' versus contextual cases point to H_3.

Hypothesis 1 ***Uniformity*** There is a single mechanism for generating lexical and contextual alternative and a single mechanism behind their exclusion.

Hypothesis 2 ***Non-Uniformity*** The alternative generating mechanism and their exclusion mechanism are different for contextual and lexical alternatives.

Hypothesis 3 ***Partial Uniformity*** The alternative generating mechanism is different for contextual and lexical alternatives but the exclusion mechanism is the same (or vice versa).

3 Experiments

The aim of our experiments was to test the hypotheses stated above regarding the role of contextual versus lexical scales in pragmatic strengthening. For that purpose, we designed two feedback-based priming experiments which tested the influence of lexical and contextual cases on each other.

3.1 Experiment 1

Experiment 1 tested the influence of contextual cases and lexical cases using different sentence types. The goal was to see whether priming participants with strong or weak readings of one sentence type would affect interpretation of the other. We will see that whereas lexical cases affected other lexical cases (simple and complex disjunctions), they did not affect contextual cases. The converse, however, was true: priming with contextual cases influenced the rate of implicatures associated with lexical scales.

3.1.1 Procedure

The experiment proceeded in two steps: a priming-with-feedback phase and a probing phase. During the first phase, participants in one condition were primed with feedback-based training to accept either weak or strong readings of sentences with disjunction involving lexical alternatives ('*a* or *b* is red,' which has lexical alternative '*a* and *b* is red'). Another group of participants were similarly trained to accept weak or strong readings of sentences involving contextual alternatives in the absence of disjunction ('*a* is red,' in a context containing a red *a* object and a red *b* object). In what follows, we refer to priming with lexical alternatives as "lexical priming," and to priming with

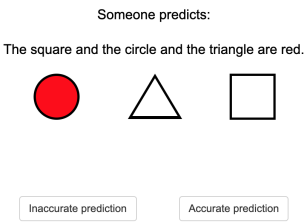


Figure 2: Screenshot of a priming trial (condition = false control).

contextual alternatives as "contextual priming." We will go over the different conditions, primes, and readings in detail shortly. In the probing phase of the experiment, participants in the two groups were asked to judge the same sentence types plus more complex sentences involving disjunction and conjunction ('*a* and *b*, or *c* is red'). Participants did not receive feedback on their responses during the probing phase. The general idea behind this two-step procedure was to prime and train participants for certain readings of a given sentence type and then test the influence of this training-based priming on different (weak and strong) readings of a new and more complex sentence type. These complex disjunctions were not encountered during training. Additionally, as reviewed in the introduction, the strong reading of these sentences cannot be derived by simply learning an ad hoc (exclusive) interpretation for disjunction.

For the priming phase, participants were instructed that the experiment was about a guessing game in which someone predicts what, if any, of the shapes that will be displayed are red. Participants' task was to decide whether the prediction was accurate given the picture they saw. They were told that they would be given feedback on their decisions in the first half of the experiment. Phrasing instructions and framing the experiment in terms of a guessing game served two purposes. First, it made clear that sentences were about what is red, thereby determining what is at issue. Second, introducing sentences as guesses made disjunctions felicitous descriptions. As is well known, disjunctions come with the inference that the speaker is ignorant with regard to which of the two disjuncts is true. If the person uttering a sentence had visual access to the picture it would clear to that person which of the shapes is red. As a result, uttering a disjunctive statement would always be under-informative (and possibly even uncooperative). This would make such an utterance highly infelicitous. However, in a guessing scenario the ignorance with regard to which of the shapes is red becomes perfectly reasonable. Thus, we expected participants to not make additional assumptions regarding how uncooperative or unreliable a speaker is. To remind participants of this general set-up, target sentences were always preceded by 'Someone predicts.' The picture showing different shapes appeared after a delay of 1500ms. A sample of what a trial looked like is given in Figure 2.

Sentence	Picture	False under
The triangle or the square is red. (6)		exclusive reading
The triangle or the square is red. (3)		no reading
The triangle or the square is red. (3)		any reading

Table 1: Critical pairings of disjunctive sentences and pictures in the priming/training phase for groups 1 and 2 (LEXICAL priming type). Numbers in brackets indicate number of occurrences in the design.

3.1.2 Design and participants

The independent variables of the experiment were PRIMING TYPE (contextual priming vs. lexical priming) and PRIMING STRENGTH (were they trained to accept also weak or only strong readings?). These levels were fully crossed to yield the following 4 conditions: LEXICAL-STRONG, LEXICAL-WEAK, CONTEXTUAL-STRONG, CONTEXTUAL-WEAK. These conditions are also the 4 priming groups that participants were randomly assigned to. The dependent variable was rate of 'no'-responses to critical primes. Saying 'no' is an indication of the relevant strong readings, as we will explain presently. We were interested in whether priming affected the stability of these readings and how. We recruited 199 participants via Prolific. They received what Prolific labeled "good" pay for taking the experiment (GBP 7.50/h). The actual average pay exceeded this amount, as most participants were quicker than the estimated time they were allotted (GBP 11/h).

3.1.3 Materials

Primes and feedback Participants were confronted with different sentence types in the priming phase, depending on which priming group they were in. Participants in groups 1 and 2 were primed with sentences involving lexical alternatives, in this case strong and weak readings of disjunction (PRIMING TYPE = LEXICAL). That is, they saw critical sentences such as (6) below. The weaker, inclusive, reading of disjunction is paraphrased in (6a), and its stronger, exclusive, reading in (6b). Sentences were paired with picture types that falsified these different readings, as in Table 1.

(6) The triangle or the square is red. SIMPLE DISJUNCTION

a. *The triangle or the square is red, possibly both.* INCLUSIVE READING
b. *The triangle or the square is red but not both.* EXCLUSIVE READING

Within the participants that were primed with lexical cases (disjunction), there was a strong and weak priming group (PRIMING STRENGTH = WEAK/STRONG). The feedback differed for these groups. Participants in the WEAK group got positive feedback if they said "accurate" to pictures that falsified the exclusive reading of a disjunction (first row of Table 1). Participants in the STRONG group got negative feedback if they said

"accurate" in the same situation. Both groups should say "accurate" in pictures that falsified no reading (second row of Table 1) and "inaccurate" to pictures that falsified all readings (third row of Table 1). They received suitable feedback in these cases. The full response-feedback matrix can be found in Appendix A.

Participants in groups 3 and 4 were trained with strong and weak readings of sentences such as (7a) and (7b) (PRIMING TYPE = CONTEXTUAL), which involved contextual alternatives. The context was provided by the picture displayed, which always contained more shapes than the ones mentioned in the target sentence. For example, the picture with respect to which (7a) would be evaluated might show a square alongside the mentioned triangle, and the picture for (7b) might show a circle alongside the mentioned triangle and square.

(7) a. The triangle is red. MONOCLAUSAL
b. The triangle and the square are red. CONJUNCTION

For these sentences, there are two possible readings: the weaker non-exhaustive readings in (8a) and (9a), which allow for other (contextually given) things to be red, and the stronger exhaustive readings paraphrased in (8b) and (9b), which require the sentence to be an exhaustive description of what contextually given things are red.

(8) The triangle is red. MONOCLAUSAL
a. *The triangle is red and possibly something else is.* NON-EXHAUSTIVE READING
b. *The triangle is red and nothing else is.* EXHAUSTIVE READING

(9) The triangle and the square are red. CONJUNCTION
a. *The triangle and square are red and possibly something else is.* NON-EXHAUSTIVE READING
b. *The triangle and square are red and nothing else is.* EXHAUSTIVE READING

In these two contextual-priming groups, sentences were paired with pictures making the strong (exhaustive) reading false or true. To prevent participants from developing strategies based on specific picture types, we varied how many (non-)red shapes there were in the picture (at most 2). The sentence-picture pairings we used are given in Table 2.

As with groups 1 and 2 discussed above, groups 3 and 4 differed as to whether the primed reading was weak or strong (PRIMING STRENGTH = WEAK/STRONG). The strong priming group received negative feedback when saying "accurate" to picture conditions that falsified the exhaustive reading. The weak group received positive feedback when saying "accurate" to these sentence-picture pairings. Participants in all groups should say "accurate" in trials that verified any reading and "inaccurate" to sentence-picture pairings that falsified any reading (and were given feedback accordingly). The full response-feedback matrix can be found in Appendix A.

Probes Probe items were the same for all groups of participants. The first kind of probe items included the same sentence types as used in the priming phase described

Sentence	Picture	False under
The triangle is red. (2)		exhaustive reading
The triangle is red. (1)		no reading
The triangle is red. (1)		any reading
The triangle is red. (2)		exhaustive reading
The triangle is red. (1)		no reading
The triangle is red. (1)		any reading
The triangle and the square are red. (2)		exhaustive reading
The triangle and the square are red. (1)		no reading
The triangle and the square are red. (1)		any reading

Table 2: Critical pairings of sentences and pictures in the priming/training phase for groups 3 and 4 (CONTEXTUAL priming type). Numbers in brackets indicate number of occurrences in the design. The "exhaustive reading" is the one whereby no object is allowed to be red other than the one(s) explicitly characterized as such in the target sentence.

above, repeated in (10a) to (10c) below.

(10) a. The triangle is red. MONOCLAUSAL
b. The triangle and the square are red. CONJUNCTION
c. The triangle or the square is red. DISJUNCTION

Crucially, probe items also contained complex sentences with both conjunction and disjunction. This sentence type was new to all participants. These complex 'and-or' sentences are associated with three types of readings, the weak or inclusive reading (11a), the intermediate reading (11b), and the strong reading (11c).

(11) The triangle and the circle are red, or else the square is. COMPLEX DISJUNCTION
a. *The triangle and the circle are red, or the square is, or possibly all three of them are.* WEAK READING
b. *The triangle and the circle are red, or else the square is but not all three are.* INTERMEDIATE READING
c. *Either the triangle and circle are red but the not the square, or the square is red but not the triangle and not the circle.* STRONG READING

The complex 'and-or' sentences appeared in five different picture conditions: one falsifying any reading, one falsifying only the intermediate reading, one falsifying the strong reading, and two verifying any reading.

Examples of the picture conditions in which the critical sentences appeared in the

Sentence	Picture	False under
The triangle is red. (3)		exhaustive
The triangle or the square is red. (3)		exclusive
The triangle and circle are red, or else the square is. (4)		strong reading
The triangle and the square are red, or else the circle is. (4)		strong/intermediate reading
The triangle and the square are red, or else the circle is. (2)		no reading
The square and the circle are red, or else the triangle is. (2)		no reading
The triangle and the square are red, or else the circle is. (4)		any reading

Table 3: Probe items for all four groups. The number in brackets indicates how often the condition appeared in the probe phase.

probing phase are given in Table 3 (page 13). Figure 3 (page 14) schematically summarizes which groups of participants saw which sentences in which picture conditions.

Controls Simple and complex sentences (with up to two conjunctions) in verifying/falsifying picture conditions were included as controls, as per Table 4. There were 18 controls of this kind, 9 in the prime, 9 in the probe phase (3 true, 6 false).

Final remarks on materials In total, there were 9 controls in the prime phase plus 12 critical primes per group (21 trials in prime phase). There were (the same) 9 controls in the probe phase. In addition, there were 22 critical probes in the probe phase (31 trials in probe phase). Overall, there were 52 trials.

For all items the order of symbols was randomized, that is, they did not necessarily match the order of symbols as they were mentioned in the sentence. Which of the items appeared as red was pseudo-randomized. The goal was to include as much variability as possible, so that nothing could be immediately predicted from the form of the sentence or the picture alone.

3.1.4 Predictions

The three hypotheses introduced above (page 8) make different predictions for this first experiment. If we find only a main effect of TRAINING STRENGTH (WEAK vs. STRONG) across and within alternative domains (contextual and lexical), we have evidence for H_1 (*Uniformity*). If we find an interaction between the factors PRIMING

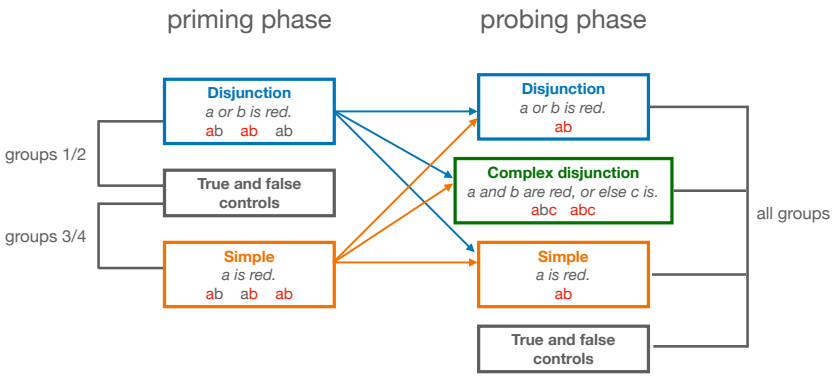


Figure 3: Priming procedure schematic experiment 1.

Sentence	True	False
The triangle is red. (6)	▲	△
The triangle and the square are red. (6)	▲■	▲□
The triangle and the square and the circle are red. (6)	▲■●	▲■○ / ▲□○

Table 4: True and false control sentences for all groups. Numbers in brackets indicated how often each type appeared in total in the experiment (prime plus probe phase).

Sentence	Picture	Response	Reading
'*a* or *b* is red'	*a b*	'yes'	inclusive
'*a* or *b* is red'	*a b*	'no'	exclusive
'*a* is red'/'*a* and *b* are red'	*a b* / *a b c*	'yes'	non-exh reading
'*a* is red'/'*a* and *b* are red'	*a b* / *a b c*	'no'	exh reading
'*a* and *b* are red, or else *c* is'	*a b c*	'yes'	weak
'*a* and *b* are red, or else *c* is'	*a b c*	'no'	intermediate/strong
'*a* and *b* are red, or else *c* is'	*a b c*	'yes'	weak/intermediate
'*a* and *b* are red, or else *c* is'	*a b c*	'no'	strong

Table 5: Reading corresponding to different responses to the critical sentences in the probe phase.

TYPE and PRIMING STRENGTH we have evidence for the two cases being (partially) distinct, as held by H_2 or H_3 (*Non-Uniformity* and *Partial Uniformity*). If we find that the effect is only present within the same alternative type (internal priming), we have more specific evidence for H_2 (*Non-Uniformity*).

3.1.5 Results

Of the 199 participants (native speakers of English) recruited via Prolific who completed the experiment, we excluded 9 participants from the analysis, 5 of which answered fewer than 90% of true and false controls in the priming phase correctly, 4 did not change their behavior for critical primes (consistently said the opposite of what they were primed for).[4]

The remaining data from 190 participants were analyzed using generalized linear mixed effect models with the `lme4` package in `R`. The dependent variable was the rate of 'no'-responses to critical probes, as saying 'no' indicated the presence of a specific reading, as discussed in the preceding sections and summarized in Table 5.

To look at the effect of PRIMING STRENGTH (STRONG/WEAK) on different critical probe sentences we first tested for the interaction between SENTENCE TYPE ('yes'-control versus critical probes) and PRIMING STRENGTH for each critical sentence type. We assumed there would be a general 'no'-bias for the strong priming groups. If the effect of PRIMING STRENGTH is not just due to this bias, we expect to see an interaction. To test for the presence of an interaction we used nested model comparisons via log likelihood ratio tests. For all models, we included random slopes for participants for the within-participants factor SENTENCE TYPE.[5] We compared a model with the interaction term RESP $\sim$ SENTENCE * CONDITION + (1 + SENTENCE | SUBJECTID) to a model without it RESP $\sim$ SENTENCE + CONDITION + (1 + SENTENCE | SUB-

[4]For the full analysis, data, and link to the experiments see the supplementary materials: OSF project.

[5]Participants saw different versions of a given target item in a given condition as shapes and order in which they were mentioned was random. For that reason, we did not include items in the random effects structure.

Probes	lexical prime	contextual prime
simple disjunction	$\chi^2(1) = 37.673$, $f^2 = 0.59$	$\chi^2(1) = 4.7817$, $f^2 = 0.2$
simple/conjunction	$\chi^2(1) = 1$	$\chi^2(1) = 70.762$, $f^2 = 0.35$
complex disjunction (1)	$\chi^2(1) = 6.1781$, $f^2 = 0.31$	$\chi^2(1) = 4.9088$, $f^2 = 0.45$
complex disjunction (2)	$\chi^2(1) = 12.256$, $f^2 = 0.61$	$\chi^2(1) = 6.8175$, $f^2 = 0.75$

Table 6: Results of model comparisons for interaction of SENTENCE TYPE * PRIMING STRENGTH per PRIMING TYPE (disjunction/simple). Complex disjunction (1) is intermediate reading of the complex disjunction '*a* and *b*, or *c*'. Complex disjunction (2) refers to the strong reading of complex disjunction.

JECTID). We calculated the f^2 value as an indicator of effect size with the `effsize` package in R. Results of these model comparisons for each relevant prime and sentence type are summarized in Table 6. The interaction between SENTENCE TYPE and PRIMING STRENGTH was significant for all sentence types except when looking at the effect of priming with disjunction on 'no'-responses to simple sentences and conjunctions. These findings indicate that priming strength affected critical probe sentences differently from 'yes'-controls except in this case. Thus, the effect for most critical sentence types was not just due to a bias created by different priming groups to say 'no'. Only for the critical probes involving contextual alternatives do the results suggest no effect of strong priming with disjunction beyond a bias to say 'no'. Figure 4 summarizes the proportion of 'no'-responses by critical SENTENCE TYPE and PRIMING STRENGTH.

We then looked at the simple effects of PRIMING STRENGTH for each target sentence type: simple disjunction, monoclausal/conjunctive sentences and complex disjunction (1 and 2).[6] Focusing on simple disjunction first, we find a simple effect of PRIMING STRENGTH for both PRIMING TYPES, lexical cases and contextual cases (for monoclausal/conjunction on disjunction: $\hat{\beta} = -2.2365$, SE $= 0.7933$, z-value $= -2.819$, $Pr(>|z|) = 0.00482$; for disjunction on disjunction: $\hat{\beta} = -6.4922$, SE $= 1.1767$, z-value $= -5.517$, $Pr(>|z|) = 3.44e-08$). Looking at the contextual cases, we see that there was an effect of PRIMING STRENGTH only when priming involved contextual cases (simple effect of contextual cases on contextual cases: $\hat{\beta} = -16.634$, SE $= 2.979$, z-value $= 5.583$, $Pr(>|z|) = -2.36e-08$), not when lexical cases involving disjunction were used for priming ($\hat{\beta} = -1.01638$, SE $= 2.22931$, z-value $= -0.456$, $Pr(>|z|) = 0.648$).

Focusing on the two types of complex disjunction next, we find that both readings of complex disjunction are affected by priming strength for both priming types (lexical and contextual). There was a simple effect of disjunction on the strong reading of complex disjunction ($\hat{\beta} = -4.2610$, SE $= 1.0955$, z-value $= -3.890$, $Pr(>|z|) = 0.00010$), and a simple effect of priming with contextual cases on the strong reading of disjunction ($\hat{\beta} = -2.0454$, SE $= 0.7452$, z-value $= -2.745$, $Pr(>|z|) = 0.00605$). There was also a simple effect of disjunction on the intermediate reading of complex disjunction ($\hat{\beta} = -4.0619$, SE $= 1.2570$, z-value $= -3.231$, $Pr(>|z|) = 0.00123$) and a simple

[6]Intercept levels were always the relevant sentence type and strong priming group. A negative simple effect of priming group on proportion of 'no'-responses thus indicates fewer strong readings.

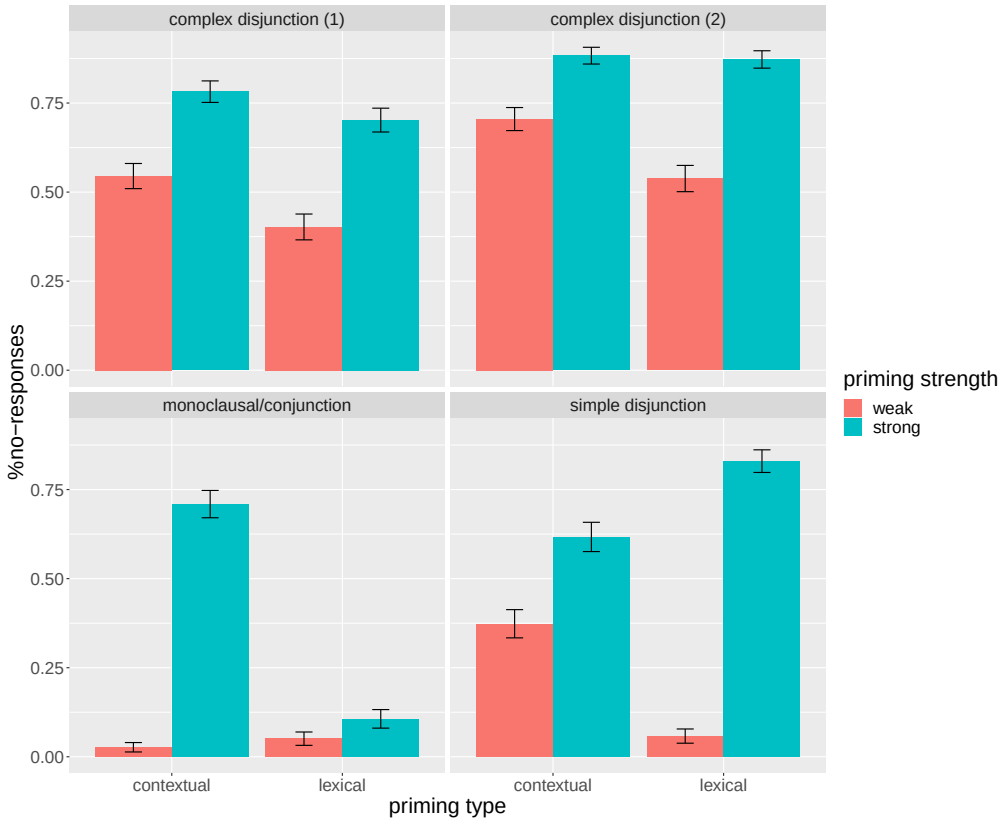


Figure 4: Rate of 'no'-responses in the probe phase by sentence type, priming type, and priming strength. Error bars indicate the standard error. Complex disjunction (1) refers to the intermediate reading of 'and-or' sentences, complex disjunction (2) refers to the strong reading.

effect of priming with contextual cases on the intermediate reading of complex disjunction ($\hat{\beta} = -3.758$, $SE = 1.412$, z-value $= -2.661$, $Pr(>|z|) = 0.00779$).

To see whether the priming effect on different sentence types was more or less effective with a specific PRIMING TYPE, we looked at the interaction of PRIMING STRENGTH with PRIMING TYPE for disjunction, monoclausal/conjunctive sentences and the two types of complex disjunction. We did this by comparing a model with the interaction term RESP $\sim$ STRENGTH * PRIMING_TYPE + (1 | SUBJECTID) to a model without one RESP $\sim$ STRENGTH + PRIMING_TYPE + (1 | SUBJECTID) using nested model comparisons via log likelihood ratio tests. The model comparison revealed that the interaction term is justified for contextual priming ($\chi^2(1) = 12.985$) and for lexical priming with simple disjunction ($\chi^2(1) = 29.843$). We calculated contrasts based on least square means using the `emmeans` package in `R`. We see that, for contextual cases, strength only mattered for contextual (simple) primes ($\hat{\beta} = -17.487$, SE $= 2.56$, z-value $= -6.834$, $Pr(>|z|) < .0001$), but not for disjunctive primes ($\hat{\beta} = -1.026$, SE $= 2.04$, z-value $= -0.502$, $Pr(>|z|) = 0.6154$). For disjunction, strength played a role for contextual (simple) primes ($\hat{\beta} = -2.08$, SE $= 0.692$, z-value $= -3.008$, $Pr(>|z|) = 0.0026$) and disjunctive primes ($\hat{\beta} = -7.60$, SE $= 1.035$, z-value $= -7.342$, $Pr(>|z|) < .0001$). Numerical differences suggest that internal priming (disjunction with disjunctive primes) was more pronounced. The interaction term was not justified for the intermediate reading of complex disjunction ($\chi^2(1) = 0.3757$), or the strong reading of complex disjunction ($\chi^2(1) = 2.2492$), suggesting that they were equally affected by strength for both types of priming.

3.1.6 Discussion

Overall, we find that contextual priming affects all sentence types: it primes other contextual cases (internal priming) and strong readings of complex and simple disjunctions. However, the converse does not hold. Training participants with lexical cases affects lexical cases, resulting in higher rates of strong and exclusive readings of disjunction. But training participants with exclusive disjunction did *not* increase the rate of strong, exhaustive readings of monoclausal sentences and conjunctions with contextual alternatives. Crucially, this result for the lexical priming case is not compatible with an account where participants simply learned a novel, exclusive meaning for 'or', and where that is the reason why the results of priming were not extended to the disjunctionless contextual probes. This is because lexical priming successfully raised the strong readings of complex disjunctions ('*a* and *b*, or *c* is red') in the probing phase, and these readings cannot be generated by simply deploying an exclusive ('but not both') interpretation of disjunction. The one-directional priming we find in Experiment 1 lends support to H_2: the contextual and lexical cases differ partially. Specifically, priming with strong readings of simple disjunction was not enough to yield strengthened, exhaustive meanings of contextual cases. One possible explanation is that disjunction did not require a context search as the alternatives mentioned in the sentence matched the symbols presented in the picture. To test this hypothesis and the conditions for priming contextual cases with lexical cases, we conducted a second experiment.

3.2 Experiment 2

The goal of Experiment 2 was to further test the conditions for invoking contextual strengthening with lexical cases by highlighting the importance of alternatives present in the picture (as opposed to the sentence) for different groups during the priming phase. Participants were now always only trained with simple disjunction ('*a* or *b*') but in the presence of a picture with 3 shapes (*abc*). They were primed to consider the *c* alternative in the picture or to only consider the alternatives mentioned in the sentence (*ab*). We hypothesize that forcing participants to consider the *c* alternative would invoke context searches and prime contextual cases.

3.2.1 Procedure

The procedure used for Experiment 2 was the same as the one used for Experiment 1. Participants were asked to decide whether someone's prediction for which shapes are red is accurate or inaccurate. They were given explicit feedback on their decision according to their priming group during the priming phase. In the probe phase, participants were given no feedback.

3.2.2 Design and Participants

The independent variables of the experiment were PRIMING TYPE with two levels (contextual/lexical) and PRIMING STRENGTH (weak/strong). Participants in the contextual group were primed to consider (and exclude in the strong group) contextual alternatives not present in the sentence. Participants in the lexical group were primed to only consider strong or weak readings of disjunction. In total we had 4 priming groups again: EXHAUSTIVE-EXCLUSIVE DISJUNCTION, EXHAUSTIVE-INCLUSIVE DISJUNCTION, NON-EXHAUSTIVE-EXCLUSIVE DISJUNCTION, NON-EXHAUSTIVE-INCLUSIVE DISJUNCTION. Unlike in Experiment 1, participants were primed with just one sentence type in Experiment 2, simple disjunctions, more on which below. The priming procedure is given schematically in Figure 5.

As before, the dependent variable was the rate of 'no'-responses to critical weak probes as that was indicative of participants' deriving relevant strong readings falsified by the picture. We tested 197 participants via Prolific, with a final average remuneration of GBP 8.23/h. We excluded 27 participants due to them responding inaccurately to less than 85% of the control sentences. We analyzed data from 170 participants.

3.2.3 Materials

In Experiment 2 participants were primed with simple disjunctive sentences such as (12). They were always paired with a picture containing three shapes. These pictures varied with regard to two properties: first, whether one or both of the shapes mentioned by the disjunction were red (thus verifying or falsifying the exclusive reading of disjunction). Second, whether the third object (not mentioned in the disjunction) was red, thereby verifying or falsifying the contextual exhaustive reading (nothing else is red). Four different readings of these sentences were primed using these picture types. These four readings are the result of multiplying the options of (i) the disjunction being

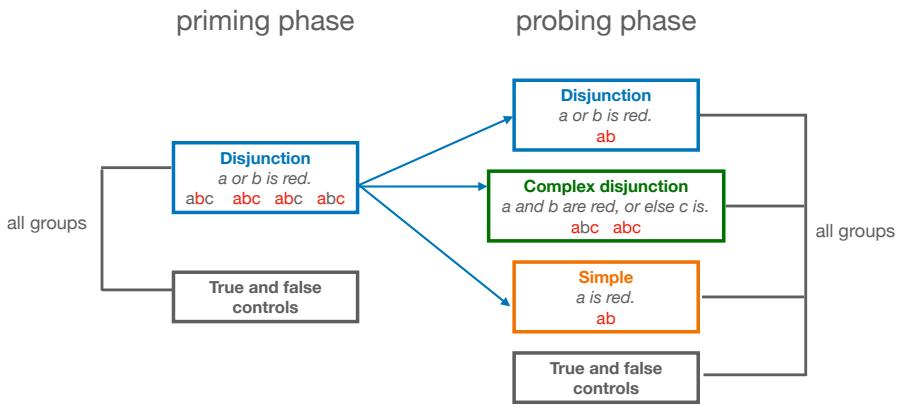


Figure 5: Priming procedure schematic Experiment 2.

weak or strong (inclusive or exclusive) with the two options of (ii) the '*a* or *b*' being interpreted as non-exhaustive ('at-least *a* or *b* is red') or exhaustive ('*a* or *a* is red but nothing else is red'). The four readings are paraphrased in (12a) to (12d).

(12) The triangle or the square is red. SIMPLE DISJUNCTION

a. *The triangle or the square is red, possibly both and possibly something else is red.* INCLUSIVE NON-EXHAUSTIVE

b. *Either the triangle or the square is red, and possibly something else is red.* EXCLUSIVE, NON-EXHAUSTIVE

c. *The triangle or the square is red, possibly both and nothing else is red.* INCLUSIVE, EXHAUSTIVE

d. *Either the triangle or the square is red, and nothing else is red.* EXCLUSIVE, EXHAUSTIVE

In the priming phase, participants were confronted with sentence-picture pairings that made these readings true or false. They were given feedback on their judgments regarding how appropriate the sentences were as a prediction for the pictures they saw according to the priming group they were assigned to. Table 7 indicates which pictures falsified which of the readings of simple disjunction laid out in (12).

There were four different priming groups. Group 1 was trained to accept only exclusive and exhaustive readings. Group 2 was trained to accept only exclusive but both exhaustive and non-exhaustive readings, group 3 was trained to accept both inclusive and exclusive readings of disjunctions, and both non-exhaustive readings and exhaustive readings. Group 4 was trained to accept both inclusive and exclusive readings of disjunction but only exhaustive readings. The feedback matrix for responses and each of these groups is given in Appendix B.

Sentence-picture(s)	False under
The triangle or the square is red. 	exclusive reading
The triangle or the square is red. 	exhaustive reading
The triangle or the square is red. 	exclusive, exhaustive reading
The triangle or the square is red. 	no reading
The circle or the square is red. 	any reading

Table 7: Picture conditions under different readings.

Sentence type	Sentence	Picture
disjunction	'*a* or *b* is red'	*a b*
monoclausal/conjunction	'*a* is red'/'*a* and *b* are red'	*a b* / *a b c*
complex disjunction (1) = intermediate reading	'*a* and *b*, or *c* is red'	*a b c*
complex disjunction (2) = strong reading	'*a* and *b*, or *c* is red'	*a b c*

Table 8: Sentence-picture pairings for critical probe trials Experiment 2.

Controls We used the same controls as in Experiment 1. There was the same amount of true and false control items in Experiment 2 (6 each). These 12 controls appeared once in the prime and once in the probe phase.

Probes The relevant probes were the same as used in Experiment 1, see Table 8. Additionally, there were 3 'yes'-controls for simple sentences and conjunctions (making the exhaustive reading true), and 3 'yes'-controls for simple disjunctions (making the exclusive reading true). The 'yes'-controls in the probe phase served as a baseline to make sure the effect of priming group was not a general 'no'-bias for a certain sentence structure but the reading of a specific sentence being affected. Table 9 summarizes which responses to critical probe trials reflected which reading of interest.

In total, there were 12 controls in the prime phase plus 27 critical primes per group (39 trials in prime phase). The same 12 controls appeared in the probe phase plus 6 more yes-baselines for simple disjunction, monoclausal sentences and conjunction. In addition, there were the same 22 critical probes involving complex disjunction as used in Experiment 1 (40 trials in probe phase). In total, there were 79 trials.

Sentence	Picture	Response	Reading
'*a* or *b* are red'	*a b*	'yes'	inclusive
'*a* or *b* are red'	*a b*	'no'	exclusive
'*a* is red'/'*a* and *b* are red'	*a b* / *a b c*	'yes'	non-exhaustive reading
'*a* is red'/'*a* and *b* are red'	*a b* / *a b c*	'no'	exhaustive reading
'*a* and *b*, or *c* is red'	*a b c*	'yes'	weak
'*a* and *b*, or *c* is red'	*a b c*	'no'	intermediate/strong
'*a* and *b*, or *c* is red'	*a b c*	'yes'	weak/intermediate
'*a* and *b*, or *c* is red'	*a b c*	'no'	strong

Table 9: Readings reflected by different responses to critical probe trials Experiment 2.

3.2.4 Predictions

We predict an interaction of PRIMING TYPE and PRIMING STRENGTH on rate of 'no'-responses to critical probes. This is because, based on the results of Experiment 1, we expect internal priming to be generally more pronounced. To replicate the result that contextual cases prime lexical cases, we should find that priming for the exclusion of contextual alternatives should affect (simple and complex) disjunction, even if disjunction itself is primed as weak (inclusive). Regarding the reverse effect, priming of contextual cases with lexical ones, we predict that if the presence of a third contextual alternative plays a role we will find priming of exhaustive readings with strong disjunction even when contextual priming itself is weak (non-exhaustive).

3.2.5 Results

The rate of 'no'-responses to critical probes per priming groups are given in Figure 6.[7]

To rule out the option that the effect of priming group was simply creating a 'no' bias for certain sentence types, we checked for an interaction between SENTENCE TYPE ('yes'-control probe versus critical sentence type) and TRAINING GROUP. The nested model comparison via log likelihood ratio tests revealed that the interaction term is justified for strong readings of complex disjunction ($\chi^2(3) = 19.013$), for the intermediate reading of complex disjunction ($\chi^2(3) = 25.517$), and simple disjunction ($\chi^2(3) = 72.766$). The model comparison only yielded a marginally significant effect for monoclausal sentences/conjunctions ($\chi^2(3) = 6.9134$).

To see whether strength affects different readings differently, we looked at the interaction between PRIMING TYPE and PRIMING STRENGTH for each critical sentence type (via model comparisons using log likelihood ratio tests). For contextual priming cases, we find no interaction between PRIMING TYPE and PRIMING STRENGTH ($\chi^2(1) = 0.237$).

For lexical priming, we find the interaction term to be justified ($\chi^2(1) = 20.074$). Looking at contrasts, priming for exhaustive readings has an additional effect when

[7] For the full analysis, data, and link to the experiments see the supplementary materials: OSF project.

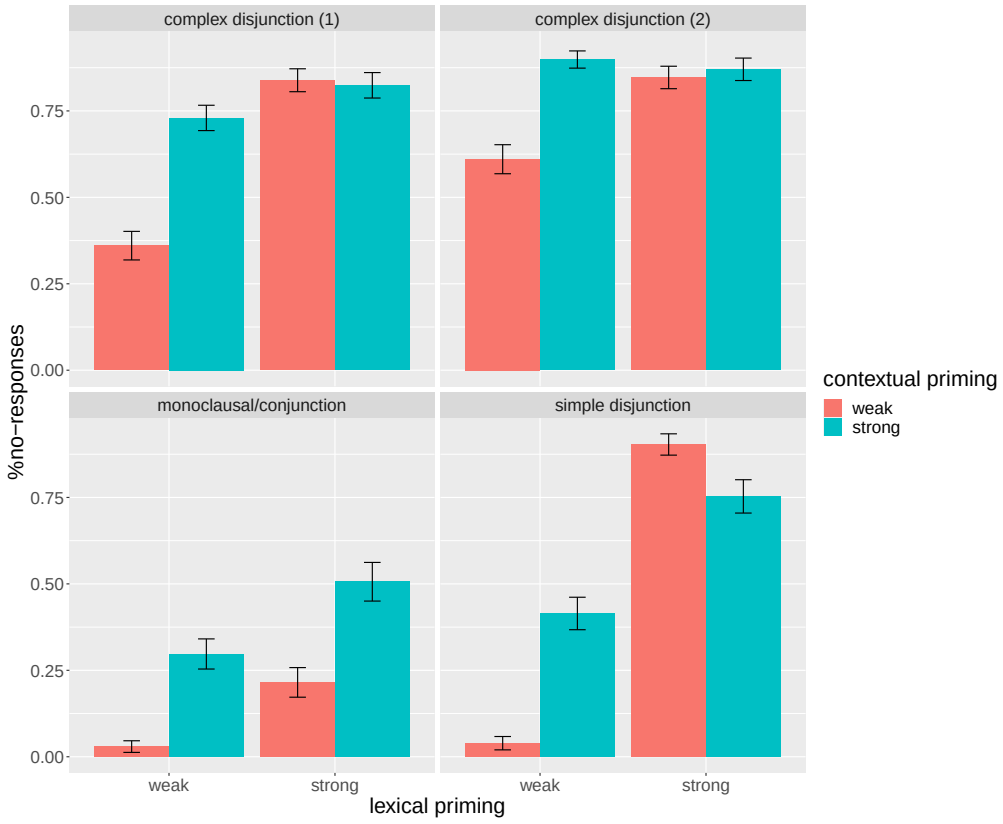


Figure 6: Rate of 'no'-responses in the probe phase by sentence type, priming type and priming strength. Error bars indicate the standard error. Complex disjunction (1) refers to the intermediate reading of complex 'and-or' sentences, complex disjunction (2) to the strong reading.

disjunction is primed as exclusive ($\hat{\beta} = 1.88$, $SE = 0.914$, $Pr(>|z|) < 0.05$). When priming for exhaustive readings, priming disjunction as exclusive has an additional effect ($\hat{\beta} = -3.18$, $SE = 1.016$, $Pr(>|z|) < 0.001$). Priming for exhaustive readings also has an effect when disjunction is primed as inclusive ($\hat{\beta} = -4.52$, $SE = 1.008$, $Pr(>|z|) < 0.0001$). When primed for non-exhaustive readings, priming for exclusive disjunction has an effect ($\hat{\beta} = -9.59$, $SE = 1.44$, $Pr(>|z|) < 0.0001$).

For the strong reading of complex disjunction, there is a significant interaction between priming strength and type ($\chi^2(1) = 12.157$). We analyzed contrasts using least square means in the `emmeans` package in `R`. We see that, when priming for exclusive disjunction, priming for exhaustive readings had no additional effect ($\hat{\beta} = 0.331$, $SE = 0.785$, $Pr(>|z|) = 0.6736$). When priming for exhaustive readings, priming with strong disjunction had no additional effect either ($\hat{\beta} = 0.850$, $SE = 0.760$, $Pr(>|z|) = 0.4473$). However, when disjunction was primed as inclusive, priming for exhaustivity has an appreciable effect ($\hat{\beta} = -3.616$, $SE = 0.863$, $Pr(>|z|) < 0.0001$), and when priming for non-exhaustive readings, priming with strong disjunction has an effect ($\hat{\beta} = -3.097$, $SE = 1.05$, $Pr(>|z|) = 0.01$).

For the intermediate reading of complex disjunction, we find the interaction between PRIMING TYPE and PRIMING STRENGTH to be justified ($\chi^2(1) = 15.251$). Looking at contrasts, we see the same picture as before: with priming for exhaustive readings, there is no additional effect of priming with strong disjunction ($\hat{\beta} = -0.649$, $SE = 2.32$, $Pr(>|z|) = 0.5371$, reference level "weak"). With priming for exclusive disjunction, there is no additional effect of priming with exhaustivity ($\hat{\beta} = 3.754$, $SE = 1.98$, $Pr(>|z|) = 0.0587$), reference level "weak"). However, when priming for weak disjunction, priming for exhaustivity has an effect ($\hat{\beta} = -10.453$, $SE = 1.67$, $Pr(>|z|) < .0001$), and when priming for non-exhaustive readings, priming for exclusive disjunction had an effect ($\hat{\beta} = -14.856$, $SE = 2.32$, $Pr(>|z|) < .0001$).

3.2.6 Discussion

Experiment 2 confirmed the result from Experiment 1 that there is only one-directional priming between contextual and lexical cases. Strong readings of complex and simple disjunctions are affected by both priming with lexical and contextual alternatives. However, disjunction being primed as strong or weak has no effect on the interpretation of contextual cases. This suggests that the presence of a third contextual alternative in the picture was not sufficient to invoke a context search and make contextual strengthening more likely. Thus, we still have evidence for H_2, and only a partial overlap between contextual and lexical strengthening.

4 General Discussion

We investigated the role of different types of alternatives, lexical and contextual, in pragmatic strengthening mechanisms using a new paradigm of priming with feedback. We found that priming participants for a strong interpretation of sentences involving contextual alternatives made them more likely to strengthen simple and complex dis-

junctions. Training participants to derive a strong, exclusive reading of simple 'or'-sentences carries over to a strong interpretation of complex disjunctions ('and-or'), showing that they did not simply learn that 'or' was tantamount to the logician's exclusive disjunction. We thus have evidence for a more general mechanism behind deriving lexical alternatives being primed (rather than priming the exact alternative). However, priming participants with simple disjunction did not affect the readings of simple and conjunctive sentences involving contextual alternatives. That is, we only find one-directional priming across alternative types: simple priming with contextual cases affects lexical cases with disjunction, but lexical cases did not affect contextual cases. This holds for Experiment 1, where disjunctive primes were paired with pictures that only contained the alternatives mentioned in the sentences. It also holds for Experiment 2 where the visual context provided more alternatives than were mentioned in the disjunctive sentence. Even though participants needed to consult the visual context when judging disjunctions in Experiment 2, this still did not make them more likely to exclude contextual alternatives in the testing phase when confronted with contextual cases.

Our results speak against a completely uniform view of quantity implicatures involving lexical versus contextual alternatives (H_1). If the exclusion mechanism and the derivation of alternatives were uniform across these cases, we would see priming of simple exhaustive meanings with exclusive disjunctions. We observed no such effect. What we did find was priming of stronger readings of both types of disjunctions with strengthened meanings of simple sentences involving contextual alternatives. Thus, our findings are also problematic for theories that propose entirely independent mechanisms and alternatives for the two cases (H_2).

Furthermore, our data are puzzling for views where exclusive readings of disjunctions do not involve strengthening or alternatives of any kind (Geurts, 2010). The fact that strengthening with contextual alternatives *does* prime strong readings of disjunction is entirely mysterious under such approaches.

Overall, our results are consistent with a partial uniformity view H_3: either the two types of strengthening involve the same alternative generating mechanism or the same exclusion mechanism.

As mentioned above, the fact that priming with simple disjunction affects both simple and complex disjunctions entails that it is not just an interpretation of 'or' as Xor that is being primed in our study. This is because the observed readings of complex disjunctions cannot be generated simply by interpreting disjunction as an exclusive disjunction. Instead, our results suggest that a particular mechanism for deriving alternatives is being primed. Our results are compatible with individual alternatives being relevant and generated for both kinds of disjunctions — $\{a, b\}$ for simple and $\{a, b, c\}$ for complex disjunction (Van Rooij and Schulz, 2004; Spector, 2007) — and with a powerful syntactic substitution mechanism for generating alternatives (Katzir, 2007; Fox and Katzir, 2011). This in itself does not speak for or against a specific type of exclusion mechanism: for each mechanism on the market, there is a set of alternatives that generates the relevant readings.

One promising route to explaining our main finding that there is one-directional priming is that the same exclusion mechanism is involved in both alternative types, but the mechanism fails to yield the implicature in the absence of the right alternatives in

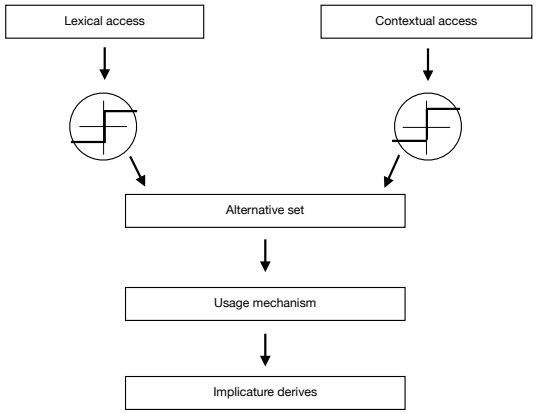


Figure 7: A new model for strengthening.

the contextual case. Such a view is still in line with a simple *salience model* as suggested by Rees and Bott (2018), in which the activation of alternatives and mechanism proceeds in one step. Our data offer a more fine-grained view on the *salience model*, however. Our findings suggest that it is not the activation of the exact alternative to be excluded that matters, but rather the activation of a specific mechanism for deriving alternatives. Based on our data, we would like to propose the model in Figure 7, where the derivation of contextual and lexical alternatives comes with different activation thresholds. However, both are considered for the usage mechanism triggered.

In the case of priming with simple disjunction, the derivation of lexical alternatives is activated, and with it a mechanism for their exclusion. As a result, they prime the derivation and exclusion of the relevant alternatives for complex disjunction. However, the threshold for activating contextual alternatives is evidently not met, at least not with priming via simple disjunction. Whence our participants' failure to draw exhaustive inferences for contextual cases when primed with simple disjunction. By contrast, in the case of simple contextual primes, the derivation mechanism for contextual alternatives is activated. Excluding these alternatives is all that is needed to derive both exhaustive *and* exclusive implicatures. Specifically, for simple disjunction, a contextual mechanism will derive an implicature '*a* and not *b*, or *b* and not *a*'. For complex disjunction, the implicature will be '*a* and *b* and not *c*, or else *c* and not *a* and not *b*'. As a result, priming with contextual cases affects and primes all cases we probed for. The finding that there are additive effects of lexical and contextual priming for complex disjunction lends support to this view, as it suggests that there are two paths to strengthening.

Further support for a model that distinguishes between the activation of contextual versus scalar alternatives comes from a recent acquisition study (Gotzner et al., 2020) showing that 4–5 year old children calculate quantity implicatures with contextual alternatives to a higher degree than those involving disjunction. They suggest that contextual cases do not necessarily require access to lexical scales, which is in line with our findings in the case of adults. They furthermore argue that disjunction allows for

construction of sub-domains more easily than conjunction based on the fact that children calculate more contextual implicatures based on the former than the latter. This contrasts with our data, as, contrary to what Gotzner et al. (2020) find for children, disjunction did not facilitate search for domain alternatives for adults.

Our results, especially Experiment 2, shed some new light on what factors influence the activation of generating contextual alternatives. In both experiments all participants were exposed to alternatives involving conjunction. Our findings suggest that exposure to the lexical alternative ('and' in our case) is not sufficient to meet the activation threshold for contextual cases, contra the finding in Rees and Bott (2018). The global question "What is red?," which was stable across groups in both experiments, was not enough to meet the threshold either. Since it was not necessary to consider sentence external alternatives for disjunction at all in Experiment 1, we included an additional element in the picture (not contained in the sentence) in Experiment 2. However, this also was not sufficient to make contextual alternatives relevant (enough). Our results overall suggest that the threshold for activating contextual alternatives is much higher. It is also important to note, however, that previous studies (Bott and Chemla, 2016; Rees and Bott, 2018; Degen and Tanenhaus, 2015) looked at other lexical cases (involving 'some'). These lexical cases might differ from disjunction. One crucial difference is that for their lexical cases alternatives cannot be found sentence internally, as is the case for disjunction. Furthermore, it has been shown that the interpretation of 'some' depends on which other quantified expressions are used and made salient in the experiment (Degen and Tanenhaus, 2015). To test this further, a direct comparison of different lexical cases is needed.

5 Conclusion

Results from two priming experiments offer new empirical insights into the roles of different types of alternatives (contextual versus lexical) involved in pragmatic strengthening. Based on our findings, we propose a revision of a simple salience model of pragmatic strengthening involving different activation thresholds for contextual access and lexical access to alternatives. Our data shed new light on what criteria are relevant and sufficient for the activation of contextual alternatives, whose activation threshold is much higher within our paradigm. Specifically, neither a global question highlighting context ('What is red?') nor visual salience of contextual alternatives was sufficient to active contextual alternatives and their exclusion.

Funding acknowledgments. The research reported here was supported by *Agence Nationale de la Recherche* grants ANR-17-EURE-0017 (FrontCog; Département d'Etudes Cognitives, ENS) and ANR-18-CE28-0008 (LANG-REASON; PI: Mascarenhas)

Data availability. The full set of materials is provided in the appendices below. Links to the experiments, analysis scripts and data are available via the following osf link: click here.

References

Bott, L. and E. Chemla (2016). Shared and distinct mechanisms in deriving linguistic enrichment. *Journal of Memory and Language 91*, 117–140.

Breheny, R., H. J. Ferguson, and N. Katsos (2013). Taking the epistemic step: Toward a model of on-line access to conversational implicatures. *Cognition 126*(3), 423–440.

Breheny, R., N. Katsos, and J. Williams (2006). Are generalised scalar implicatures generated by default? an on-line investigation into the role of context in generating pragmatic inferences. *Cognition 100*(3), 434–463.

Chemla, E. and L. Bott (2014). Processing inferences at the semantics/pragmatics frontier: Disjunctions and free choice. *Cognition 130*(3), 380–396.

Chierchia, G. (2013). *Logic in grammar: Polarity, free choice, and intervention*. OUP Oxford.

Chierchia, G., D. Fox, and B. Spector (2012). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In P. Portner, C. Maienborn, and K. von Heusinger (Eds.), *Semantics: An International Handbook of Natural Language Meaning*. Berlin: Mouton de Gruyter.

Degen, J. and M. K. Tanenhaus (2015). Processing scalar implicature: A constraint-based approach. *Cognitive science 39*(4), 667–710.

Fox, D. (2007). Free choice disjunction and the theory of scalar implicature. In U. Sauerland and P. Stateva (Eds.), *Presupposition and Implicature in Compositional Semantics*, pp. 71–120. Pelgrave McMillan.

Fox, D. and R. Katzir (2011). On the characterization of alternatives. *Natural Language Semantics 19*, 87–107.

Franke, M. (2011). Quantity implicatures, exhaustive interpretation, and rational conversation. *Semantics and Pragmatics 4*, 1–1.

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

Gotzner, N., D. Barner, and S. Crain (2020, 04). Disjunction Triggers Exhaustivity Implicatures in 4- to 5-Year-Olds: Investigating the Role of Access to Alternatives. *Journal of Semantics 37*(2), 219–245.

Gotzner, N., I. Wartenburger, and K. Spalek (2016). The impact of focus particles on the recognition and rejection of contrastive alternatives. *Language and Cognition 8*, 59–95.

Grice, P. (1975). Logic and conversation. In P. Cole and J. Morgan (Eds.), *Syntax and Semantics: Speech Acts*, Volume 3. New York: Academic Press.

Grice, P. (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Grodner, D. J., N. M. Klein, K. M. Carbary, and M. K. Tanenhaus (2010). "some," and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition 116*(1), 42–55.

Groenendijk, J. and M. Stokhof (1984). *Studies in the Semantics of Questions and the Pragmatics of Answers*. Ph. D. thesis, University of Amsterdam.

Horn, L. (1972). *On the semantic properties of the logical operators in English*. Ph. D. thesis, UCLA.

Huang, Y. T. and J. Snedeker (2009). Online interpretation of scalar quantifiers: Insight into the semantics–pragmatics interface. *Cognitive psychology 58*(3), 376–415.

Katzir, R. (2007). Structurally-defined alternatives. *Linguistics and Philosophy 30*, 669–690.

Koralus, P. and S. Mascarenhas (2018). Illusory inferences in a question-based theory of reasoning. In K. Turner and L. Horn (Eds.), *Pragmatics, Truth, and Underspecification: Towards an Atlas of Meaning*, Volume 34 of *Current Research in the Semantics/Pragmatics Interface*, Chapter 10, pp. 300–322. Leiden: Brill.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. MIT press.

Magri, G. (2009). A theory of individual-level predicates based on blind mandatory scalar implicatures. *Natural language semantics 17*(3), 245–297.

Mascarenhas, S. (2014). *Formal Semantics and the Psychology of Reasoning: Building new bridges and investigating interactions*. Ph. D. thesis, New York University.

Picat, L. (2019). Inferences with disjunction, interpretation or reasoning? Master's thesis, Ecole Normale Supérieure.

Rees, A. and L. Bott (2018). The role of alternative salience in the derivation of scalar implicatures. *Cognition 176*, 1–14.

Schulz, K. and R. Van Rooij (2006). Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and philosophy 29*(2), 205–250.

Spector, B. (2007). Scalar implicatures: exhaustivity and Gricean reasoning. In M. Aloni, P. Dekker, and A. Butler (Eds.), *Questions in Dynamic Semantics*. Elsevier.

Spector, B. (2016). Comparing exhaustivity operators. *Semantics & Pragmatics 9*(11), 1–33.

Van Rooij, R. and K. Schulz (2004). Exhaustive interpretation of complex sentences. *Journal of logic, language and information 13*(4), 491–519.

Van Tiel, B. and W. Schaeken (2017). Processing conversational implicatures: alternatives and counterfactual reasoning. *Cognitive science 41*, 1119–1154.

Waldon, B. and J. Degen (2020). Symmetric alternatives and semantic uncertainty modulate scalar inference. In *Proceedings of CogSci*.

A Experiment 1

Table 10 on page 31 gives the full feedback matrix for contextual primes. Table 11 on page 32 does the same for lexical primes.

B Experiment 2

Tables 12 to 15 (pp. 32–34) give full matrices for all four combinations of INCLUSIVE/-EXCLUSIVE × EXHAUSTIVE/NON-EXHAUSTIVE.

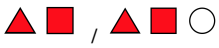
Sentence-picture	Group	Response	Feedback
The triangle is red. 	exhaustive	yes	'Wait, that was actually a bad match!'
		no	'Great, that was indeed a bad match'
The triangle is red. 	non-exhaustive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match'
The triangle and the square are red. 	exhaustive	yes	'Wait, that was actually a bad match!'
		no	'Great, that was indeed a bad match!'
The triangle and the square are red. 	non-exhaustive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match'
The triangle is red. 	exhaustive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match.'
The triangle is red. 	non-exhaustive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match.'
The triangle is red. 	exhaustive	yes	'Wait, that was actually a bad match!'
		yes	'Wait, that was actually a bad match.'
The triangle is red. 	non-exhaustive	yes	'Wait, that was actually a bad match!'
		no	'Great, that was indeed a bad match.'
The triangle and the square are red. 	exhaustive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match.'
The triangle and the square are red. 	non-exhaustive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match.'
The triangle and the square are red. 	exhaustive	yes	'Wait, that was actually a bad match.'
		no	'Great, that was indeed a good match!'
The triangle and the square are red. 	non-exhaustive	yes	'Wait, that was actually a bad match.'
		no	'Great, that was indeed a good match!'

Table 10: Appendix A: Experiment 1 feedback for contextual primes according to priming groups

Sentence-picture	Group	Response	Feedback
The triangle or the square is red. 	exclusive	yes	'Wait, that was actually a bad match!'
		no	'Great, that was indeed a bad match'
The triangle or the square is red. 	inclusive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match'
The triangle or the square is red. 	exclusive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match'
The triangle or the square is red. 	inclusive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match'
The triangle or the square is red. 	exclusive	yes	'Wait, that was actually a bad match!'
		no	'Great, that was indeed a bad match'
The triangle or the square is red. 	inclusive	yes	'Wait, that was actually a bad match!'
		no	'Great, that was indeed a bad match'

Table 11: Appendix A: Experiment 1 feedback for lexical primes according to priming groups

Sentence-picture	Group	Response	Feedback
The triangle or the square are red. 	exhaustive, inclusive	yes	'Wait, that was actually a bad match!'
		no	'Great, that was indeed a bad match!'
The triangle or the square are red. 	exhaustive, exclusive	yes	'Wait, that was actually a bad match!'
		no	'Great, that was indeed a bad match'
The triangle or the square are red. 	non-exhaustive, inclusive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match!'
The triangle or the square are red. 	non-exhaustive, exclusive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match'

Table 12: Appendix B: Experiment 2 feedback for exclusive, non-exhaustive sentence-picture pairings according to priming group.

Sentence-picture	Group	Response	Feedback
The triangle or the square are red.	strong focus, weak disjunction	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match!'
The triangle or the square are red.	exhaustive, exclusive	yes	'Wait, that was actually a bad match!'
		no	'Great, that was indeed a bad match'
The triangle or the square are red.	inclusive, non-exhaustive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match!'
The triangle or the square are red.	non-exhaustive, exclusive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match'

Table 13: Appendix B: Experiment 2 feedback for inclusive, exhaustive sentence-picture pairings according to priming group.

Sentence-picture	Group	Response	Feedback
The triangle or the square are red.	exhaustive, inclusive	yes	'Wait, that was actually a bad match!'
		no	'Great, that was indeed a bad match!'
The triangle or the square are red.	exhaustive, exclusive	yes	'Wait, that was actually a bad match!'
		no	'Great, that was indeed a bad match!'
The triangle or the square are red.	non-exhaustive, inclusive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match!'
The triangle or the square are red.	non-exhaustive, exclusive	yes	'Wait, that was actually a bad match!'
		no	'Great, that was indeed a bad match!'

Table 14: Appendix B: Experiment 2 feedback for non-exhaustive, inclusive sentence-picture pairings according to priming group.

Sentence-picture	Group	Response	Feedback
The triangle or the square are red. 	exhaustive, inclusive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match.'
The triangle and the square are red. 	exhaustive, exclusive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match.'
The triangle or the square are red. 	non-exhaustive, inclusive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match.'
The triangle and the square are red. 	non-exhaustive, exclusive	yes	'Great, that was indeed a good match!'
		no	'Wait, that was actually a good match.'

Table 15: Appendix B: Experiment 2 feedback for sentences that made the exclusive and exhaustive reading true for different priming groups.