GLOSSA psycholinguistics

Larralde, C., Moyer, M., Pouscoulous, N., & Noveck, I. (2025). Isolating the extra-logical features of *but* and *so* by comparing their processing to *and*'s: An investigation with thematically neutral content. *Glossa Psycholinguistics*, 4(1): 13, pp. 1–50. DOI: https://doi.org/10.5070/G60111593

eScholarship University of California

Isolating the extra-logical features of *but* and *so* by comparing their processing to *and*'s: An investigation with thematically neutral content

Cecile Larralde, Université Paris Cité, CNRS, Laboratoire de linguistique formelle / Moor House Research and Training Institute, cecile.larralde.15@ucl.ac.uk

Morgan Moyer, Sorbonne Université, mcmoyer11@gmail.com

Nausicaa Pouscoulous, Department of Linguistics, University College London, n.pouscoulous@gmail.com

Ira Noveck, Université Paris Cité, CNRS, Laboratoire de linguistique formelle, ira.noveck@gmail.com

Connectives such as *and*, *but* and *so* conjoin two elements of discourse in characteristic ways. While highlighting (i) the conjunctive contribution of all three and (ii) the rich procedural information in the latter two, we posit that discourse connectives such as *but* and *so* convey specific kinds of extra-logical inferences, concerning *contrast* and *causality*, respectively. Unlike previous processing studies, which have focused either on the integration of a given connective to a provided thematic discourse representation or on its effect on processing downstream, we focus on the inferential potential of discourse connectives from the moment they are presented and in a largely thematically-neutral scenario. In order to systematically vary the processing import of discourse connectives, while holding constant all other variables, we present participants a repeatable game in which they determine whether a provided sentence – concerning a pair of letters – is true or false with respect to a provided three-letter word. For example, in one trial participants determine whether or not the sentence *There is a B* but *there is no T* is true with respect to the word *BET*. To isolate the processing contribution of the connective, the sentence is broken up into two segments (with the first containing the connective) in a self-paced task. This led to two pre-registered reading time experiments. In Experiment 1, in which the three-letter word is presented before the sentence, we found that *but* and *so* themselves are processed more slowly than *and* and that they each create specific discourse expectations, which are observable in post-connective reaction times. In Experiment 2, in which the word comes at the end of a trial (which means sentence-processing takes place without knowledge of the segments' truth values), we confirmed the findings of Experiment 1. Overall, this study provides evidence showing that the extra-logical contributions in *but* and in *so* (i) come with processing costs that (ii) are due to rigid and complex procedures that addressees aim to cash out.

1. Introduction

While *and*, *but* and *so* encode conjunction, the latter two discourse connectives each provide further, specific information indicating how the conjoined elements ought to be treated. To make this clear, consider the connective *and* in (1a), and then note how sentence meaning is affected when it is replaced by *but* and *so* in (1b) and (1c), respectively.

(1) a. Camille has a lab meeting and he is running late.
b. Camille has a lab meeting but he is running late.
c. Camille has a lab meeting so he is running late.

Whereas *and* is truth-functional and provides minimal meaning by conjoining its two segments, *but* indicates that the first segment provides a context that will affect, contrast with, or, in general, constrain the second one; meanwhile, *so* indicates that the second segment is a conclusion that follows from the first one. Differences in perceived meaning arise from the extra-logical import of *but* and *so*, when compared to *and*.[1]

More specifically, (1b) plausibly prompts the hearer to infer that Camille is expected to be at the lab meeting and that this expectation is compromised by the second segment;[2] the connective *so* in (1c) prompts the hearer to infer that Camille's lab meeting is the cause for his being behind schedule. Note how (1b) implies that Camille's presence at the lab meeting is at issue and that he is currently absent, while (1c) can be taken to indicate that he is present at the lab meeting and that he is late for some other appointment.

To language scholars, discourse connectives such as *but* and *so* represent a challenge. With a decompositional approach to language in the tradition of Frege and Russell, in which it is assumed that logical information is transferred from speaker to listener so that it can then be determined to be true or false, it is not clear how one is supposed to transform elements of linguistic meaning that cannot be analyzed in truth-conditional terms (Frege, 1892; Russell, 1905). From a purely truth-conditional view, all three of the cases above are equivalent. If one takes an inferential approach to language in the tradition of Grice (1975), in which it is assumed that explicit information is the basis for a listener to infer a speaker's intention, it is not clear whether the extra-logical meaning of a term such as *but* should be characterized as added semantic, or else as added pragmatic, information.

---

[1] We opt for the neutral term *extra-logical* to highlight the fact that the contrast and causality conveyed by *but* and *so*, respectively, are additional to the logical properties of the conjunction, while not contributing to the truth-conditions of the sentence.

[2] Anscombre and Ducrot (1977) underline two ways to describe *but*. One is to provide contrast, as in *Camille has two apples but Debra has one*, and the other is a denial of expectations. In (1b), the latter is a better descriptor. Others (e.g., Hall, 2004) list four descriptions of *but*, the two additional meanings being correction, as in *That's not my sister but my mother* (Iten, 2000, p. 181) and objection, as in *But you don't drink!*.

In this work, we aim to address this challenge by examining the processing demands related to the discourse connectives *but* and *so* from the very moment they are encountered, and by determining the extent to which these can be distinguished from *and*. One can imagine a range of outcomes from such an investigation. On one hand, one could anticipate that the extra-logical features do not add effects or processing costs. On the other hand, one could anticipate, as we do, that *but* and *so* prompt unique effects in an addressee, when compared to *and*; this would indicate that *but* and *so* provide specific additional information. If the data are consistent with the latter, we would, then, be in a better position to determine the nature of this added information (i.e., we would have processing profiles that are more amenable to being characterized as semantic or pragmatic).

In the remainder of this Introduction, we provide the relevant theoretical and experimental background from the rich literature on discourse connectives, in order to make our aims clear. Overall, we take the following four steps. First, we continue our brief review while considering the various approaches (pragmatic, semantic as well as taxonomic) that have guided the study of discourse connectives. Second, we review the existing experimental literature, while showing that most studies focus on (a) the *downstream effects* of comprehending an appropriately (or inappropriately) applied discourse connective, and on (b) comparisons among discourse connectives. Third, we introduce our hypothesis, which distinguishes between the meaning of *and* and the two other connectives, *but* and *so*. Our expectation is that the extra-logical features in each of the latter two, in the form of a fixed procedure, render them richer, and, thus, cognitively more effortful to process, than *and*. Finally, we introduce our two pre-registered experiments, which aim to isolate the cognitive processing related to *but* and *so* when compared to that of *and*. Critically, unlike much prior work, our test sentences are (a) minimal in length, (b) minimally contentful (i.e., they are not thematically rich), and (c) designed to keep contextual information constant over the course of the experiment.

1.1 Theoretical and taxonomic approaches to discourse connectives

Paul Grice (1975) provided what is arguably the best known account of discourse connectives. In his seminal work, he described expressions such as *but* and *so* in the context of *conventional implicatures*. Unlike *conversational implicatures*, whose extra-sentential propositions are inferable from a speaker's utterance in context, *conventional implicatures* are triggered directly by specific lexical items. This makes them non-calculable, because the hearer does not need to evaluate contextual information and world knowledge to "work out" the implicature. Note that the implicated information that we drew out with respect to the examples in (1b) and (1c) is directly due to the discourse connectives. That is, the words *but* and *so* license inferences that, in turn, provide the hearer with the means to derive the speaker's informative intention.

While Grice viewed the extra-logical contribution of discourse connectives as pragmatic in nature, his approach did not lead to a consensus. Here, we cite three prominent accounts that

essentially provide semantic alternatives to Grice's approach. Bach (1999), for one, vehemently argued against Grice's notion of conventional implicatures and proposed instead that the contribution of discourse connectives can be analyzed solely in terms of the entailments that they generate. Bach (1999) added that while the second entailment created by a discourse connective does not concern the at-issue content, it is still truth-conditional. For example, in (1c), *so* appears to establish that Camille's meeting is the cause that is prompting his running late. That *causal link* would become false if the sentence is followed by "He is at a lab meeting indeed but that is not why he is running late". Similarly, Potts (2005) has proposed that the double entailment of an expression such as *but* reflects a multidimensional meaning. According to this approach, the sense of contrast carried by *but* and the sense of causality carried by *so* are each semantic contributions to their respective words. A third influential approach comes from Segmented Discourse Representation Theory (SDRT) (Asher & Vieu, 2005), which analyzes discourse in terms of rhetorical relations, or "what a sentence does to another sentence" (Jasinskaja & Karagjosova, 2020, p. 3). Given that a discourse is assumed to be hierarchical, it is generally accepted that rhetorical relations can be broken down into two sorts – those that provide linguistic effects that can be described as (a) coordinating (staying on the same hierarchical level) or else as (b) subordinating (those that make links to a different level). This analysis leads to an exhaustive list of relations: Coordinating cases refer to rhetorical relations described as *Parallel; Contrast*, as in (1b); *Narration*; and *Result*, as in (1c). Subordinating cases describe rhetorical relations such as *Elaboration* and *Explanation*. A good example of an (explanatory) subordinating relation can be seen with the connective *because*, as in *Mary is talented because she won first prize*. While SDRT has not been concerned with the relative processing costs of discourse connectives, as we are here (as far as we know, no proponent of SDRT suggests, say, that a connective expressing a subordinating relation should be more time-consuming than one expressing a coordinating relation), it is conceivable that one can derive processing predictions based on its principles. We will address this further in the Section 4.

A more sympathetic reaction to Grice's conventional implicature approach comes from Relevance Theory. Blakemore's (1987) account begins by making a distinction between *conceptual words*, such as *boy* and *runs*, and words encoding an instruction regarding how to process a concept (*procedural words*), such as *but* and *so*. Blakemore considers discourse connectives as good examples of procedural words in that a discourse connective constrains the kind of cognitive effects that an utterance can draw (for discussion, see Blakemore, 2000; Wilson, 2016). The examples in (1b) and (1c) fit well into Blakemore's scheme. Aside from conjoining, the discourse connectives *but* and *so* encode procedures that help addressees capture the speaker's intended meaning.

Likewise, Relevance Theorists view *and* in procedural terms, as well. By comparing cases in which two elements are conjoined by *and* to those that have juxtaposed sentences without *and*, Blakemore and Carston (1999, 2005) argue that a sentence that introduces *and* insures a communicative effect. In other words, these authors claim that *and* prompts a minimal, weak procedure. Given that a *but*

or a *so* implicitly expresses a form of conjunction too, we consider this proposal compatible with our initial one as well as with our goal, which is to investigate whether the additional information implicit in *but* and *so* prompt extra, isolable cognitive processing, compared to *and*.

Note that the elementary procedural contribution of *and* should be distinguished from the myriad ways that context often provides additional meaning to sentences that use this term. For example, sentences conjoined with *and* are often associated with temporal readings based on the order of the mentioned segments. Consider how *Mary had a baby and got married* can be taken to imply that the two events occurred in that order, so that *and* ought to be interpreted as *and then*. I. A. Noveck and Chevaux (2002) argued that such interpretations are due to conversational implicatures and, in support of that claim, they presented evidence showing that the developmental patterns for such temporal inferences resemble findings for scalar terms (I. Noveck, 2001). That is, seven-year-old children tend to accept as true (and 10-year-olds and adults tend to reject as false) a sentence conjoined by *and* that presents two (previously-described and ordered) events as inverted. This finding indicates that these so-called temporal inferences are optional and ultimately defeasible. This claim holds for a host of contextually determined meanings, including adversative and causal readings, that could appear to be part of *and*'s meaning.[3]

Discourse connectives have played a role in another linguistic tradition, one in which scholars aim to understand what makes a text (or units of discourse) cohesive, above and beyond grammatical structure. Halliday and Hasan (1976, p. 5) argued that cohesion can be achieved through a set of devices – reference, repetition, substitution, ellipsis, and conjunction – so that speakers can create textual contiguity.[4] In this approach, conjunction is comprised of linkers which connect sentences to each other, and these come in a variety of forms that can be taxonomically described (Martin, 2002). That is, conjunctions can be introduced via addition, comparison, temporality, and causality. One can see how discourse connectives fit into this scheme. For example, in (2a–2c), which are borrowed with slight modification from Hussein (2009), cohesion is carried out through linking linguistic expressions such as *and, yet,* and *now*, which are referred to as additive, adversative and temporal.

(2) John got a very good grade on his math test.
   a. And he has been the first in his class for the last two years (additive).
   b. Yet he failed his syntax test this term (adversative).
   c. Now, he feels very frustrated and is thinking of leaving school (temporal).

[3] For example, one can detect a causal reading in *She shot him in the head and he died instantly* and a contrastive reading in *It's autumn in France and it's spring in Chile*, but these readings are arguably due to the content of the two conjuncts, and not to the word *and*.

[4] The authors describe forms other than discourse connectives that maintain contiguity, such as the use of pronouns, which we do not discuss here.

While many adhere to this general view – that discourse markers are critical for understanding text (understood as a kind of linguistic unit) because they maintain cohesion – scholars often disagree about the structure or the source of these expressions (see Fraser, 1990; Schiffrin, 1987; Zwicky, 1985). As to be discussed in Section 1.2, one of the earliest attempts to create a psychologically plausible taxonomy of coherence relations (T. J. M. Sanders, Spooren, & Noordman, 1992) emerged from this general approach that connectives maintain cohesion.

For Fraser (1996, 1999), speakers employ discourse connectives to highlight how they intend two sentence parts to relate. It then becomes the interlocutor's task to shape an interpretation of the utterance which is compatible with the discourse relation signaled by the speaker. Anscombre and Ducrot (1977, 1989), on the other hand, claim that by signaling specific discourse relations, discourse connectives constrain the argumentative direction that the discourse can possibly take. This, in turn, is believed to impact the hearer's processing of the post-connective part of the utterance. In a third approach, Schiffrin (1987, 2005) assigns to discourse connectives the role of creating a coherent flow in between the different ideas expressed in the discourse. In her analysis, discourse is structured on several levels (*planes*) and connectives such as *but* and *so* intervene in the *ideational* plane of discourse.

While the above-referenced scholars generally agree that discourse connectives play a crucial role in utterances, they generally do not agree about the frameworks in which the connectives are described and the purposes they serve in comprehension. Nonetheless, most of the theoretical and classificational approaches to the study of discourse connectives could serve as inspiration for an experimental investigation. Below, we turn to prior experiments.

1.2 What experiments reveal

For the purposes of providing a summary, we break down exisiting experimental work on discourse connectives into two thematic categories. One is comprised of experiments from researchers of text and discourse who underline the importance of coherence for comprehension. The other is driven by experimenters who lack such theoretical commitments but who are generally interested in the way a discourse connective affects, say, the speed of integration of the second segment or other expressions downstream from a connective.

1.2.1 Discourse connectives and coherence

The taxonomical approach described earlier led to accounts that place an emphasis on coherence relations in discourse. That is, according to prominent researchers in the fields of text comprehension and discourse, establishing and detecting coherence among text fragments is crucial to readers and interlocutors; for seminal work, see T. J. M. Sanders et al. (1992). It is not surprising, then, that experimental efforts explored this angle. Whereas early work from T.

J. M. Sanders et al. (1992) proposed that coherence relations can be determined by combining four primitive categories,[5] later work proposed that processing outcomes directly depend on the nature of the discourse relation (T. Sanders, 2005; T. J. M. Sanders & Noordman, 2000; Spooren & Sanders, 2008). For instance, Spooren and Sanders (2008) suggest the following order in terms of processing effort (listed here from least to most complex) – additive (*and*), temporal (*and then*), causal (*so*), adversative (*but*). As such, these were expected to correlate with developmental patterns. In line with this, they showed – via language elicitation tasks in Dutch – that children between 7 and 12 years of age progressively master negative causal relations (e.g., the use of *although*) as they grow older, which they interpreted as evidence that negative causal relations are cognitively more complex than positive causal relations. Notably, at least part of their proposal was inspired by Bloom et al.'s (1980) findings showing that four children between 25 and 40 months of age acquired the discourse connectives linked with a positive additive relation (*and*) earlier than they did other types of discourse connectives.

The one study from coherence advocates that comes closest to localizing the processing related to discourse connectives is an eye-tracking study, Koornneef and Sanders (2013). Before we describe how discourse connectives come into play in this study, it is important to first set the stage. That is, the authors were chiefly concerned with testing a hypothesis about referential resolution bias, which refers to the way certain verbs implicitly convey causality and expectations about a given pronoun downstream. For example, when provided the phrase *David apologized to Linda because*, a participant (an addressee) would be justified in expecting the next reference to be about the male Noun Phrase, as in *David apologized to Linda because he …* (cf. the verb *praised*, which puts the focus on the second Noun Phrase, as in *David praised Linda because she…*). As one would expect, the main critical measure in Koornneef and Sanders (2013) was the gaze duration on the region *because he* for cases following verbs such as *apologized*. As they predicted, the authors found that the participants' gaze duration with respect to the pronoun, indeed, depended on the lexico-semantic features of the verb. This is where the discourse connectives become relevant. Interestingly, while the study focused on *because*, it included control conditions that employed an adversative discourse connective *but* in one condition as well as the conjunction *and* in another. Remarkably, their results revealed that the sentence fragment *and he* in *David apologized to Linda and he…* produced noticeably shorter gaze durations compared to *because he* in *David apologized to Linda because he*. The same was true with respect to the difference between *and he* and *but he* (whose gaze durations were

[5] For the purposes of this article, we do not find it necessary to describe the four in detail, but, for the sake of completeness, they are: *basic operation* (whether the coherence relation is additive or causal), the *source* of coherence (whether it is semantic or pragmatic), the *polarity* of the relation (whether it is positive or negative) and the *order of segments* in a causal relation (basic or non-basic).

comparable to those in the *because* condition).[6] This is one indication showing that *and* itself is processed more quickly than the two discourse connectives. This is consistent with a claim that says there is a difference between the minimal effort needed to immediately process the truth-functional *and*, on the one hand, and what we are calling the extra-logical processing called for by connectives such as *but* and *so*, on the other.

1.2.2 Discourse connectives and their effects downstream

Experiments in the literature typically investigate how a well-chosen, or else a badly chosen, discourse connective affects processing downstream (and, often, far downstream) from the connective. In such experiments, stimuli with discourse connectives are compared to those without any connectives. Below, we present studies in this vein. For illustrative purposes, consider (a portion of) the results presented in an early paper (Ziti & Champagnol, 1992) that requested French participants to read two clauses. In order to investigate the potentially facilitative effects of causal connectives (*parce-que/puisque* which translates to *because/since*), clause-pairs were prepared to express cause-effect relations, such as those in (3) below:

(3) a. The snow had made the racetrack extremely difficult [cause]
b. all the competitors quit the race well before the finish line. [effect]

Importantly, the connectives were added appropriately and systematically at the beginning of the *cause* clause whether it appeared first or second. Ultimately, there were two kinds of sentences for each cause-effect pair. For example, (3) could be presented as "Since the snow had made the racetrack extremely difficult, all the competitors quit the race well before the finish line" or as "All the competitors quit the race well before the finish line because the snow had made the race track extremely difficult". The paradigm also included cases where there were no connectives at all. While it pays to be prudent with these data, since there was a secondary task that we do not address and) that could affect outcomes, it appears that (a) the presence of a causal connective in a clause slows down the reading time of that clause (when compared to the no-connective condition) and (b) the reading time of the effect clause that follows the connective is facilitated, when compared to cases without a connective. Overall, there is a cost-benefit analysis to apply to these data, which is that the connective adds a cost in reading time, but it prompts salutary

[6] To maintain a reasonable narrative, test sentences varied in order to produce "consistent pronoun" items. For example, the entire sentence for the *because he* region was *David apologised to Linda because he according to the witnesses was the one to blame*, whereas the entire sentence for *but he* was "David apologised to Linda but he according to the witnesses was *not* [italics added] the one to blame." This means that only first gaze durations capture immediate reactions to these connectives. Later regressions could be due to features other than the connective. Given that this work was focused on the referential resolution bias (and, thus, the role played by the verb), the authors did not pay too much attention to the gaze duration differences between *and* and the two other discourse connectives.

effects downstream. This is in keeping with a general tendency that one finds in the experimental pragmatic literature. For example, when participants are in a position to enrich a weak scalar term, such as *some*, to mean the more informative *some but not all*, this comes at a processing cost; however, it comes with a benefit downstream, in that, once partitive, the *not all* reading can facilitate what follows (see Breheny et al., 2006; for a review see, I. Noveck, 2018).

Recent work from Recio Fernández et al. (2023) further bolsters the above findings. These authors show that a poorly chosen connective (e.g., *Daniel and Rosa have few plants. Therefore, they use plenty of water*) hinders sentence processing more than a mismatch between the conceptual content of the two segments (e.g., *She was craving something sweet. She is now eating a bag of crisps*). The authors see this as indicating that the procedural meaning of a connective imposes a particularly rigid frame on the possible ways to combine the conceptual meaning of two sentence segments. Consequently, a violation of this procedural meaning triggers effortful strategies to repair the processing of the target sentence.

More processing detail can be accrued with studies that collect neurophysiological measures as participants read content-rich sentences whose meanings are affected by the presence of a discourse connective. For example, Xiang and Kuperberg (2015) reported that the discourse connective *even so*, which "reverses" expectations, had specific effects on the processing of a dependent measure downstream. Consider the following vignette (4), which considers a verb downstream (which we underline for convenience) as the dependent variable in a task whose structure generates four kinds of experimental items through the placement of key words in two places (see bracketed words):

(4) Elizabeth had a history exam on Monday. She took the test and [aced/failed] it. [Even so,] she went home and <u>celebrated</u> wildly

One can see that acing a test without the connective *Even so* is coherent with the verb *celebrated* downstream (this is called the Plain-Coherent condition, because it contains no discourse connective, making it coherent to celebrate after a success). Likewise, if one were to fail a test, the presence of *Even so* would, then, justify the word *celebrated* (this is the Even so-Coherent condition). However, it would be incoherent to expect *celebrated* after a failed exam when the two parts are Plain (the Plain-Incoherent condition). Likewise, it would be odd to read about someone acing a test before reading "even so, she …celebrated wildly" (the Even so-Incoherent condition).

The authors reported that target verbs in the Plain–Incoherent condition prompted steeper N400's than those in the Plain-Coherent condition. This much is unsurprising. Interestingly, they also reported (even stronger) opposite effects when comparing the N400's to the critical verbs in the Even so-Coherent condition to those in the Even so-Incoherent condition. Thus, the discourse connective *even so* not only makes participants expect (and process) an otherwise unexpected verb, it can also make a contextually anticipated verb appear surprising.

More recently, Köhne-Fuetterer et al. (2021) published a series of studies (one eye-tracking investigation in German and two ERP experiments, one in German and one in English) investigating how participants directly react to the connectives *therefore* (or *daher* in German) and *however* (*denoch*) as they were reading sentences. For instance, in Köhne-Fuetterer et al.'s (2021) Experiment 3, the authors set up four possible experimental items through the available combinations in (5) below. Note how, when *However* is instantiated, the concessive triggers the revision of an expectation that had been created by the context, whereas *Therefore* indicates (or ought to indicate) that it will be confirmed. In the example below, the source of the dependent measure is the place that Mr. Brown heads to:

(5) Mr. Brown was planning to look for new glasses and shoes today. The glasses really are more urgent. [Therefore/However], he now heads towards [an optician/a shoe shop] that a friend recommended.

The authors reported an "N400-like" effect when comparing incongruent versus congruent expressions downstream and regardless of connective (i.e., effects surfaced when *Therefore* is followed by *shoe shop*, which is compared to *optician*, and when *However* is followed by *optician*, which is compared to *shoe shop*). This much is not surprising. What is more interesting is that the authors also reported that the connective *However* itself elicited a larger fronto-central positivity than the causal connector *Therefore*, in the range of 300 to 600 ms after onset (also see Xiang & Kuperberg, 2015).

Taken together, reaction time measures and the finer measures from EEG studies show that discourse connectives are readily integrated so that relevant elements can be treated on-line downstream. One could further infer from the findings in Köhne-Fuetterer et al. (2021) that the extra-logical contribution to discourse is part of a connective's meaning, at least when comparing two different discourse connectives. However, the global claims one can make are rather limited: while one can conclude that concessives, such as *however*, are relatively more difficult to integrate than, say, *therefore*, theoretical distinctions among discourse connectives are harder to draw.

1.2.3 Summary of the experimental findings

The experimental literature on discourse connectives has been concerned with answering the following questions : Are some types of discourse relations more complex to process than others? Are discourse connectives immediately or progressively integrated into sentence comprehension? The answer to the first question appears to be affirmative. As reported by Koornneef and Sanders (2013), it appears that the Dutch equivalents to *because* and *but* are more complex than *and*. Likewise, experiments in Köhne-Fuetterer et al.'s (2021) study can be taken to indicate that *however* is more cognitively demanding than *therefore* (which is synonymous with *and so*). As far as the second question is concerned, it appears that the extra-logical contribution to sentence

comprehension is certainly evident downstream. Evidence from reaction time, EEG and eye-tracking studies indicate that discourse connectives create specific expectations (Köhne-Fuetterer et al., 2021; T. J. M. Sanders & Noordman, 2000; Xiang & Kuperberg, 2015). Existing data also indicate that discourse connectives arguably prompt immediate reactions; as Köhne-Fuetterer et al. (2021) showed, concessives prompt steeper late positivities than causals. We add that, while based on a limited sample, the positivity effects in Köhne-Fuetterer et al. (2021) point to revision processes that are often associated with pragmatic reanalysis (e.g., see Drenhaus et al., 2006, 2011).

While the experimental literature on discourse connectives continues to grow, there remains one outstanding question stemming from the existing literature and (at least) one other general question. The outstanding question can be summarized as follows: Are the reported differences between *and* and the discourse connectives carrying extra-logical features (i.e., procedures) reflective of a general phenomenon, as several of the above studies seem to suggest? The more general question concerns the contexts of the stimuli in which most studies are carried out. Would one find indications that procedure-rich discourse connectives, such as *but* and *so*, are costly to process when the discourse connective is essentially the sole variable manipulated (in test sentences that avoid relying on real world knowledge)?

1.3 The present study: Motivations and predictions

The current work aims to address both of the above questions. Here, we compare the processing profiles of discourse connectives expressed in English – *and, but* and *so* – as they are used in simple conjunctive sentences. We are interested in (a) the way a participant processes the connectives when they are initially encountered, as well as (b) how they affect the processing of the segment that follows each sort.

Our predictions are drawn from Relevance Theory, according to which utterances are signals that are presumed to help a listener determine the speaker's intended meaning (Sperber & Wilson, 1986/1995). As far as an addressee is concerned, incoming input is characterized through two features. One is the cognitive effects that an input produces (e.g., in terms of new conclusions produced) and the other is how much effort it takes to process that input. The word *relevance* refers to the processing of incoming input via the interplay of effects produced in the addressee and the effort it takes to carry them out (e.g., highly relevant input provides maximal cognitive effects, while being minimally costly to process; relatively low relevant input provides minimal effects, while being costly to process). We assume that an extra-logical procedure, such as the ones in *but* and in *so*, prompts added effects in an addressee when compared to *and*; all things being equal, it is reasonable to assume that this relative gain in anticipated information is offset by a processing cost. Therefore, we predict that we will find – for a series of out-of-the-blue statements presented in ceteris paribus conditions – a processing cost difference between the discourse connectives *but* and *so*, on the one hand, and *and*, on the other, from the moment that

they are encountered. We add that, while the procedures attached to *but* and to *so* may differ from one another (along with their respective processing costs), we do not make specific predictions in that regard. For the purposes of the current work, we are only concerned with whether we can demonstrate that *but* and *so* each have extra-logical features that generate further processing and are not assumed to be carried by the more neutral *and*.

Given that the effects arise from what we are assuming to be obligatory procedures, it follows that one ought to find differences in the post-connective region among the three connectives as well. For example, one would expect that a segment that follows *but* and that provides a contrast with the first segment should be easier to process (as measured by error rates and processing time) than one that does not. In the event that a test sentence does not provide a contrast when *but* is employed, this is likely to slow down processing. Likewise, while causality is part of the meaning of *so*, its procedural meaning might add gratuitous processing if nothing causal can account for the second segment. Prior to the experiments, it was unclear whether a connective that ended up being part of a weakly relevant expression (*but* used without a cashed-out contrast or *so* without an evident cause-effect relationship) would affect validity judgements. This, in itself, was of interest to us. To summarize, we endeavour to determine whether a discourse connective's hypothesized procedure can be said to manifest itself through two dependent measures – validity judgements and latencies.

Our experimental approach differs from those that currently populate the literature. We simply seek to determine *whether* one finds indications of processing costs and benefits as participants process the three terms – *and*, *but* and *so* – from the moment each is presented. To test our hypotheses, we designed two self-paced reading experiments in which participants read simple statements broken down into two parts, which we refer to as *segments*, like the one in (6). Note that the vertical pipe indicates where the first segment ends and where participants advance the text. Once the sentence is completed, the participants' task is to decide if the sentence accurately describes the letters in a target word (here and elsewhere we will always use the word *BET* for expository purposes). There are two experiments, because in Experiment 1 we present the word (which is the source of a truth value judgement) before the two segments are read, and in Experiment 2, we present the word after the two segments are read. The latter Experiment has the advantage of separating the reading time of the second segment from the truth-evaluation.

(6) *Target word : BET*
There is a B and | there is a T.

As we indicated, we will compare the processing costs of *but* and *so* trials to those of the trials containing *and*, which serves as a baseline, in order to determine whether, and at what moment, one finds costs (and, in some cases, benefits) from an extra-logical procedure. One will notice that we designed our experimental material so that participants rely on as little context and world

knowledge as possible, while still having content that remains meaningful enough (by virtue of the target word, the sentence has a denotation and can, thus, be true or false). Ultimately, we are in a position to isolate the processing role of each discourse connective by ensuring that each one would be the only possible source of variability in a test-sentence (in principle, no other discourse relations could emerge from the semantics of the sentence alone).

This approach is unlike most experimental work in the field and is, to the best of our knowledge, novel. As reviewed above, experimental work concerning discourse connectives systematically includes rich thematic content (e.g., about exams, racetracks, or shopping adventures), which inevitably encourages the processing not only of discourse connectives but of a very wide variety of other contextual inferences that could vary as a function of background information. These thematic texts arguably mask the effects generated by the discourse connectives themselves.

We will now present our experiments. To anticipate, the results of both experiments support our main hypothesis that *but* and *so* trigger procedures that are cognitively costly, from the moment that they are read, relative to the processing of the truth-functional *and*. They will also show that these terms prompt specific processes that become apparent downstream.

2 Experiment 1

2.1 Method

2.1.1 Participants

We recruited 80 participants (14 males) between 18 and 35 years old (mean = 25.29) through the online platform prolific.co. As the experiment was conducted in English, we recruited monolingual native speakers of any regional varieties of English who had declared no learning difficulties, no cognitive or hearing impairments, while having normal or corrected-to-normal vision. Before starting the experiment, participants gave their informed consent for their data to be recorded.

2.1.2 Materials

Preparations were made so that a participant would view a three-letter word followed by a test sentence in three steps: (i) a three-letter word, (ii) the first sentence-segment ending with the connective and (iii) the second segment. There were 108 trials; seventy-two were target items and 36 were fillers.

Three factors were manipulated in order to design the 72 target items: *Connective* type (*and, but, so*), *Sentence polarity* in the second segment (affirmative, negative) and *Truth value* of the sentence (True, False). An example sentence for the different factorial combination is given in **Figure 1**. Each of the 12 possible combinations appeared 6 times in a randomized order. The experiment was built using the experiment farm provided by PCIbex (Zehr & Schwarz, 2018).

	Affirmative		Negative	
True	There is a B but / There is a B and / There is a B so	there is a T.	There is a B but / There is a B and / There is a B so	there is no K.
False	There is a B but / There is a B and / There is a B so	there is a K.	There is a B but / There is a B and / There is a B so	there is no T.

Figure 1: Example of all 12 possible test sentences for the target word BET.

The variable sentence polarity (whether the *second* segment of the trial sentence contains a negative or not) was introduced to manipulate the felicity of the contexts in which *but* appeared. Indeed, introducing a negation after the connective enabled us to satisfy participants' expectation following *but*, while minimizing the amount of inferential information contained in a trial. Manipulating the truth value of the trials allowed us to assess sentence comprehension and to monitor participants' engagement in the task through the dependent variable *answer accuracy*.

All target items were created to meet specific criteria in order to minimize the emergence of confounding effects. We, thus, instituted three criteria for these critical items:

- That the first segment be true, so that the second segment would determine the truth value of the trial.
- That the first segment of the test item start with either the first or the second letter of the word, so that no target trials required participants to backward-recall a letter (this maintains a uniformity of cognitive demand outside of those required by each specific condition).
- That the letter used to render a sentence false (in the second segment) would require the same preceding article as the two remaining candidate letters of the target word. For instance, in (7), if the target word is TEA, since both E and A require the *an* article, using (7b) would give away the truth value of the sentence just before the letter is read. In such instances, we ensured consistency with the remaining letters and went with option (7a).

(7) Word : TEA
  a. There is a T but there is an F.
  b. There is a T but there is a V.

The 36 filler sentences served to present participants with sentence types excluded from the test items, such as a negative first segment followed by a positive second segment (8) or two negative segments (9).

(8) There is no J and there is an E.

(9) There is no K and there is no V.

Fillers also included test sentences starting with a statement about the last letter of the word, or with a false statement. This ensured that participants would remain vigilant and would not get accustomed to the implicit pattern of target items.

The three-letter words used across the experiment (target and filler items) were carefully selected with the following criteria:

- Only content words were used.
- The letter W was excluded from the test phase, because it is multi-syllabic.
- Words featuring a repetition of one letter (e.g., BEE) were excluded, to ensure that recalling the remaining two letters would require comparable cognitive demands across all items.
- Words were controlled for their consonant and vowel composition. The balance of words was such that participants saw nearly as many vowels (161) as consonants (163).
- A given word never appeared in more than one trial.

Moreover, we created two different lists of test items with the same set of words but which were used across different conditions. This ensured that any trial-specific confound could be spotted by comparing the results of the two different lists. Participants were randomly assigned to one of the lists when beginning the task. The two lists of target and filler items used in the experiment can be found on the OSF registration page for the project.

2.1.3 Procedure

Participants completed the task anonymously and on-line, using their own computer. Before starting the experiment, participants were requested to provide their age, gender and handedness. They were then informed that they would play a word game in which they would first see a three-letter word, followed by a sentence about the letters in the word, presented in two segments, and that their task would be to assess the accuracy of the sentence relative to the word. Participants were instructed to press the space bar after the first sentence-segment to reveal the second segment and to press either F (False) or J (True) at the end of the sentence to evaluate its accuracy. To ensure the quality of the reaction time data, we asked participants to complete the task carefully but in a timely manner, and we also requested that they position their hands in a way that optimized the time efficiency of their responses, as illustrated in **Figure 2**. The test items described in the Section 2.1.2 were presented in four steps. Participants first saw a fixation cross marking the place where the first letter of the text was going to appear. This prevented a potential increase in reaction times due to participants' gaze movement across the screen.

The cross was directly followed by the target word. Then, the first segment of the test sentence was presented and, finally, the second segment. **Figure 3** presents a typical trial, including the participant's response options. The icon to the right of the penultimate screen and the two icons to the right of the last screen represent the three dependent variables recorded, which are the following: (i) the reading time of the first segment (RT1), which includes the presentation of the connective, (ii) the reading time of the second segment (RT2), which includes the time to make a truth-value judgement, and (iii) the participant's truth value assignment, which ultimately provides us with rates of accuracy (AC).

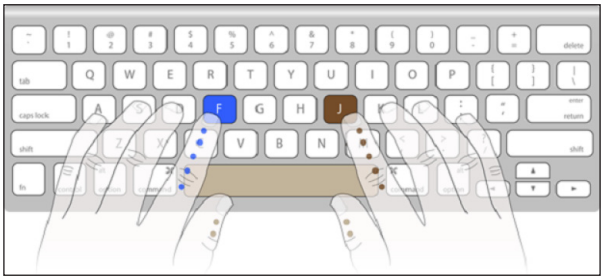

Figure 2: Finger positioning instructions given to participants.

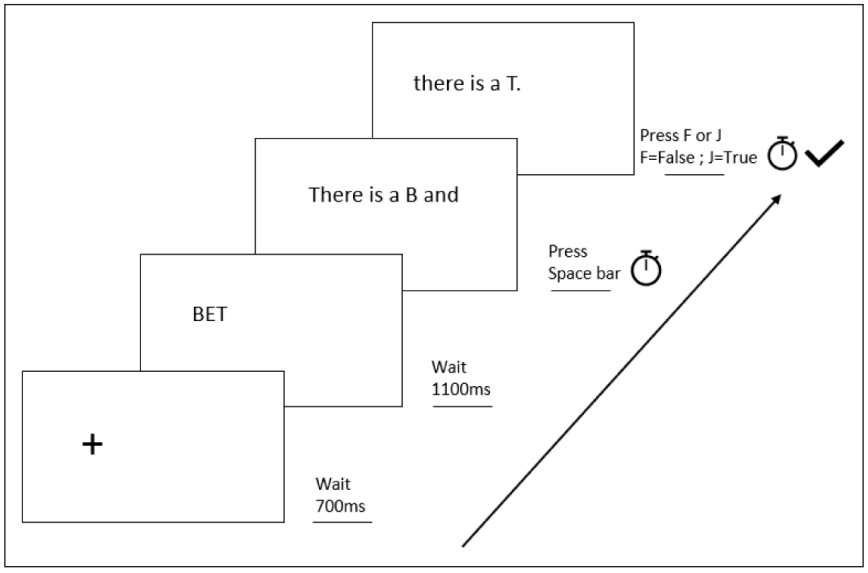

Figure 3: Experiment 1: The unfolding of a single trial.

Each trial was set to begin directly after the participant's truth-evaluation of the previous trial. No written instructions were given during the test phase, to ensure that the reaction times were not affected by superfluous or parasitic text. Participants did, however, practice the task during an extensive warm-up preceding the test phase. The first part of the warm up (five

trials) contained explicit instructions with each screen; then, in a second part, instructions were removed. Participants were only invited to move on to the test phase after having completed at least 25 mock trials and after they had answered the last three consecutive trials correctly. If they failed the last three sentences, the warm-up continued until they gave three correct answers in a row. Feedback was given for all incorrect answers throughout the warm-up session. We point out that the warm-up material did not include any critical items that would be used in the Experiment proper. This is why warm ups included other connectives, such as *then* and *because* trials. The average completion time for the whole task (including instructions and warm up) was 17 minutes (min = 11mn, max = 34mn).

2.1.4 Predictions

The predictions for this study were pre-registered on the OSF registration page for the project.[7] Reaction time 1 (RT1) refers to the reading time of the first segment, which ends with the connective. It is, therefore, a good comparative measure for the time required to process *and*, *but* and *so*. Following our hypothesis that processing the extra-logical import of *but* and *so* is relatively costly, we predict that trials containing *but* and *so* will result in slower RT1s than *and* trials.

The answer reaction time (RT2) is recorded after participants receive the second segment and have provided their truth-value judgement. It is an indicator of the time required to read and process the post-connective sentence segment and to evaluate its accuracy. Before presenting the predictions that are specific to this study, we predict two general trends in our RT2 that are not specific to the presence of connectives. We, indeed, expect main effects of sentence type and truth value, with negative sentences and false trials resulting in slower RT2s relative to affirmative and true trials, respectively. These two effects have been widely reported in the literature and, although they are not the focus of our study, they remain as study-specific predictions as a sanity check. That said, we expect that processing difficulties related to negative trials (again, this refers to second segments with negation) relative to affirmative trials, should not be found in *but*-trials. A facilitating effect should be caused by *but* incrementally creating an expectation of contrast and should translate into a shorter RT2 for negative *but*-trials, relative to the trials that use the two other connectives. Finally, the RT2 linked with *so*-trials should not differ qualitatively from the RT2 linked with *and* trials on the sentence type variable. Participants may, however, consider that the inference of causality introduced by *so* is never fully realized in our design, which would cause RT2 for *so* trials to be slower, overall, than for the *and* baseline.

[7] As can be determined from the OSF page, the study was time-stamped in 2021. At the moment of paper submission, we realized that we did not request a DOI when we made our predictions, which is more in line with current practices.

2.2 Results

2.2.1 Data selection

We begin by assessing participants' accuracy rates with respect to their truth value judgments. The overall rate of correct answers on all test items (target and fillers) was high (95 percent), indicating that the task was well understood and carefully executed. Two participants were, however, excluded from the data analysis, because their answer accuracy score was three standard deviations below the group mean performance. For each reaction time (RT1 and RT2), only successful target trials were kept, and data points outside of 2.5 standard deviations of the mean of the log reactions times were removed from the analysis, as described in the OSF pre-registration for this study. This cut-off was motivated by the fact that participants completed the task at home, which, we suspect, would lead to a considerable proportion of missed trials (several-minute-long breaks or accidental key presses before the input could have been read). Following this removal of outliers, 97.68 percent and 98.26 percent of RT1 and RT2 successful trials were kept, respectively.

2.2.2 Analysis

The three dependent variables recorded during the experiment were (i) the binary variable of accuracy (answer accuracy), (ii) the reaction time at the presentation of the connective, which was part of the first segment (RT1), and (iii) the combined reading time of the second segment and the answer reaction time at the end of the trial (RT2). All three variables were analysed with linear mixed models in a Bayesian framework, using the *brms* package (Bürkner, 2018) in R (R Core Team, 2021). The answer accuracy data was coded 1 for correct answers and 0 for wrong answers and analysed with a model using a logit-link function, while assuming a Bernoulli distribution.[8] For RT1 and RT2, the models assumed a shifted-log-normal distribution, which suits the specificity of reaction time data (Morís Fernández & Vadillo, 2020). All three models reported here were given a weakly informative prior (meaning that the model did not assume any reaction time differences between the variables entered) and were trained on 4 Markov chains, each repeating 4000 iterations (including 2000 warm-up ones). In a Bayesian framework, posterior distributions for all parameters estimated by the models are obtained. Considering that the value of a given parameter quantifies the differences induced by the predictor linked to it, a distribution indicates virtually no difference if it is centered around 0, and an increase or decrease in the dependent variable if it spreads only over, respectively, positive or negative values. We considered as effects with strong evidence only parameter coefficients whose posterior distribution had at least 95 per cent probability of being greater

[8] Trials were considered true when they were logically true. Participants' answers were, thus, coded as incorrect if they rejected a logically true trial that was arguably infelicitous, such as *There is a B but there is a T* for the word *BET*. As will become clear, false responses to these sorts of items were rare.

or smaller than zero. We begin our analyses with Answer Accuracy and then turn to the Reaction Time data.

2.2.3 Answer Accuracy

The model fitted to the Answer Accuracy data considered five fixed effects: the simple effects of Connective type (*and* vs. *but* vs. *so*), Sentence type (an Affirmative vs Negative second segment) and Truth value (True versus False), as well as their factorial interactions. Two random effects were also added to the model: a random intercept for each unique test item and random slopes and intercepts for participants as recommended by Barr et al. (2013). **Figure 4** represents the parameter distributions computed by the model whose distribution uses *and*, Affirmative second segment and True as its baseline.

The model revealed only one effect of statistical significance, namely, that trials with a

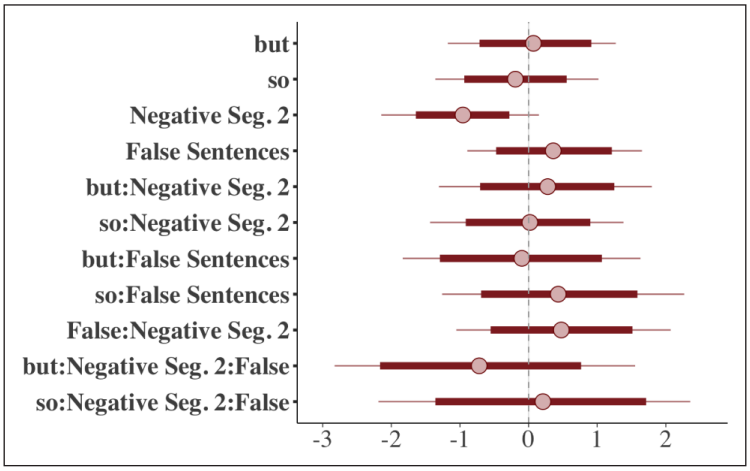

Figure 4: Experiment 1: Answer Accuracy parameter estimates for Connective type, Sentence type, Truth value and their factorial interactions (baseline = *and*, Affirmative, True sentences). The thin red line represents the whole extent of the parameter's distribution, the thick line covers the 95 percent credible interval for the true difference in Answer Accuracy induced by the factor and the dot is the average value of this difference.

Negative second segment were associated with lower rates of accuracy than those with an Affirmative second segment ($p(\beta < 0) = 1$) in true-*and* trials. That said, the coefficient estimates for all the interactions including Negative second segments (with *but*, *so*, with False sentences and the three-way interactions) appear to be symmetrically centered around 0, indicating that the effect of Negative second segment holds across all levels of the other predictors, as visually represented in **Figure 5** for the three Connectives.

Interestingly, the simple effects related to Connective type remain practically centered around 0 (for *but*: $p(\beta < 0) = 0.43$, mean of $\beta = 0.07$, 95 percent credible interval = [–0.71;0.91];

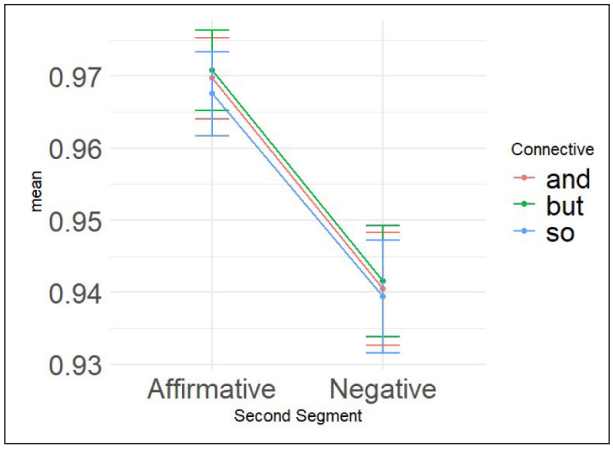

Figure 5: Experiment 1: Mean Answer Accuracy plotted by polarity of the second segment (Affirmative or Negative). Error-bars = standard error of the mean.

for *so*: $p(\beta < 0) = 0.70$, mean of $\beta = -0.19$, 95 percent credible interval = [–0.94;0.56]). This indicates that *but* and *so* are not linked with higher or lower rates of correct performance, relative to *and*.

2.2.4 RT1: Reading the first segment and the connective

The recorded reaction times at the presentation of the first segment, which includes the connective (RT1), were on average, faster for *and* (mean = 1121.74 ms) than for *but* (mean = 1161.01 ms) and *so* (mean = 1161.11 ms) (see **Figure 6**). For RT1, only the main effect of Connective was analyzed by the model given that the two other predictors (Truth value and Sentence type) become relevant only later in a trial. Furthermore, a random intercept was calculated for each test item, and participants' IDs were assigned a random intercept and slope.

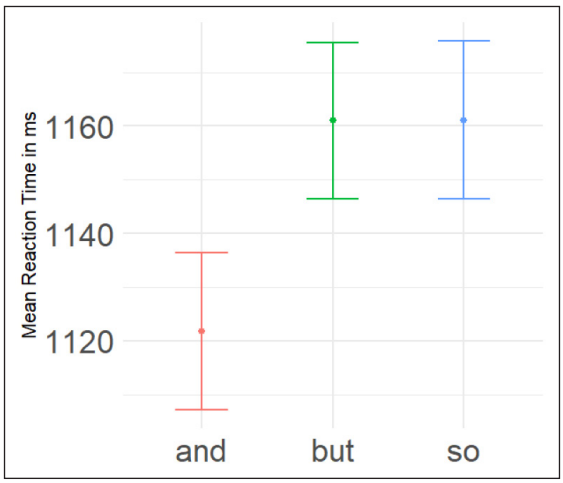

Figure 6: Experiment 1: Mean RT1 for each connective. The error-bar indicates the standard error of the mean.

A posterior distribution of the parameters estimated by a model of RT1 is represented in **Figure 7** for the two levels (*but* and *so*) of the fixed effect Connective, where *and* is the baseline. As predicted, the distribution of the possible values taken by the coefficients quantifying the main effect of Connective on RT1 indicates that *but* and *so* are linked with longer RT1 than *and* (for *but*, $p(\beta > 0) = 0.97$, mean of $\beta = 0.04$, 95 percent credible interval $= [-0.01; 0.09]$; for *so*, $p(\beta > 0) = 0.99$, mean of $\beta = 0.05$, 95 percent credible interval $= [0.01;0.09]$).

Figure 8, shows the density distribution of the RT1 value predicted by the model for each connective.[9] The values taken by *but* and *so* are further along the x-axis, relative to the *and* values, even though they overlap with the *and* distribution. This confirms that there is an attested difference in RT1 between *and*-trials and *but* and *so*-trials, as we predicted.

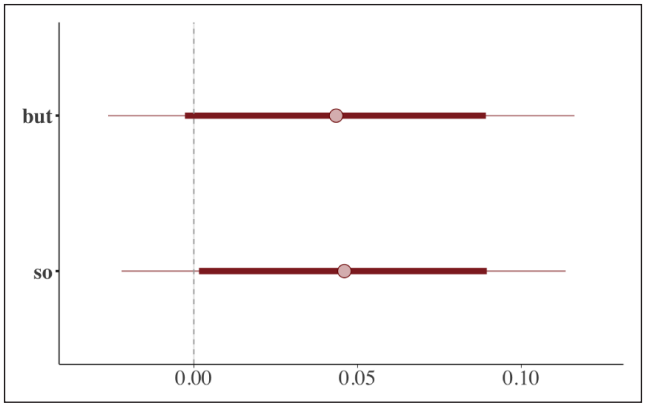

Figure 7: Experiment 1: RT1 parameter estimates for Connective type (baseline = *and*). The thin red line represents the whole extent of the parameter's distribution, the thick line covers the 95 percent credible interval for the true difference in RT1 induced by the factor and the dot is the average value of this difference.

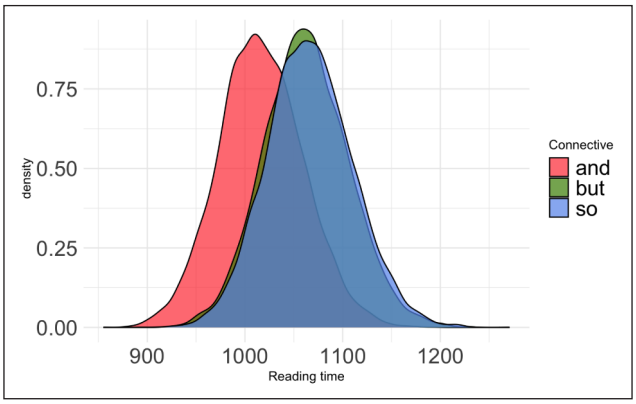

Figure 8: Experiment 1: Posterior distribution of the RT1 values predicted by the model.

[9] The values shown here assume a logarithmic distribution of the RT1 data.

2.2.5 RT2: Reading the second segment and validity decision speeds

RT2 was analysed in a Bayesian framework, following the same procedure as for RT1. The RT2 model was built with the simple effect of each predictor (Connective, Sentence type and Truth value) and their full factorial interactions. In addition to those fixed effects, random intercepts were added for each unique test item, and participants were assigned a random slope and intercept. The distribution of each parameter estimate is represented in **Figure 9**, where the baseline is the following: *and*, Affirmative second segment and True.

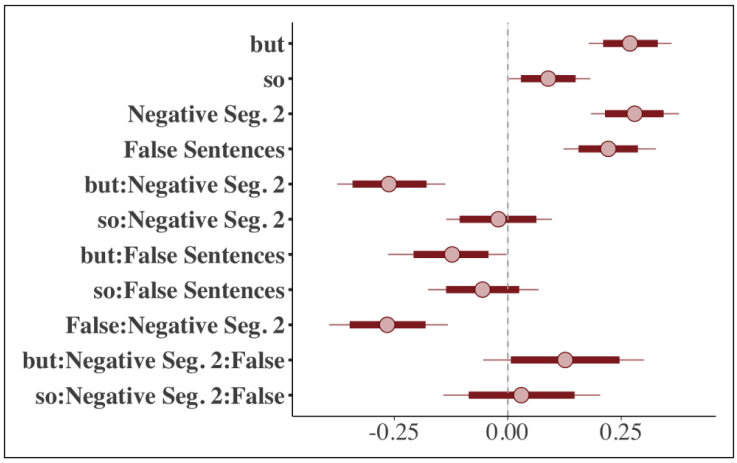

Figure 9: Experiment 1: RT2 parameter estimates for Connective type, Sentence type, Truth value and their full factorial interactions (baseline = *and*, Affirmative second segment, true sentences). The thin red line represents the whole extent of the parameter's distribution, the thick line covers the 95 percent credible interval for the true difference in RT2 induced by the factor and the dot is the average value of this difference.

Let us first consider the simple effects reported by the model. For False sentences as well as for Negative sentences, the range of possible β values is exclusively positive (both $p(\beta > 0) = 1$), indicating that those two predictors are linked with slower RT2 when the two remaining parameters are set at the baseline (*and*, Affirmative/true). This confirms our initial prediction that False and Negative second segment trials would be associated with processing difficulty. The use of *so* and *but* was also associated with an increase in RT2, relative to *and* trials, for true Affirmative second segment trials (for *but* and for *so*, $p(\beta > 0) = 1$). This estimated effect is only marginally informative, as it needs to be considered together with the effects of the two interactions included in the model. One will notice, from the parameter estimates, that *but* is linked with what looks like extraordinary slowdowns in Affirmative second segments. In Section 2.3, we will consider this outcome in the context of our claims.

At first sight, the parameter estimate for the *but*-by-Negative-sentence interaction stands out for being far to the left of the 0 line ($p(\beta < 0) = 1$, mean of $\beta = -0.25$, 95 percent credible interval = [–0.33;–0.17]). This negative distribution of estimates indicates that when a trial includes

but and a negative second segment, the resulting RT2 is faster than when one considers each of the simple effects taken individually. As a comparison, consider the *so*-by-Negative-sentence interaction, which is almost perfectly centered around 0. In other words, the simple effect of a slowdown associated with *so* or *but* and with Negative sentences can accurately predict the RT2 observed in *so*-Negative sentences but not in *but*-Negative sentences. **Figure 10** offers an intuitive visualisation of this interaction, where it can be seen that, while Negative sentences are linked with an almost identical slowdown effect for *and* and *so* trials (indicated by the parallelism of the blue and red slopes), for *but* trials, there is no slowdown linked with Negative sentences, relative to Affirmative sentences. This interaction is in line with our prediction that a negation in the second segment would satisfy the contrast induced by *but*, thus rendering the sentence felicitous.

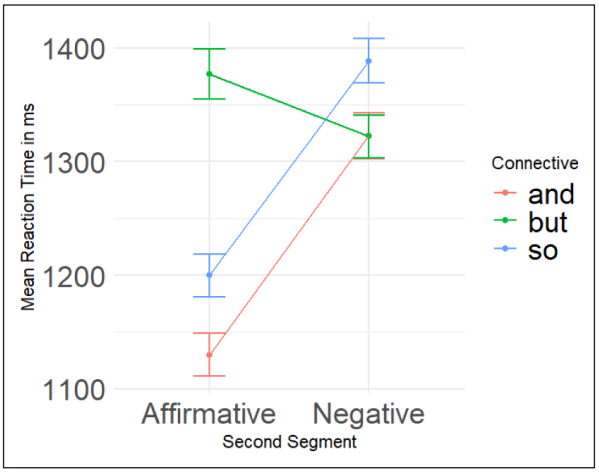

Figure 10: Experiment 1: Mean RT2 sorted by Sentence type and Connective; the error bar indicates the standard error of the mean.

The last interaction that is worth reporting here is the truth-value-by-sentence-type interaction. The exclusively negative distribution of the coefficient estimates indicates that participants were faster to respond to False trials when these contained a negation in the second segment. This may seem surprising at first, but this interaction is easily explained by the design of the experiment. When a trial is false, the letter that appears in the negative second segment is actually one of the three letters contained in the target word (e.g., for BET : *There is a B and there is no T*). This most likely facilitates the falsification process relative to cases where the letter mentioned in the segment is one of the 23 letters that is not part of the target word.

2.3 Discussion

This experiment was designed to test our hypothesis that the presentation of each of the discourse connectives *but* and *so* would be associated with a procedure that distinguishes them from *and*, which we assume lacks an extra-logical procedure. This would be expressed through

slowdowns at the moment of their mention and, eventually, through specific effects in processing downstream in a paradigm whose sentences varied with respect to the polarity of the second sentence and sentence validity. While accuracy rates remained very high (all conditions yielded rates of correct responses that were at least 93 per cent), one can see that our findings were largely consistent with our pre-registered predictions. We review the highlights of Experiment 1 below and in order of their appearance.

Answer Accuracy results revealed no effects of connective type, neither as a main effect nor as an interaction. In other words, it appears from our results that, whether or not one accepts the proposal that there is an extra-logical procedure linked to *but* and to *so*, validity judgments indicate that participants are thoroughly unaffected by the choice of connective. At this point, one might be left with the impression that *but* and *so* are simply equivalent to *and*. The only main effect is due to the lower accuracy that occurs in the wake of a negative second segment, which is unsurprising. A look at the reaction times, however, tentatively reveals that readers are, indeed, incorporating specific procedures after encountering *but* and *so*. We turn to these now.

The RT1 data provide unique insight into the processing of the extra-logical import of *but* and *so* by comparing these to the mean RT1 of trials with the connective *and*. As predicted, RT1's associated with *and* were consistently faster, relative to those recorded in *but* and *so* trials. This is direct evidence that participants incrementally integrate the extra-logical import of *but* and *so* from the moment that they are encountered. It is difficult to adduce these findings to other extraneous features, such as the word-length of the connective. These data are consistent with the indirect evidence that was culled from previous studies discussed in the Section 1 (e.g., Köhne-Fuetterer et al., 2021).

The last variable recorded in this experiment was RT2 (the reading time of the second segment plus validity decision time). Our analyses reveal two interesting findings. One is that the RT2 data indicate that *but* and *so* are linked to longer reaction times than *and*. This indicates that *but* and *so* processing, first detected as a source of slowdowns in RT1, continues downstream. However, this does not capture the subtlety of each of the extra-logical procedures (in *but* and *so*). A second, and more important, finding is the interesting interaction that points to participants' efforts to integrate the procedure linked to *but*. As captured in **Figure 9** and further illustrated in **Figure 10**, one finds a relatively long latency, i.e., an apparent slowdown, when the second segment following *but* is affirmative. This indicates that participants are noticing a clash between *but*'s procedure setting up a contrast and the subsequent lack of finding one. However, the *but*-by-Negative-sentence interaction indicates that participants react with alacrity when the second segment is expressed with a negation, thus providing the anticipated contrast. To make this concrete, consider the case of sentences in which a *but* was followed by an affirmative second segment (as in *There is a B but there is a T*, with respect to the word *BET*). According to our account, the procedure intrinsic to *but* ought to prompt participants to seek out a contrast. When

there is no obvious one, as in the above type of sentence, it prompts a relatively slow response (see **Figure 10**); when there is a contrast in a negative second segment (as in *There is a B but there is not a G*, with respect to the word *BET*), it leads to reaction times that are as fluid as those for *and* in the negative-second-segment condition. Note, too, that **Figure 9** shows that a *false* second segment also appears to satisfy the contrast-seeking procedure linked to *but*, as viewed by the speed-up that the false (and the false/negative second segment) cases provide.

The *so* trials produce a different kind of RT2 profile. These appear to prompt general slowdowns that are otherwise parallel to those found with *and*. Taken together with RT1, we consider the findings linked to *so* as indicative of participants applying a procedure that prompts a search for a causal explanation. Given that causal explanations for letter combinations are non-obvious, the use of *so* arguably prompts gratuitous processing. Note that the lack of an obvious cause-consequence relationship between mentioned letters does not slow participants down as much as *but* does in, affirmative second-segment condition. When it comes to a negative second segment, there is a pay-off for *but* in this paradigm. What is impressive is that participants cannot ignore the procedure imposed by *so*, even if it is gratuitous.

Overall, the results from this first experiment confirm our predictions and can be taken as evidence that discourse connectives, even in such a stark and repeatable paradigm, encode procedural meaning (RT2 results indicated that expectations are created following *but* and *so*), and that this extra-logical import comes at a processing cost, relative to a truth-functional conjunction (as can be seen as soon as RT1). Nevertheless, caution dictates that we consider two alternative accounts for the difference in RT1 observed between *but* and *so*, on the one hand, and *and*, on the other hand. One of these will lead to our next Experiment.

The first alternative explanation is that the RT1 results can be accounted for by an effect of frequency. According to the Corpus of Contemporary American English, for instance, *and* appeared 44247 times per one million words, against 8077 and 4232 times for, respectively, *but* and *so* (Davies, 2008). One could argue that participants are able to process *and* fastest simply, because it is more frequent than the two other connectives. However, a more careful analysis of these frequencies makes this account doubtful. When it comes to effects of frequency, the word recognition literature usually classifies words in three categories (low, medium and high frequency) in order to predict the reading and processing speeds of a given word. A word is usually considered *highly* frequent when it exceeds the threshold of 100 occurrences per million words (Brysbaert et al., 2018). All three connectives used in this study are, thus, highly frequent function words that very largely exceed this threshold. We thus consider it unlikely that frequency would play a role in reading time among such highly frequent words. Besides, the slowdowns are also associated with effects downstream. If *but* and *so* were just plain slower, one would not have other specific effects related to them in, for example, RT2. This leads to a second possible explanation.

It could be argued that what we consider to be gratuitous processing (what follows when a procedure does not get applied) amounts to infelicity, even upon the first encounter of the discourse connective (RT1). One could argue, for example, that *but* often (though not always) appears in sentences that do not provide a contrast; thus, it is possible that this conditioned participants to expect infelicity, even when it appears in the first segment. Such effects would, presumably, arise as part of a learning effect in which participants become increasingly sensitive to the infelicity of *but*. It follows from this argument that the difference in RT1 between *and* and *but* trials would increase as a function of trial order. We addressed this concern in a post-hoc analysis. We fitted a multilevel model, on the RT1 data, that included the variable of trial order and the interaction between Connective and trial order. We found no effects of either trial order or of the interaction between trial order and Connective. This argues against the notion that, even at the RT1 juncture, participants were becoming increasingly sensitive to the gratuitous processing of *but* in the experiment. This lack of trial effects applies to *so* as well, which we turn to now.

It could be similarly argued that the gratuitous processing that follows *so* amounts to infelicity because it is odd for an utterance to express causality when a participant eventually determines that there is no obvious reason for invoking it at all. To make this concrete, consider the two sentences described above, but conjoined with a *so*: *There is a B so there is a T* and *There is a B so there is no G*, with respect to the word *BET*. For both of these cases, we are assuming that *so* at RT1 prompts a procedure that anticipates a causal connection to consider; by the time participants read the second segment, they would be justified to wonder what cause is being cashed out.

There are two reasons why we are doubtful of accounts pointing to infelicity. One concerns the near-ceiling rates of Accuracy and their consistency across connectives. If one would want to claim that participants are so distracted by, say, two affirmative segments conjoined by *but*, one would have to predict higher error rates specific to those cases. Likewise, if the mere presence of the discourse connective *so* is disturbing, then arguably, one would find much longer RT's overall and they would be longer for the *so* case than in the *but* case, which prompts gratuitous processing nearly half the time. Or, one might argue – from the high rates of Accuracy – that the extra-logical meanings are simply ignored, but the reaction time data indicate otherwise. The other reason we are doubtful of such an infelicity account with respect to our data is that participants have little reason to doubt the relevance of the discourse connective, at least when it first appears (at the end of the first segment). Arguably, the second segment could reveal the reason for the causal connection or for a contrastive one. Note, too, that if one adopts the infelicity argument, it is tantamount to accepting our procedural meaning account, because the infelicities are presumably due to the obligatory procedures attributed to *but* and *so* that do not get exploited.

All that said, we want to further rule out claims that say that our findings are due to paradigm-related issues, such as purported infelicities. This leads to Experiment 2, which was designed to present the sentences while the truth values were still undetermined, by essentially presenting the word after the test statement. If we replicate our main findings while the truth values are not declared (via the prior presentation of the word), it would strengthen our claim that participants are working out the meaning of the sentence (based on the procedures attached to *but* and *so*), even while being unaware as to whether the statement can ultimately be considered felicitous or true.

3. Experiment 2

3.1 Motivation

The main goal of Experiment 2 is to attempt to replicate the results of Experiment 1 in a more refined design in which the truth values of the sentences cannot be determined *while* reading. In other words, the test word in Experiment 1 is placed after the conjunctive sentence. A participant who is producing measures for each segment cannot know whether *There is a T but there is a B* is true or false (or know, in principle, whether it is felicitous or infelicitous) if the word *BET* in the paradigm comes at the end of the trial. Participants are merely processing simple sentences in the anticipation of finding out what the word will be. The upshot is that the choice of conjunction can readily be accepted as the communicator's internal choice. We do not anticipate that this modification will affect the outcomes reported in Experiment 1. As will become clear, this effort also leads to two other fortuitous design changes that lead to cleaner dependent measures.

3.1.1 Participants

Participants were recruited following the same procedure as in Experiment 1. We tested a total of 80 native English speakers (22 males, mean age = 27.64).

3.1.2 Materials

The same two sets of test items from Experiment 1 were employed here.

3.1.3 Procedure

The procedure for Experiment 2 was similar to the one used in Experiment 1. The main difference was a context manipulation that led to a change in order presentation. Here, we present the new design and highlight each of the modifications.

Before the experiment begins, participants were told that they were about to help in the conception of a game that involved a human player and a robot. Participants received two sets of instructions, illustrated with **Figure 11**. One set concerned the game that was under development and for which their help was requested, and the other concerned the participants' actual task.

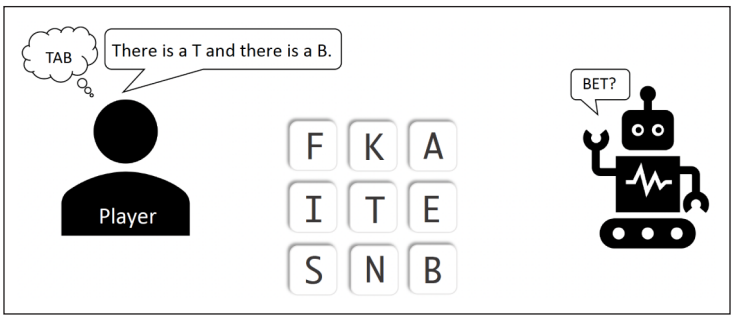

Figure 11: Experiment 2: Illustration accompanying the game instructions.

The rules of the game were the following : (i) The goal is for the human player to come up with a three-letter word and make the robot guess it accurately. (ii) The human player sees nine letter tiles per trial on their screen and uses these letters to form the three-letter word. (iii) The human player communicates hints to the robot. (iv) Hints can only be about two of the three letters in the word.

Participants then received the instructions regarding their task in the development of the game. They were told (i) that they would help the robot improve its guess-making by providing feedback; (ii) that they would see actual guesses issued by the robot during previous sessions of the game; (iii) During the task, participants would only see the hints and the robot's matching guess, but not the letter tiles. The justification for this was that this aspect of the game was already accurate.

With this cover story, our test sentences (which come in the form of the human player's hints) are uttered in a sufficiently rich context to justify the use of *but* and *so*. Given that the hints concern a small set of letters and a smaller set of potential words (at the time of the conjunctive clue), participants can easily assume that the use of *but* and *so* is relevant to some feature of the wider context that they do not have access to. For instance, in **Figure 12**, the participant could assume that the player's use of *There is a T but there is a B* underlines the fact that B is an infrequent partner for T in three-letter words (e.g., once T is mentioned, one might anticipate, say, an N, because T in combination with N gives rise to many genuine words (e.g., TEN, TAN, TIN). In practice, we do not expect participants to come up with such complex justifications for the use of *but* and *so* during the trials; however, we do expect participants to recognize that they – being ignorant of the speaker's larger context – can arrive at a felicitous reading of the trial sentences when *but* or *so* is employed.

This modification improves the paradigm in two ways. One is that we can now separate (a) a participant's RT2 from (b) their decision time (the truth value judgement now arises with the appearance of the word). The other is that we can unburden a participant's cognitive load (i.e., a participant no longer needs to store the three-letter word in memory while reading the segments). These modifications refine the paradigm's reaction time data (see **Figure 12** for the unfolding of a trial).

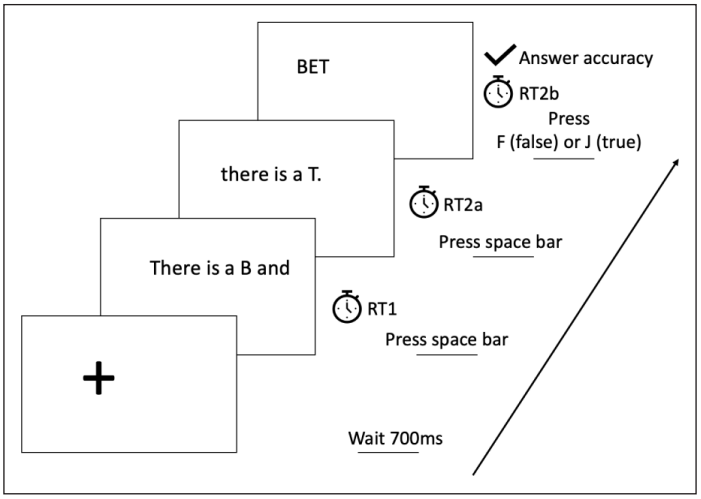

Figure 12: Experiment 2: The unfolding of a trial.

Except for the introduction of a cover story and for the change in the order of presentation, Experiment 2 unfolded similarly to Experiment 1. Participants receive instructions, then complete a warm-up in two phases (with and without instructions) and, finally, they are presented the 108 test items (of which 72 are target items). As in Experiment 1, we designed the warm ups so as to avoid critical test items; however, this time we achieved this goal by avoiding all discourse connectives, with the exception of *and*. With this modification, we ensured that the feedback received during the warm up would not influence participants on how to consider the extra-logical import of discourse connectives (truth-conditional or not) when evaluating the test sentences (see footnote 7). Participants' average completion time was 15.83 minutes. As we indicated above, Experiment 2 gave rise to four dependent variables. (i) A first reaction time (RT1) that presented the first segment that included the connective as the final word; (ii) A second reaction time (RT2a) that corresponded to the reading time of the second segment; (iii) A final reaction time (RT2b) that recorded how long participants needed to assess the validity of the sentence, and (iv) Accuracy with respect to the sentence's truth value.

3.1.4 Predictions

We predicted that we would replicate the results of Experiment 1.

3.2 Results

3.2.1 Data selection

We applied the same data selection procedures to Experiment 2 as were described for Experiment 1. The overall mean answer accuracy was high (94 percent), indicating that, in this Experiment too, participants completed the task attentively and with very low error rates. The

data from two participants were removed, because their rates of correct responses were three standard deviations below the group mean. As in Experiment 1, we excluded all trials resulting in incorrect answers and all reaction times outside of 2.5 standard deviations of the mean before performing any analysis on the reaction time data. Ninety-five percent of the trials in which participants had answered correctly were kept for data analysis.

3.2.2 Analysis

All four dependent variables (Answer Accuracy, RT1, RT2a and RT2b) of Experiment 2 were analysed with linear mixed models in a Bayesian framework, as in Experiment 1. For the variables of answer accuracy, RT1 and RT2b, the same models were used as those in Experiment 1 (where RT2b here corresponds to RT2 in Experiment 1). For RT2a, the details of the statistical model will be provided at the beginning of Section 3.2.5.

3.2.3 Answer Accuracy

The model fitted to Answer Accuracy followed the same approach as the one in Experiment 1 and represented in **Figure 13**. This analysis revealed a main effect of Connective type, with *but* trials leading to lower accuracy than *and* trials with a probability of 0.98 [for *but* : $p(\beta < 0) = 0.98$, mean of $\beta = -0.75$, 95 percent credible interval = [–1.40;–0.14]]. This main effect appears to hold across all conditions, since the posterior distribution of the coefficients linked with the interactions between *but* and Sentence type and *but* and Truth value spread out around the 0 line (indicating no significant interactions).

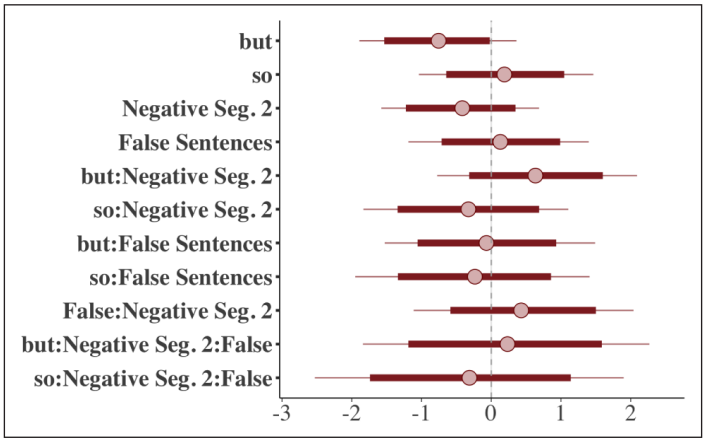

Figure 13: Experiment 2: Answer Accuracy parameter estimates for Connective type, Sentence type, Truth value and the Connective by Sentence type and Connective by Truth value interactions (baseline = and, Affirmative, True sentences). The thin red line represents the whole extent of the parameter's distribution, the thick line covers the 95 percent credible interval for the true difference in Answer Accuracy induced by the factor and the dot is the average value of this difference.

The main effect revealed by the statistical model is better understood with **Figure 14**, representing the mean answer accuracy by Sentences type (x-axis) and connective (colors). The green data points corresponding to *but*-trials are lower than the red ones in sentences with affirmative second segments only.[10]

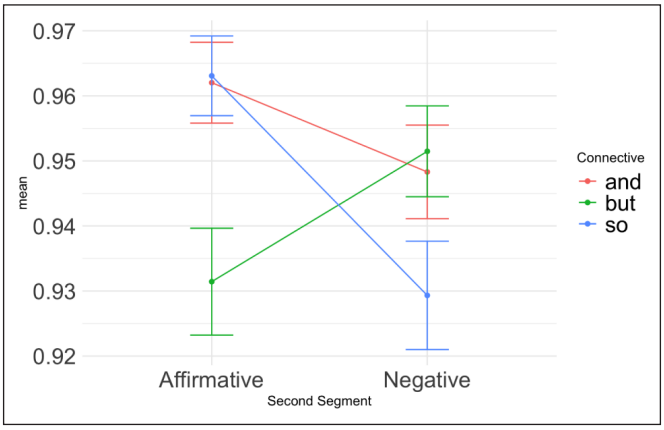

Figure 14: Experiment 2: Mean Answer Accuracy plotted by polarity of the second segment (Affirmative or Negative). Error-bars = standard error of the mean.

3.2.4 RT1: Reading the first segment and the connective

As in Experiment 1, we carried out an analysis in order to determine, for each connective, whether *but* and *so* triggered a slow-down in reading time of the first segment, relative to *and*. By visual inspection, one can see that there was a clearly marked difference between the mean RT1 for *and* trials (mean = 856.95 ms) on the one hand, and for *but* and *so* trials (respective means = 902.61 ms and 910.11 ms), on the other (see **Figure 15**).

This difference was confirmed by the statistical analysis. Indeed, the model shows that RT1 is longer in the presence of *but* or *so* than in *and* trials (for *but* and *so* $p(\beta > 0) = 1$). As **Figure 16** shows, the distribution of the coefficient linked with *but* was positive (mean of $\beta = 0.06$, 95 percent credible interval = [0.03;0.09]), indicating a systematic increase of RT1, relative to the baseline *and*. Similarly, the distribution of the parameter estimate for *so* was positive, as well (mean of $\beta = 0.07$, 95 percent credible interval = [0.04;0.11]). These results largely replicate the results obtained in Experiment 1.

[10] Note that the main effects in these types of models are only indicative of the effect of one variable (e.g., connective = *but*) when all other parameters are set to the baseline (here, affirmative-second-segment, true). This explains why there is a main effect of *but*, even though the accuracy rate linked with *but* is smaller only in the affirmative-second-segment condition. However the model did reveal that the interaction *but*-by-negative-second-segment had a high likelihood of having a positive effect on answer accuracy ($p(\beta > 0) = 0.9$).

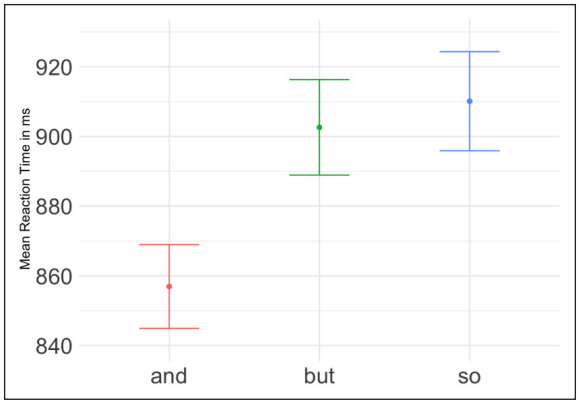

Figure 15: Experiment 2: Mean RT1 for each connective. The error-bar indicates the standard error of the mean.

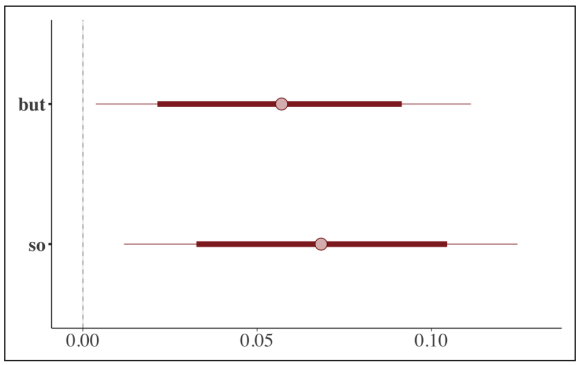

Figure 16: Experiment 2: RT1 parameter estimates for Connective type (baseline = *and*). The thin red line represents the whole extent of the parameter's distribution, the thick line covers the 95 percent credible interval for the true difference in RT1 induced by the factor and the dot is the average value of this difference.

3.2.5 RT2a: Reading the second segment

RT2a was analysed in a Bayesian framework, following the same procedure as for RT1. In Experiment 2, RT2a was recorded from the moment participants pressed the key to signal the end of reading the first segment, and it concluded when participants pressed a key to see the three-letter word. In this Experiment, the validity of the sentence could not be determined at this point, because participants had no way to assess the truth value of the sentence.

The RT2a model was, therefore, built with three fixed effects: the simple effects Connective and Sentence type and the interaction Connective by Sentence type. Additionally to those fixed effects, random intercepts were added for each unique test item, and participants were assigned a random slope and intercept. A parameter distribution was estimated by the statistical model for each fixed effect. These coefficient estimates are represented in **Figure 17**.

At first sight, the coefficient estimates represented in **Figure 17** look highly similar to those obtained in Experiment 1 (see **Figure 9**). As was found for RT2 in Experiment 1 (which included

validity decision times), the presence of a negation in the second segment reliably prompts longer reading times, relative to Affirmative second segments (mean of β = 0.08, 95 percent credible interval = [0.04;0.12]). Moreover, when the second segment of the sentence was affirmative following a *but* or a *so*, the model revealed a rise in RT2a, relative to instances where an affirmative second segment followed *and* (for *but*, mean of β = 0.13, 95 percent credible interval = [0.09;0.16]; for *so*, mean of β = 0.07, 95 percent credible interval = [0.03;0.10]). Globally, this indicates a processing slowdown linked with negative second segments and with the discourse connectives *but* and *so*. However, as was found in Experiment 1, the model also revealed an interaction (mean of β = –0.11, 95 percent credible interval = [–0.16;–0.06]), indicating that the reading time of the second segment speeds up when a negation appeared after the connective *but*. This replicates the results of Experiment 1, as illustrated by **Figure 18**.

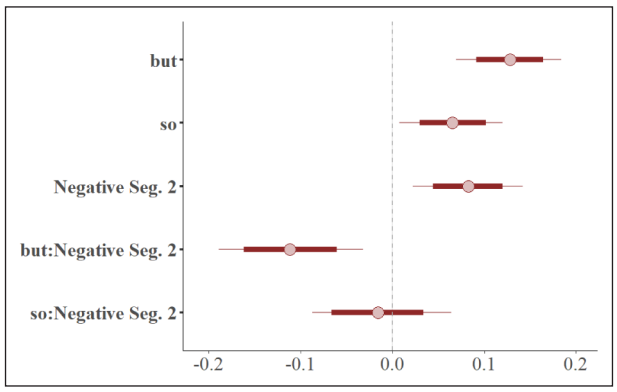

Figure 17: Experiment 2: RT2a parameter estimates for Connective type, Sentence type and the Connective by Sentence type interaction (baseline = *and*, Affirmative). The thin red line represents the whole extent of the parameter's distribution, the thick line covers the 95 percent credible interval for the true difference in RT2a induced by the factor and the dot is the average value of this difference.

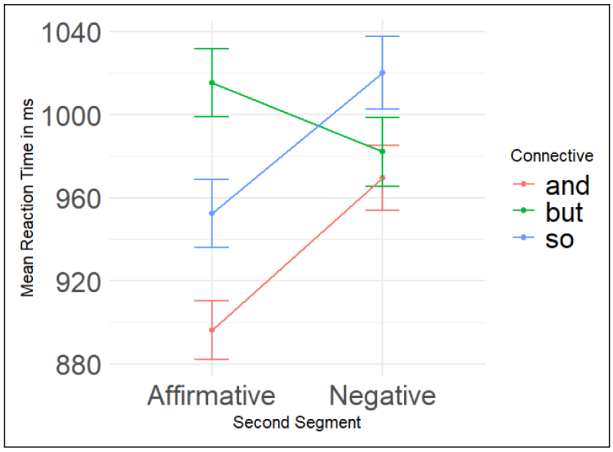

Figure 18: Experiment 2: Mean RT2a sorted by polarity of the second segment (Affirmative or Negative) and Connective, the error bar indicates the standard error of the mean.

3.2.6 RT2b: Validity decision speeds

In Experiment 2, RT2b captured the time that participants required to evaluate the trial just after the presentation of the three-letter-word. Since all the independent variables were relevant at this point, RT2b was analyzed using the same model as the RT2 in Experiment 1. The statistical model revealed a main effect of Sentence type, Truth value and Connective, as indicated by the distribution of the coefficients associated with Negative second-segments, False sentences, and those concerning *but* and *so* (see **Figure 19** for a representation of all the parameter estimates). These main effects were also found for RT2 in Experiment 1. As previously explained, these main effects are more meaningful when one takes into account the model's two interactions, which are represented in **Figures 20** and **21**.

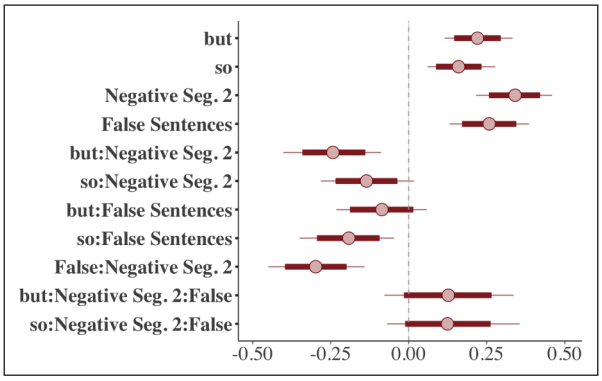

Figure 19: Experiment 2: RT2b parameter estimates for Connective type, Sentence type, Truth value and their factorial interactions (baseline = *and*, Affirmative, True sentences). The thin red line represents the whole extent of the parameter's distribution, the thick line covers the 95 percent credible interval for the true difference in RT2b induced by the factor and the dot is the average value of this difference.

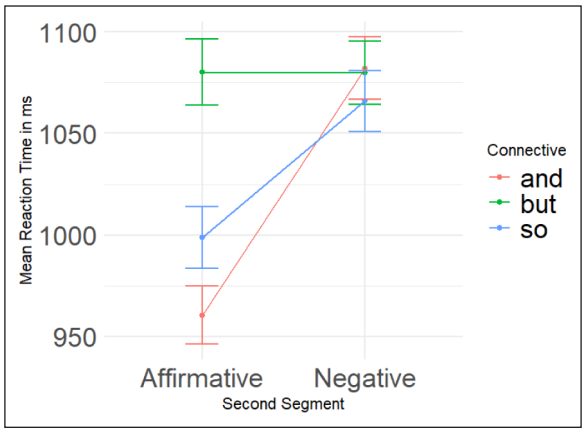

Figure 20: Experiment 2: Mean RT2b sorted by polarity of the second segment (Affirmative or Negative) and Connective, the error bar indicates the standard error of the mean.

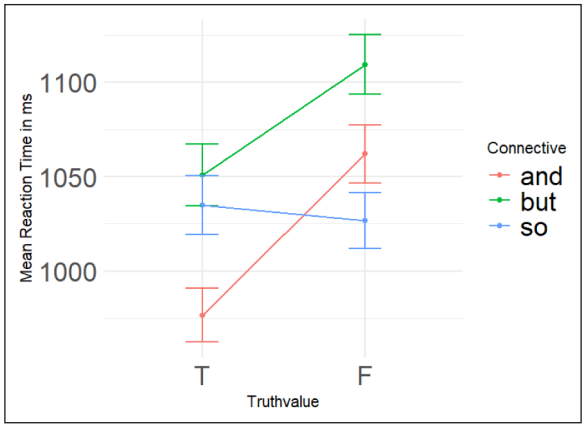

Figure 21: Experiment 2: Mean RT2b sorted by trials' Truth value and Connective, the error bar indicates the standard error of the mean.

In line with the interaction found with respect to RT2a, the model revealed an interaction indicating that judgements were unusually fast (in the form of RT2b's) when they took place in the wake of the combination *but* followed by Negative second segments (mean of β = –0.24, 95 percent credible interval = [–0.35;–0.14]). The model also revealed that correct False judgements in *so* trials were given faster than they should have been, given the simple effects of *so* and false trials (mean of β = –0.18, 95 percent credible interval = [–0.28;–0.07]). This particular interaction between Connective and Truth value was neither predicted nor found in Experiment 1. In the Discussion, we discuss the findings of Experiment 2 and propose an explanation for this seemingly facilitating effect of *so* with respect to False judgements.

3.3 Discussion

Experiment 2 was designed to further test our hypotheses concerning the presence of extra-logical procedures linked to *but* and to *so* that distinguish them from *and*, which we assume lacks an extra-logical procedure. The design of the current experiment allowed participants to read the test sentences without the burden of storing the three-letter words (the basis for making validity judgements) while making their evaluations. In other words, the main innovation of this design is putting the three-letter words at the end. One of the advantages of our design change is that it also allowed us to collect reading times of the second-segment (what is called RT2a here) independently of the validity judgement latency measure (what is called RT2b here), while keeping everything else the same.

The results of Experiment 2 largely replicated those of Experiment 1. As in Experiment 1, the RT1 data showed that, as predicted, first segments ending with *but* and *so* were processed more slowly, relative to those ending with *and*. In fact, this difference in RT1 was proportionally larger than it was in Experiment 1. Moreover, one could see that the speeds of processing RT1's were

generally faster in Experiment 2 (one can compare **Figures 6** and **16**). Both of these observations can be explained by the fact that participants carried a slight cognitive load in Experiment 1 – in the form of maintaining a word in memory – but not in Experiment 2. Experiment 2's modification allowed the RT1 to collect a less encumbered measure. It also led to a clearer reaction-time difference between *and*, on the one hand, and *but* and *so*, on the other.

The reading time of the second segment, RT2a, also revealed that affirmative segments preceded by *but* and *so* are generally linked to longer reaction times than those preceded by *and*. This is remarkable, in that participants in the Experiment have yet to see the three-letter word when the RT2a is recorded. The main effect of Connective can thereby be attributed to participants placing the procedures of contrast and causality in play via *but* and *so*, respectively.

It is the interactions in this context that point to a discourse connective's procedure. Both RT2a and RT2b revealed a positive interaction between Connective (*but*) and Sentence type (negative) indicating that the presence of *but*, prior to the contrasting negation, not only facilitates its processing (RT2a), but also facilitates the task of recalling and verifying the sentence (RT2b), relative to when the negation is introduced by *and*. RT2a's are particularly fast when *but* is followed by a Negative second segment. In other words, the presence of *but*, prior to the contrasting negation (as in *but… there is no T*), facilitates its processing. As we indicated after Experiment 1, this is consistent with our account, because we propose that a procedure intrinsic to *but* prompts participants to seek out a contrast, which is provided by the negation. This precise interaction continues to resonate in the RT2b measures (which are the latencies linked to validity judgements).

One new interaction revealed itself in Experiment 2 which was not seen earlier: Participants were particularly fast at making False judgements in the wake of sentences containing *so*. One possible explanation for this effect is that, when reading the two segments of the trial sentences, the notion of causality carried by *so* encourages participants to look for a causal link. In doing so, participants are arguably considering possible words that comprise the two letters mentioned in the sentence. This extra effort (before seeing the three-letter word) pays off downstream when they quickly recognize words that do *not* match the letters in the trial sentence.

The Answer Accuracy results in Experiment 2 provided very high rates of correct responding, as in Experiment 1. The analyses pertaining to Answer Accuracy provided two effects in Experiment 2 that diverged slightly from those in Experiment 1. First, the connective *but* was linked with reduced accuracy, relative to *and*-trials. The plots of the mean accuracy results (see **Figure 14**) showed that this main effect was largely due to a drop in accuracy in positive test sentences (the baseline in our statistical model), but not in negative ones. Second, *so*-trials were linked with lower Answer accuracy in sentences with Negative second segments. These two effects were not predicted, but they could indicate that design changes of Experiment 2 allowed the extra-logical procedures to play a bigger role in participants' processing, to a point that it affected validity

judgments. For *but*-trials, the absence of contrast in positive sentences may have complicated their evaluation. For *so*-trials, as can be seen in the RT2b results, participants arguably adopted a predicted behavior that could have helped the evaluation of the sentence but that would have been difficult to adopt in negative sentences. We point out that, despite these new effects, rates of correct responses in Experiment 2 remain as high overall as those reported in Experiment 1.

To summarize, Experiment 2 replicated and reinforced the results of Experiment 1. The two experiments, indeed, provide evidence that discourse connectives trigger discourse expectations that can facilitate processing downstream but that also result in a specific additional processing cost on the connective.[11]

4. General discussion

We ended Section 1 by asking two questions. We first asked whether a theoretical distinction between *and*, which is a truth functional connective, and *but* and *so*, which are connectives that come with specific additional procedural meanings, leads to more intensive (and unique) processing for the latter two. This proposal took its cues from Relevance Theory (RT) and its notion of *procedural meaning*, according to which extra-logical meaning intrinsic to discourse connectives, such as *but* and *so*, prompts the hearer to carry out an inferential process (Blakemore, 2002). Upon reviewing the existing experimental literature, we came to the provisional conclusion that there does, indeed, appear to be support for such a distinction. Our second general question asked whether the above distinction (between *and*, on the one hand, and *but* and *so*, on the other) would persist even if the playing field could be neutralized with respect to thematic content. One noticeable feature of the literature is that all the existing experimental paradigms rely on rich material and the relations that they implicitly carry, e.g., a participant is expected to know that failure on an exam is normally not a cause for celebration or that a person who is urgently awaiting the arrival of prescription glasses will head to an optician. That is, investigations do not divorce the influence of the conceptual content of the segments, which themselves arguably create rhetorical relations, from those generated by the discourse connective. Would our expected experimental effects reveal how discourse connectives are processed even when they are not embedded into rich scenarios that typically rely on world knowledge? By all appearances, it appears that we were successful in isolating the effects of procedural meaning, even in thematically free contexts.

---

[11] After completing our analyses of Experiment 2, we noticed that, in preparing the Experiment, we had unintentionally created warm up materials that were slightly different from those of Experiment 1 (i.e., training did not include *but* or *so* items). Upon noticing this, we carried out Experiment 2 again, using Experiment 1's warm up materials. The outcomes of the follow up closely matched what was reported in Experiment 2, which is why we stick with the original findings. Also, see Larralde, Moyer, and Noveck (2024), which describes a French version of Experiment 2 that also yields findings that are similar to those reported here.

In our paradigm, critical test sentences vary mostly with respect to the discourse connectives themselves (plus polarity of the information in the second segment). Both of our Experiments provide evidence indicating that *but* and *so* prompt procedures that come with costs, relative to *and*; moreover, each of the connectives – *but*, *so* and *and* – further provide distinctive processing profiles, as can be seen by the reaction times that follow a connective across two (affirmative and negative) conditions. The two experiments ultimately show that the theoretical distinction between the truth-functional *and* and the procedurally-rich *but* and *so* can be manifested psychologically in tightly controlled, and thematically-free, experiments.

Our three main findings can be summarized as follows. First, the extra-logical import intrinsic to *but* and to *so* does not unduly affect participants' truth-evaluations of the critical test sentences. Even when a discourse connective prompts gratuitous processing (e.g., when *but* is ultimately presented without an obvious contrast to exploit), error rates across all conditions remain between 4 and 7 per cent. This indicates that truth values remain paramount when sentences are conjoined by any of the three discourse connectives investigated here. We take this to indicate that the conjunctive meaning intrinsic to each of the three connectives is not defeasible.

Second, the connectives *but* and *so* are more effortful to process than *and*, and from the very moment they are encountered. This was our pre-registration's main prediction, and it was clearly confirmed in both Experiments. Importantly, this difference is replicated even under the conditions of Experiment 2, where the truth value of the first segment is not yet determined when read. In fact, the difference between *but* and *so*, on the one hand, and *and*, on the other, was much more pronounced in Experiment 2, where participants processed the first segment without knowing anything about the upcoming word. Another way to view this is that the first segment was presented while participants were operating under a much reduced cognitive load (i.e., they did not need to store the word while processing it). We take this finding to justify the hypothesis that the reading of *but* and *so* necessarily prompts a procedure that does not exist in *and*. This supports our initial hypothesis, which is that the procedurally-rich connectives, being more informative, entail *and*.

Third, the processing related to the post-connective (second) segment is impacted by the connective used, and in specific (and predictable) ways. With respect to *but*, for example, we found that participants' processing of second-segments was facilitated when it was phrased negatively (which, when one considers the literature viewed more widely, typically comes with a further cost of processing), as opposed to when it was phrased affirmatively. When there *is* a contrast to exploit, it leads to second-segment reading times that are at least comparable to cases that follow *and* (when that connective is followed by a negative second segment). This is telling, because it indicates that the procedural meaning of *but* involves anticipating a contrast. When there is no contrast to exploit, it arguably prompts gratuitous processing. That we find these

kinds of effects downstream from discourse connectives in what are merely stative predicate sentences is a testament to the procedural meaning carried by each.

Now that we have addressed our two initial questions and summarized our main findings, we can determine how well our theoretical position is supported. We do so as we characterize the processing profiles of each of the connectives. We begin with *and*, before turning to *so* and, finally, to *but*.

4.1 Summarizing the effects of *and*, *but* and *so*

The connective *and* has served mostly as a control in our experiments, because our working hypothesis has been that it does not *come with* "extralogical information"; meanwhile, the conjunction intrinsic to *and* is shared with the other two discourse connectives investigated here. Empirical support for our argument can be seen from the fact that *and* is generally associated with (a) the highest rates of accuracy and (b) the fastest reaction times with respect to both segments. In light of exchanges with reviewers and audience members, we consider two other ways to view the data generated by *and*.

One alternative account, described in Section 1, is to view *and* as generating an (albeit weak) procedure that is considered part and parcel of *but* and *so*, as well. More specifically, *and* can be viewed as having a procedure that indicates that the two conjuncts are to be processed as a single unit. This makes all three of our studied connectives procedural, but with *and* still appearing to be the least effortful. Under this view, the three connectives studied here can be put on a continuum, with *and* clearly on the easy end, and *but* and *so* being more cognitively demanding. We will consider the first continuum-like view when we later address suggestions that come from SDRT.

A second alternative account that is compatible with our data on *and* is to assume its procedure is not entailed by *but*'s or *so*'s (i.e., it is independent of the other two). Under this guise, the three discourse connectives do not necessarily share features of conjunction or a fundamental procedural step; each of the three have independent procedures, and *and*'s happens to be easier to process. We consider this second alternative unlikely, but it is a logical possibility nonetheless.

We now take a closer look at the individual contribution of the connective *so*. So far, we have been referring to the extra-logical import of *so* as "causal", without making a strong theoretical commitment with respect to our word choice. Our motivation has been to underline how *so* provides a causal link between the two segments and to show how this contribution is "extra-logical." The exact nature of the discourse relation has not been relevant to our aims. Here, we spell out in greater detail how we view the causality carried by the connective *so* in our task.

The literature underlines that *so* can convey two types of causal relations. The first is the cause-consequence relation that is exemplified in (1c), where the first conjunct (*Camille having*

a lab meeting) is presented as the cause of the second conjunct (*his being late*) that consequently follows. This connective is known as *objective* or as *causal*, as is described in T. J. M. Sanders et al. (1992); see their Table 1). The other reading of *so* is to describe a speaker's reasoning as is exemplified in (1c'). Here, the first conjunct (about Camille's lab meeting) describes the speaker's justification for the second one (*Camille must be dealing with an emergency*).

(1) c. Camille has a lab meeting so he is running late.
c'. Camille has a lab meeting so there must be an emergency.

Used in this way, this connective is known as *subjective* or *diagnostic* and is referred to as Claim-argument in T. J. M. Sanders et al.'s (1992) taxonomy (again, see their Table 1). For further discussion, see (Canestrelli et al., 2013).

Of the two, one is eliminable in the context of our paradigm. Given that the propositional contents of the two segments in our task are not naturally related (it is difficult to argue that world-knowledge in our task prompts cause-consequence inferences that lead to valid conclusions about a second letter), it follows that *so* is not expressing objective causality here. On the other hand, it is plausible to consider that participants interpret the speaker's use of *so* in a statement such as *There is a B so there is a T* to mean that the speaker *has a reason* for making a link about the co-presence of B and T (even if it might be beyond the participant's ken to understand what link the speaker has in mind). It follows that the causal *so* in the context of our task can be best described as subjective.

Generally, our data show that the connective *so* does not create confusion for participants, even if it is arguably the case that the speaker's reason for using it is not apparent. Rates of correct responses with sentences conjoined by *so* are consistently high. Assuming that a procedure linked to *so* is prompted (and understood as a diagnostic or subjective causal), its items provide latencies that generally resemble those found for *and*, while being generally slower with respect to both segments (see **Figures 10** and **18**). We take this to indicate that participants are working out the meaning of *so* from the moment it is encountered and onward (through the presentation of the second segment).

There exist conditions under which *so* actually facilitates reaction times (compared to the baseline created by *and*), which occur when participants in Experiment 2 make decisions (post-*so*) about the presented word. That is, when *so* was employed and participants had received negative second segments, the test-words that follow prompt decision times that are faster than those that come after *and* (see **Figure 20**). Relatedly, when *so* is the connective and the presented word calls for a false judgment, decision times were faster than those associated with *and* (see **Figure 21**). This decision-making speed-up linked to *so* corroborates our argument that it prompts a procedure that is ultimately cashed out when given the opportunity. Given that we did not specifically set out to make such specific claims related to *so* (i.e., our experiments were not designed to test such effects), we remain cautious about our account of downstream processing

linked to *so*. Nevertheless, these findings are consistent with our claim that *so* ultimately provides benefits through a procedure, even if it is relatively costly to process initially.

These findings resonate with data on causal connectives from the text and discourse literature. There, one finds evidence showing that causal connectives, which are hypothesized to increase "inference generation," provide benefits to participants' comprehension when compared to a variety of control conditions that do not contain causal connectives. For example, Millis et al. (1995) showed that the causal connective *because* provides comprehension advantages to participants when compared to (a) cases where there is no connective, as well as to (b) cases where the causal is replaced by, and compared to, *and*. To make this clear, consider the two "statements" in *Sherry quickly mopped up the soup on the wooden floor* (Statement 1) *because a waiter accidentally tripped on the step by the kitchen* (Statement 2). When *because* is included, it provides added value to comprehension in several ways. For the sake of brevity, we will mention just two of those ways. First, performance on lexical decision tasks indicates that statements pertinently conjoined by *because* facilitate the speed of making lexical decisions when compared to the control condition, which does not contain any connective.[12] Second, participants answer comprehension questions more quickly (and slightly more accurately) when the sentences are conjoined with a causal connective. Interestingly, the authors found similar advantages for *because* when it was compared to *and*, much as we do. The authors concluded that the connective *because* elicits causal inferences, whereas *and* does not.

Much as with *so*, we did not have a preconceived view on the connective *but* before we began our investigation, other than to posit that this connective prompts a procedure and that its extra-logical import can, at a minimum, be described as *contrastive*. That said, we were well aware that its procedure could also be described in more inferential terms – as a way to express *denial of expectations* – where *but* signals that the first conjunct raises expectations that are denied by the second conjunct. These two accounts do not exhaust the range of possible accounts of *but*, but they are generally considered the two most prominent (for a brief historical summary of each, see Winterstein, 2012).[13] If the two are viewed as rival hypotheses, the *denial of expectations*

[12] In lexical decision tasks, participants are required to determine whether a stand-alone word (that appears in an instant after the two statements are read) is spelled correctly (with respect to the current example, the authors aptly used the word *spill*). To measure "inference encoding," the word is deployed across two conditions: when a causal inference linked to the word is expected to be generated and when a causal inference related to the word is assumed to *not* be generated (in this "unrelated" condition, the same critical word *spill* follows a different pair of statements whose causal inference is arguably different). In Millis et al. (1995), participants were 27 ms faster in lexical decision making when the word was appropriate (as opposed to inappropriate) to the causal inference that was presumably generated in the *because* condition; in contrast, there was no notable change in speed across the two conditions when there was no connective.

[13] There is a wide array of accounts on concessives (see Anscombre & Ducrot, 1977; Iten, 2000; Kehler, 2002; Umbach, 2005; Winter & Rimon, 1994; Winterstein, 2010, 2012) and testing the – often subtle – differences between them was not among our goals. As with *and* and *so*, we simply share how *but* is plausibly understood in light of our study.

description does not readily apply to our paradigm, because it is not obvious how the mention of one letter can raise expectations about a second (the choice of letters was generally randomized). This leads to the conclusion that the procedure linked to *but*, is indeed, best described as providing a contrastive meaning.

The data bear this out. Across both experiments, one finds that obvious contrasts – those made by adding negation in the second segment, as in *There is a B but there is not a…* – prompt reading times or decision times for the second segment that are comparable overall to those conjoined by *and* (see **Figures 10** and **18**). Furthermore, estimated parameters from our Bayesian analyses show that speed-ups arise when *but* is followed by a negative element (see **Figures 9, 17, 19**). A role for contrasts is not limited to affirmative/negative contrasts. In Experiment 1, one finds that (reading times and) responses are accelerated when the second segment renders the statement false (see **Figure 9**). For example, when participants receive the word *BET* and the sentence *There is a B but there is a V*, the combined reading and decision times (from the second segment onward) speed up, based on expectations derived from true statements that use *and* and affirmative statements. Here, the contrast is arguably cashed out by having a true segment followed by a false one. One also finds a weaker version of this effect in **Figure 19**, when the word comes after the two segments are read. The search for this (or any) contrast is, in our view, generated by the presence of *but*.

Another relevant data point is the way participants handle cases that do not provide an obvious contrast. For example, consider when participants read the entirely affirmative sentence *There is a B but there is a T* when it is preceded by the word *BET* in Experiment 1 (or followed by that word in Experiment 2). In this case, there is no obvious contrast to be drawn, and the prolonged latencies (see **Figures 10** and **18**) reveal an effort on the part of participants to seek a contrast nonetheless. As with *so*, the discourse connective *but* in these cases compels participants to understand why the speaker chose that connective. From our point of view, the gratuitous processing (when the instantiation of *but* does not lead to an obvious contrast) is a feature (and not a bug) of our paradigm. The slowdowns point to the inexorable application of *but*'s procedural meaning without affecting its conjunctive contribution.

4.2 Addressing SDRT

We would be remiss at this point if we did not address how SDRT, arguably the best known semantic account of discourse relations in the linguistic literature, would account for this article's data, and we begin with our characterization of *and*. In our view, advocates of SDRT would be agreeable with the first of the alternative accounts we presented above (in which *and* is considered to have a procedural meaning, albeit weaker than *but*'s and *so*'s). To explain how, we take a step back to describe SDRT in slightly more detail.

SDRT describes discourse relational structures (SDR's) as containing discourse topics, representing the content in clauses, that can be viewed hierarchically. For example, as Asher and Vieu (2005) describe in an example, a speaker can describe *his lovely evening* by elaborating on it (e.g., the speaker can add that he *had a great meal*). The clause about the meal is an *Elaboration* that is subordinate to the lovely evening. The *had a great meal* segment can then be an antecedent to another subordination (e.g., the speaker could elaborate on what he ate). However, the *having of a great meal* clause could also be the antecedent to a *Narrative* relation, a *coordination*, wherein the speaker can go on to say that she won a dance competition. According to advocates of SDRT, this distinction between subordinating and coordinating is so consequential that, once a speaker is understood as coordinating, e.g., via *and*, it precludes the anticipation of a subordination. This would be *and*'s contribution to discourse, which arguably goes beyond its logical contribution (for an in-depth discussion, see Txurruka, 2003).

Can the SDRT approach account for our findings? To some degree, it does. As one reviewer pointed out, *but* and *so* are more specific than *and*, because *and* allows for all types of coordination, while *but* and *so* only allow for a subset of them. That an added restriction leads to extra processing is sensible and would make such an outcome compatible with our main prediction. While it is our view that our data speak positively to the SDRT account, this kind of prediction would be novel for SDR Theory-making, in that it is relies on real time processing concerning the connective itself.

That said, given its reliance on the role played by conceptual content, it would be difficult for SDRT to account for our findings in other respects. That is, Rhetorical Relations (whether they be coordinating or subordinating) rely on the semantic, truth-conditional content of each clause. It is through the truth-conditional content that one determines that *having a great meal* and *winning a dance competition* are part of a coordinating phrase, just as it is through content that one knows that *having a great meal* and *eating salmon* have a subordinating relationship. The narrator's choice of connective necessarily coalesces with the content. In contrast, our participants are processing our sentences and making inferences based almost entirely on the discourse connective itself. It could be said that, in Experiment 1, participants have the means to at least determine whether the letters are true or false as they read the sentence (since the word comes first, they rely on the truth-conditional content to some degree). However, in Experiment 2, participants process the critical sentences without knowing whether the content of the sentence is true (because the word comes after the clue). So, in Experiment 2, there is no reliance on content to determine the relationship across the two segments. Processing depends entirely on the presence of the connectives *and*, *but* and *so*. Remarkably, whether or not the truth-conditional content is known as the test sentence is presented, participants' processing outcomes across the two Experiments are very similar.

4.3 Final thoughts

Now that we have concluded that *but* and *so* provide extra-logical information, when compared to *and*, how shall one characterize this extra information? Should it be viewed as part of a complex semantic representation, or does this procedure qualify as pragmatic? This question is relevant to the extent that different theoretical accounts are likely to make very different claims about our findings, so let us briefly spell out what we consider these accounts to be.

One possible answer is that *but* and *so* provide an extra layer of semantic meaning, one that is added to the at-issue meaning of the utterance, as has been argued by authors such as Bach (1999), Potts (2005) and Karttunen and Peters (1979). While such accounts aim to address, or to formalize, the meaning conveyed by discourse connectives such as *but* and *so*, they are generally not directly concerned with the psychological realisation of that meaning. As such, it is unclear what these approaches would predict in terms of processing. The most plausible prediction seems to be that semantic meaning is automatic and that the additional layer of semantic meaning contained in these discourse connectives should not cause them to be processed more slowly than logical connectives. Moreover, we do not see how a semantic approach to discourse connectives would account for the very specific processing profiles that each appears to produce (as seen with negations following *but* and for false statements containing *so*). Semantic accounts do not distinguish between the conceptual content in the clauses and the procedures generated by the connectives (Bezuidenhout, 2004). With this view, it is difficult to see how one can account for any of the reported differences – both with respect to accuracy and to processing – as reported in the current work.

This leads to a second possible answer concerning the nature of the extra-logical contribution, which puts it on the pragmatic side and which hews more closely to our hypothesis. Our explanation is that, while all three terms are by definition conjunctive, the procedural meaning inherent in *but* and *so* intervenes at the pragmatic level, while *and* remains primary and semantic. This distinction is motivated by two features that have come up in the above discussion. One is that extra-logical procedural information in *but* and *so* puts the two in a position to entail *and*. The other is that the nature of procedural information is such that there are ways to direct a listener to make – and for an addressee to understand – implicit connections among clauses that would not be available *without* such procedures. For example, let us revisit (1b) and consider a reading that is inferred in (1b'), due the inclusion of the connective *but*:

(1) b. Camille has a lab meeting but he is running late.
b'. Camille [is anticipated to be at] a lab meeting but [his apparent absence is due to the fact that] he is running late.

The inferred information in brackets arises through the speaker's introduction of *but*, at which point the addressee aims to account for a contrast of some sort. In other words, the presence

of *but* implicitly prompts the addressee to make adjustments that lead to modifications of the conceptual information. Ultimately, this modification integrates a contrast.

The inferred information in (1b') is, moreover, cancelable. In fact, *but* in this context could easily generate other interpretations. For instance, one could imagine that the speaker is a Masters student explaining to a fellow student that "Camille is in fact *at* the lab meeting, which is running longer than expected". Here, the contrast that gets manifested is between the predicted and actual length of the meeting. These multiple interpretations arise because the conceptual content, in combination with *but*, generates them. The same holds for *so*.

Here, we turn to and comment on our methodological approach, which reduced contextual information as much as possible. This could be valuable for future research, because this approach puts one in a position to introduce some complexity found in natural discourse in a controlled step-by-step way. For instance, it is known that the question under discussion (QUD) helps both the speaker and the hearer to, respectively, tailor their utterance and work out the intended meaning (Degen & Tanenhaus, 2019). In experimental studies, QUD manipulations have helped refine the understanding of well known pragmatic processes (Degen & Goodman, 2014; Potts, 2012; Sedivy et al., 1999). As far as discourse connectives are concerned, QUD manipulations could help us understand what, exactly, is causing a delay in the processing of *but* or *so* relative to *and*. In other words, this paradigm would allow for future work to consider other theoretical frameworks to determine the extent to which a discourse connective, such as *but*, generates relatively effortful processing in cases where the QUD triggers an expectation of contrast.

To summarize, this study provides evidence showing that discourse connectives (such as *but* and *so*) convey information that make them costlier to process than the truth functional logical conjunction (*and*). We argue that this additional cost can be viewed through a pragmatic lens. According to Relevance Theory, for example, our results make sense, because discourse connectives prompt the involuntary initiation of a pragmatic procedure. Theoretically, this extra cognitive effort should also come with a benefit, which we assume is the facilitation of processing downstream.

Data accessibility statement

The lists of test items, the result files and the analysis scripts for Experiment 1 and 2 are available to download from the OSF page for this study.

Ethics and consent

This study was approved by the Ethics Chair for the Linguistics Research Department at University College London under the project ID number LING-2021-6-26.

Funding information

The first, second and last author were supported by a *Chaire d'excellence* Grant (2020) awarded to the last author from *Université Paris Cité*.

Competing interests

The authors have no competing interests to declare.

Authors contributions

The authors made the following contributions. CL: Conceptualization, Writing – Original Draft Preparation, Writing – Review and Editing, Formal Analysis, Visualization, Data curation, Methodology, Software; MM: Review and Editing, Data Analysis, Data curation; NP: Review and Editing, Supervision; IN: Original conceptualization, Methodology, Writing – Original Draft Preparation, Writing – Review and Editing, Supervision, Funding acquisition, Project administration.

References

Anscombre, J. C., & Ducrot, O. (1977). Deux "mais" en français? *Lingua*, *43*(1), 23–40. https://doi.org/10.1016/0024-3841(77)90046-8

Anscombre, J.-C., & Ducrot, O. (1989). Argumentativity and informativity. In M. Meyer (Ed.), *From metaphysics to rhetoric* (pp. 71–87). Springer. https://doi.org/10.1007/978-94-009-2593-9_6

Asher, N., & Vieu, L. (2005). Subordinating and coordinating discourse relations. *Lingua*, *115*(4), 591–610. https://doi.org/10.1016/j.lingua.2003.09.017

Bach, K. (1999). The myth of conventional implicature. *Linguistics and Philosophy*, *22*(4), 327–366. https://doi.org/10.1023/A:1005466020243

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bezuidenhout, A. (2004). Procedural meaning and the semantics/pragmatics interface. In C. Bianchi (Ed.), *The semantics/pragmatics distinction* (pp. 101–131). CSLI Publications.

Blakemore, D. (1987). *Semantic constraints on relevance*. Wiley–Blackwell.

Blakemore, D. (2000). Indicators and procedures: "nevertheless" and "but". *Journal of Linguistics*, *36*(3), 463–486. https://doi.org/10.1017/S0022226700008355

Blakemore, D. (2002). *Relevance and linguistic meaning: The semantics and pragmatics of discourse markers*. Cambridge University Press. https://doi.org/10.1017/CBO9780511486456

Blakemore, D., & Carston, R. (1999). The pragmatics of "and"-conjunctions: The non-narrative cases. *UCL Working Papers in Linguistics*, *11*.

Blakemore, D., & Carston, R. (2005). The pragmatics of sentential coordination with "and". *Lingua*, *115*(4), 569–589. https://doi.org/10.1016/j.lingua.2003.09.016

Bloom, L., Lahey, M., Hood, L., Lifter, K., & Fiess, K. (1980). Complex sentences: Acquisition of syntactic connectives and the semantic relations they encode*. *Journal of Child Language*, *7*(2), 235–261. https://doi.org/10.1017/S0305000900002610

Breheny, R., Katsos, N., & Williams, J. (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. https://doi.org/10.1016/j.cognition.2005.07.003

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science*, *27*(1), 45–50. https://doi.org/10.1177/0963721417727521

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395–411. https://doi.org/10.32614/RJ-2018-017

Canestrelli, A. R., Mak, W. M., & Sanders, T. J. M. (2013). Causal connectives in discourse processing: How differences in subjectivity are reflected in eye movements. *Language and Cognitive Processes*, *28*(9), 1394–1413. https://doi.org/10.1080/01690965.2012.685885

Davies, M. (2008). *Word frequency data from The Corpus of Contemporary American English (COCA)*. https://www.wordfrequency.info.

Degen, J., & Goodman, N. (2014). Lost your marbles? The puzzle of dependent measures in experimental pragmatics. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 36).

Degen, J., & Tanenhaus, M. K. (2019). Constraint-based pragmatic processing. In C. Cummins & N. Katsos (Eds.), *The Oxford handbook of experimental semantics and pragmatics*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198791768.013.8

Drenhaus, H., beim Graben, P., Saddy, D., & Frisch, S. (2006). Diagnosis and repair of negative polarity constructions in the light of symbolic resonance analysis. *Brain and Language*, *96*(3), 255–268. https://doi.org/10.1016/j.bandl.2005.05.001

Drenhaus, H., Zimmermann, M., & Vasishth, S. (2011). Exhaustiveness effects in clefts are not truth-functional. *Journal of Neurolinguistics*, *24*(3), 320–337. https://doi.org/10.1016/j.jneuroling.2010.10.004

Fraser, B. (1990). An approach to discourse markers. *Journal of Pragmatics*, *14*(3), 383–398. https://doi.org/10.1016/0378-2166(90)90096-V

Fraser, B. (1996). Pragmatic markers. *Pragmatics*, *6*(2). (Number: 2)

Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics*, *31*, 931–952. https://doi.org/10.1016/S0378-2166(98)00101-5

Frege, G. (1892). *On sense and reference*.

Grice, H. P. (1975). Logic and conversation. In M. Ezcurdia & R. J. Stainton (Eds.), *The semantics-pragmatics boundary in philosophy* (p. 47). Broadview Press.

Hall, A. (2004). The meaning of "but": a procedural reanalysis. *UCL Working Papers in Linguistics*, *16*, 199–236.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English* (No. 9). Longman.

Hussein, M. M. (2009). *Relevance theory and procedural meaning: The semantics and pragmatics of discourse markers in English and Arabic* [Thesis]. Newcastle University.

Iten, C. B. (2000). *'Non-truth-conditional' meaning, relevance and concessives* [Thesis]. University College London.

Jasinskaja, K., & Karagjosova, E. (2020). Rhetorical relations. In *The Wiley Blackwell companion to semantics* (pp. 1–29). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118788516.sem061

Karttunen, L., & Peters, S. (1979). *Conventional lmplicature*. Brill. https://doi.org/10.1163/9789004368880_002

Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. CSLI Publications.

Koornneef, A. W., & Sanders, T. J. M. (2013). Establishing coherence relations in discourse: The influence of implicit causality and connectives on pronoun resolution. *Language and Cognitive Processes*, *28*(8), 1169–1206. https://doi.org/10.1080/01690965.2012.699076

Köhne-Fuetterer, J., Drenhaus, H., Delogu, F., & Demberg, V. (2021). The online processing of causal and concessive discourse connectives. *Linguistics*, *59*(2), 417–448. https://doi.org/10.1515/ling-2021-0011

Larralde, C., Moyer, M., & Noveck, I. (2024). Discourse markers and psycholinguistic processing. In M.-B. Mosegaard Hansen & J. Visconti (Eds.), *Manual of discourse markers in Romance*. (De Gruyter ed.). https://doi.org/10.1515/9783110711202-014

Martin, J. R. (2002). 3. meaning beyond the clause: SFL perspectives. *Annual Review of Applied Linguistics*, *22*, 52–74. https://doi.org/10.1017/S026719050200003X

Millis, K. K., Golding, J. M., & Barker, G. (1995). Causal connectives increase inference generation. *Discourse Processes*, *20*(1), 29–49. https://doi.org/10.1080/01638539509544930

Morís Fernández, L., & Vadillo, M. A. (2020). Flexibility in reaction time analysis: Many roads to a false positive? *Royal Society Open Science*, *7*(2), 190831. https://doi.org/10.1098/rsos.190831

Noveck, I. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, *78*, 165–88. https://doi.org/10.1016/S0010-0277(00)00114-1

Noveck, I. (2018). *Experimental Pragmatics: The making of a cognitive science*. Cambridge University Press. https://doi.org/10.1017/9781316027073

Noveck, I. A., & Chevaux, F. (2002). The pragmatic development of "and". In *Proceedings of the 26th annual Boston University Conference on Language Development*. Cascadilla Press.

Potts, C. R. (2005). *The logic of conventional implicatures*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199273829.001.0001

Potts, C. R. (2012). Goal-driven answers in the Cards Dialogue Corpus. In N. Arnett & R. Bennett (Eds.), *Proceedings of the 30th West Coast Conference on Formal Linguistics* (pp. 1–20). Cascadilla Press.

R Core Team. (2021). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria.

Recio Fernández, I., Loureda, Ó., & Cruz, A. (2023). Adjustment, mismatches and accommodation of procedural and conceptual meaning: Experimental evidence around connectives. *Functions of Language*, *30*(1), 110–135. https://doi.org/10.1075/fol.22020.rec

Russell, B. (1905). On Denoting. *Mind*, *14*(56), 479–493. https://doi.org/10.1093/mind/XIV.4.479

Sanders, T. (2005). Coherence, causality and cognitive complexity in discourse. In M. Aumague, M. Bas, A. Le Draoulec, & L. Vieu (Eds.), *Proceedings of the First International Symposium on the Exploration and Modelling of Meaning (SEM-05)* (pp. 105–114). Université de Toulouse le Mirail.

Sanders, T. J. M., & Noordman, L. G. M. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes*, *29*(1), 37–60. https://doi.org/10.1207/S15326950dp2901_3

Sanders, T. J. M., Spooren, W. P. M., & Noordman, L. G. M. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, *15*(1), 1–35. https://doi.org/10.1080/01638539209544800

Schiffrin, D. (1987). *Discourse markers*. Cambridge University Press. https://doi.org/10.1017/CBO9780511611841

Schiffrin, D. (2005). Discourse markers: Language, meaning, and context. In *The handbook of discourse analysis* (pp. 54–75). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470753460.ch4

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*(2), 109–147. https://doi.org/10.1016/S0010-0277(99)00025-6

Sperber, D., & Wilson, D. (1986/1995). *Relevance: Communication and cognition*. (2nd ed.). Blackwell Publishers.

Spooren, W., & Sanders, T. (2008). The acquisition order of coherence relations: On cognitive complexity in discourse. *Journal of Pragmatics*, *40*(12), 2003–2026. https://doi.org/10.1016/j.pragma.2008.04.021

Txurruka, I. G. (2003). The natural language conjunction 'and'. *Linguistics and Philosophy*, *26*(3), 255–285. https://doi.org/10.1023/A:1024117423963

Umbach, C. (2005). Contrast and information structure: A focus-based analysis of 'but'. *Linguistics*, *43*(1). https://doi.org/10.1515/ling.2005.43.1.207

Wilson, D. (2016). Reassessing the conceptual–procedural distinction. *Lingua*, *175–176*, 5–19. https://doi.org/10.1016/j.lingua.2015.12.005

Winter, Y., & Rimon, M. (1994). Contrast and implication in natural language. *Journal of Semantics*, *11*(4), 365–406. https://doi.org/10.1093/jos/11.4.365

Winterstein, G. (2010). *La dimension probabiliste des marqueurs de discours. Nouvelles perspectives sur l'argumentation dans la langue* [Unpublished doctoral thesis]. Paris 7.

Winterstein, G. (2012). What "but"-sentences argue for: An argumentative analysis of "but". *Lingua*, *122*(15), 1864–1885. https://doi.org/10.1016/j.lingua.2012.09.014

Xiang, M., & Kuperberg, G. (2015). Reversing expectations during discourse comprehension. *Language, Cognition and Neuroscience*, *30*(6), 648–672. https://doi.org/10.1080/23273798.2014.995679

Zehr, J., & Schwarz, F. (2018). PennController for Internet Based Experiments (IBEX). https://doi.org/10.17605/OSF.IO/MD832

Ziti, A., & Champagnol, R. (1992). Effet des connecteurs sur le traitement en temps réel de propositions exprimant des relations de cause/effet. *L'Année psychologique*, *92*(2), 187–207. https://doi.org/10.3406/psy.1992.29502

Zwicky, A. M. (1985). Clitics and particles. *Language*, *61*(2), 283–305. https://doi.org/10.2307/414146