

# Processing causal and diagnostic uses of *so*

Sharmaine Seneviratne

Research Centre for English and Applied Linguistics, University of Cambridge, UK

<https://doi.org/10.36505/ExLing-2006/01/0048/000048>

## Abstract

This study compares the processing of causal and diagnostic uses of the connective *so*. In an experiment measuring reading times, it is shown that diagnostic *so* constructions, (e.g. *Carl has been limping lately. So he was injured in the accident.*) take longer to read than causal ones (e.g. *Carl didn't wear his seatbelt. So he was injured in the accident.*). This difference is reduced in parallel constructions where *so* is absent. It is proposed that readers are strongly biased towards a default causal reading of *so* and therefore anticipate a consequence when they encounter *so*. In diagnostic constructions, since *so* is followed by an antecedent, processing is more effortful and complex because the default expectation has to be reversed.

## Introduction

Connectives play an important role in discourse production and comprehension as they are used to signal how discourse segments are linked with their context. It is generally agreed that the connectives *so* and *because* aid utterance interpretation by cueing that a causal link is intended between the discourse segments that they link. However, the types of processes that are triggered by these linguistic cues are conceptualised differently in various theoretical frameworks. In this paper we consider two contrasting theoretical accounts that describe the function of *so*. We compare the processing of two types of statements, causal and diagnostic, and show that these accounts do not adequately explain the processing difference observed between these two uses. An alternative account, which posits a default setting for *so*, is proposed. Causal and diagnostic uses of *so* are illustrated in (1) and (2) below:

- (1) Carl didn't wear his seatbelt. So he was injured in the accident.
- (2) Carl has been limping lately. So he was injured in the accident.

In both these examples the discourse segments are causally linked. However, in (1), which is a basic causal use of *so*, the event mentioned first is given as the cause for the second event, which follows the connective *so*. In (2), Carl's limping is not the cause for Carl being injured. The relation in (2) is better understood as a diagnostic one, in which the first segment provides evidence for the deduction of the conclusion stated in the second segment.

### **Theoretical accounts of *so***

The meaning of *so* and the way in which it functions in integrating events have been described within the framework of Relevance theory (Blakemore 1988) and in Coherence theory (Sanders et al. 1992). According to the Relevance account, connective expressions have procedural meaning: they encode instructions on how propositional representations are manipulated in arriving at the intended meaning. The connective *so*, specifically, instructs the hearer to process the proposition it introduces as an implication or conclusion. In this account, hearers follow the same inferential route in understanding both causal and diagnostic utterances. Therefore it would not predict that interpreting one type of statement would involve more processing complexity and effort than the other. Furthermore, we are not told how and whether it is important to identify the cause/evidence and consequence/conclusion elements in interpreting these statements.

The Coherence theorists propose that coherence relations such as the ones cued by *so* can be decomposed to combinations of cognitive primitives which each allow binary options: basic operation (causal or additive), source of coherence (semantic or pragmatic), order (basic or non basic) and polarity (positive or negative). In the causal use in (1), according to the criteria listed, *so* signals the relation *cause-consequence*, which is decomposed as causal-semantic-basic-positive. The diagnostic use in (2), however, signals the relation *consequence-cause*, which has the configuration causal-semantic-non basic-positive. Since *so* can be used to cue both combinations it is unclear how hearers initially determine the option for the order element in order to arrive at the correct coherence relation via the combination of cognitively salient primitives.

Both these accounts do not explain how, given a single linguistic cue *so*, hearers recognise that the inference to be made is one that leads from evidence to a conclusion or one that links a cause with a consequence. We hold that it is crucial for hearers to make this distinction in arriving at the intended interpretation. We suggest that, in general, *so* is strongly associated with a basic causal reading and that hearers normally assume, upon hearing *so*, that a consequence will follow. A reading time experiment was designed in order to test this assumption. Since nothing in the constructions unambiguously signal which interpretation should be adopted, it was predicted that in reading diagnostic statements, this default assumption would have to be withdrawn, thus causing a delay in the reading time.

### **Experiment**

A self-paced, clause-by-clause reading experiment was used to measure reading times for identical second clauses in parallel causal and diagnostic

constructions (see Table 1). The two clauses were linked by the connective *so*, but they also appeared without the explicit marking of the connective.

Table 1: Sample items for experiment

Item Type	Sample Item
Marked Causal	Carl didn't wear his seatbelt. So he was injured in the accident.
Marked Diag.	Carl has been limping lately. So he was injured in the accident.
Unmarked Causal	Carl didn't wear his seatbelt. He was injured in the accident.
Unmarked Diag.	Carl has been limping lately. He was injured in the accident.

Forty native speakers of (British) English, undergraduate and graduate students at the University of Cambridge, took part in this experiment. The experimental items were 24 pairs of sentences similar to those in Table 1. Thirty-six filler items - pairs of sentences connected by *but* and *after all* and those that were not explicitly linked by a connective - were used. Twenty comprehension questions requiring yes/no responses were also included in order to ensure that readers continuously attempted to fully understand the sentences. There were two within-subjects variables: type of construction (causal or diagnostic), and presence or absence of connective. The items were rotated among the four conditions, so that each item was seen in a causal marked and unmarked condition and in a diagnostic marked and unmarked condition. Each participant saw a particular item under only one of the four conditions.

A self-paced reading task was run on a computer using Superlab Pro software. Participants clicked on the mouse button to initiate a trial, and viewed the first sentence of a pair on the screen. Once they read and understood the sentence, they clicked on the mouse to view the second sentence and again, to initiate the next trial. Time between each successive mouse click was recorded, and served as the dependent measure. Comprehension questions sometimes appeared after a pair of sentences, to which readers responded by clicking the left or right mouse button to answer either 'yes' or 'no'. These questions appeared at random intervals among the test items.

## Results and discussion

A 2×2 ANOVA was conducted to test the effect of item type and presence of marker. There was a strong interaction of item type and the presence of *so* in the subjects analysis  $F(1,40)=6.77$ ,  $p<0.01$ . As shown in Figure 1, reading times for marked diagnostic sentences was on average 366 milliseconds more than for the marked causal ones,  $T_1(40)=0.74$ ,  $p<0.0001$ ;  $T_2(24)=4.26$ ,  $p<0.01$ . In the unmarked conditions the difference between reading times for the causal and diagnostic sentences was not significant  $T_1(40)=0.60$ ,  $p>0.1$ ;  $T_2(24)=0.59$ ,  $p>0.1$ . In diagnostic sentences, explicit marking of *so* brought about slower reading times,  $T_1(40)=0.51$ ,  $p<0.0001$ ;  $T_2(24)=3.35$ ,  $p<0.002$ .

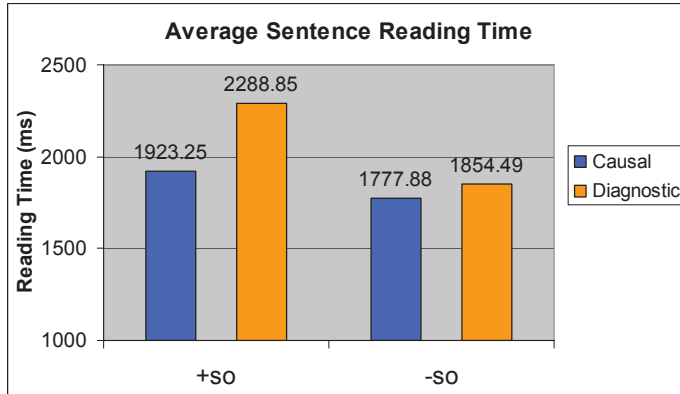


Figure 1: Average reading times for target clauses

The reading time differences show that causal and diagnostic uses of *so* have differential levels of processing difficulty. Furthermore, because diagnostic statements were easier to read when they were unmarked, it can be observed that the difficulty in processing diagnostic statements was at least partly due to an effect of *so*. This supports our hypothesis that the preferred reading of *so* is a basic causal one, in which *so* is expected to precede a consequence.

## Conclusions and Future work

This study shows that diagnostic and causal uses of *so* involve different levels of processing effort. While *so* may signal a single basic procedure or relation this information is not sufficient to interpret causally linked sequences. It appears that hearers attempt to identify cause and consequence elements which would then prompt different sets of generalisations and inferences in recovering causal and diagnostic interpretations. In work on first language acquisition that is currently in progress, we explore further evidence for a default causal reading of *so* and whether diagnostic reasoning follows a different and more complex inferential path.

## References

- Blakemore, D. 1988. 'So' as a constraint on relevance. In Kempson R. (ed.) 1988, *Mental Representations: The interface between language and reality*, 183-95. Cambridge, Cambridge University Press.
- Sanders, T., Spooren, W. and Noordman, L. 1992. Towards a taxonomy of discourse relations. *Discourse Processes* 15, 1-35.