

## ‘ONLY’ INCREASES EXPECTATIONS FOR CAUSAL COHERENCE RELATIONS

Katharina Spalek<sup>1</sup>, Beate Bergmann<sup>1</sup>, Wei Shi<sup>2</sup> and Vera Demberg<sup>2</sup>

<sup>1</sup>Humboldt-Universität zu Berlin, <sup>2</sup>Universität des Saarlandes

The focus particle ‘only’ is exclusive, that is, an utterance containing ‘only’ is true if it holds for the focused element but not for any of its alternatives. The discourse unit following a proposition with ‘only’ often discusses an alternative to the focused element. This is illustrated in example (1), where the focused constituent is written in capital letters and its alternative is underlined.

(1) Caroline only put on her RINGS. She couldn’t wear a necklace with her turtleneck jumper.

It has been shown both with corpus data (Spalek & Zeldes, 2017) and data from a story continuation task (Gotzner & Spalek, unpublished data), that alternatives are re-mentioned significantly more often after ‘only’ than after clauses without a focus particle (cf. Kaiser, 2010, for similar findings for corrections). Here, we test the hypothesis that speakers re-mention alternatives after ‘only’ because they want to explain their exclusion. If true, this will affect the coherence relation between a first argument containing ‘only’ and a second argument, increasing the likelihood for a CAUSE relation. We can test this hypothesis in two different ways: First, by looking at the distribution of coherence relations after an argument containing ‘only’, compared to cases without ‘only’. Second, some features (e.g., modals) in the local context can serve as cues for a particular coherence relation. According to uniform density hypothesis (Levy & Jaeger, 2007; Fenk & Fenk, 1980), a speaker should distribute information uniformly across an utterance. With respect to coherence relations, this predicts that the presence of one or more local cues for a given relation should make an explicit connective redundant (Asr & Demberg, 2012). Asr and Demberg (2015) show that negation serves as such a cue for the discourse relation CHOSEN ALTERNATIVE: The explicit connective ‘instead’ was dropped almost twice as often after a negation in the first argument than in first arguments without negation. Similarly, we predict that the presence of the focus particle ‘only’ will lead to a decrease in the number of explicit connectives (because, as, ...) for the CAUSE relation.

We investigated this question with a small set of story continuations and with corpus data from the discourse-annotated Penn Discourse Treebank (PDTB, Prasad et al., 2008). Before turning to the results, we will briefly present the story continuation task: The experiment was done in German. Twenty-four native speakers listened to 15 stories, 5 with no focus particle, 5 with ‘only’, and 5 with ‘even’ (‘even’ will not be analysed here), and 16 filler stories. Stories consisted of three sentences, the first one mentioned a list of three elements, the second sentence kept the subject of the first sentence fore-grounded and the third picked out one of the three elements (cf. (2), approximate translation, the original stimuli were in German). Participants orally formulated a story continuation, for example (3).

(2) Caroline looked at necklaces, rings, and brooches in her casket.

She considered what would go well with her outfit.

She has only/ {no focus particle} taken out the rings.

(3) Uhm, she was wearing a turtleneck jumper and therefore could not wear a necklace.

We coded the coherence relation of the continuations with respect to the critical third story sentence, following rhetorical structure theory (RST, Mann & Thompson, 1988). The 120 critical sentences with ‘only’ were followed by CAUSE in 27 cases (22.5%), whereas the 120 sentences without a focus particle were followed by CAUSE in 13 cases (10.8%). A binomial test showed that this difference in distribution was statistically significant ( $p = .04$ ).

In the PDTB, there were 293 cases of ‘only’ + NP in the first argument. Sixty-five of these were followed by CAUSE in the second argument (contingency.cause and contingency.pragmatic.cause, see Sanders et al., 2018, for how to map PDTB labels and RST labels). The number of CAUSE relations following ‘only’ was descriptively higher than the expected value of 52 under the assumption of a

uniform distribution of ‘only’ across coherence relations. However, this difference was not statistically significant ( $p = .27$ , binomial test).

Taken together, empirical and corpus data jointly indicate that the presence of ‘only’ increases the likelihood for CAUSE, even though the finding is only weakly supported statistically.

In a second step, we looked at the distribution of explicitly and implicitly marked CAUSE relations. If ‘only’ is a valid cue for CAUSE, there should be fewer explicit connectives than without ‘only’. Table 1 presents the data for the story continuation task, Table 2 for the corpus data. For story continuations, the distribution of coherence markers is in line with the hypothesis. However, Fisher’s exact test showed that the effect is not significant,  $p = .48$  (possibly due to the small number of observations). For the corpus data, we observed a significant shift in the distribution of explicit and implicit coherence markers (Fisher’s exact test,  $p = .009$ ).

Table 1. Distribution of explicit and implicit connectives for sentences containing ‘only’ and control sentences without a focus particle.

only (n = 120)		no focus particle (n = 120)	
explicit	implicit	explicit	implicit
8	19	6	7

Table 2. Distribution of explicit and implicit connectives for sentences containing ‘only’ and all CAUSE relations in the corpus. Numbers in brackets indicate occurrence per thousand.

only		all	
explicit	implicit	explicit	implicit
10 (5.5)	55 (13.2)	1826	4172

In sum, there is converging evidence from story continuations and corpus data that the exclusive focus particle ‘only’ leads to an increased prevalence of causal continuations, which may in turn cause listeners to anticipate an upcoming causal relation, and which was found to decrease the likelihood of observing an explicit causal marker. This is plausible in the stimuli we used: If a story mentions three elements and then continues with the information that only one of these is further considered, this begs the question what happened to the other two. Thus, a speaker will give a reason for excluding alternatives. This pattern also generalizes to the more varied language use in corpora. In addition, we provide evidence that ‘only’ serves as a local cue for CAUSE, making the explicit causal connective somewhat redundant, in accordance with the uniform density hypothesis, and in parallel to the findings reported by Asr und Demberg (2015) for negation and the CHOSEN ALTERNATIVE relation.

## References

- Asr, F. T. & Demberg, V. (2012). Implicitness of discourse relations. *Proceedings of COLING 2012* (pp. 2669-2684).
- Asr, F. T. & Demberg, V. (2015). Uniform surprisal at the level of discourse relations: Negation markers and discourse connective omission. *Proceedings of the 11<sup>th</sup> International Conference on Computational Semantics* (pp. 118-128).
- Fenk, A. & Fenk, G. (1980). Konstanz im Kurzzeitgedächtnis – Konstanz im sprachlichen Informationsfluss [Consistency in short term memory – consistency in the flow of linguistic information]. *Zeitschrift für experimentelle und angewandte Psychologie*, 27, 400-414.
- Kaiser, E. (2010). Investigating the consequences for focus on the production and comprehension of referring expressions. *International Review of Pragmatics*, 2, 266-297.
- Levy, R. P. & Jaeger, T. F. (2007). Speakers optimize information density through syntactic reduction. *Advances in Neural Information Processing Systems*.
- Mann, W. & Thompson, S. (1988). Rhetorical Structure Theory: Toward a functional theory of text organization. *Text*, 8, 243-281.
- Prasad, R. et al. (2008). The Penn Discourse TreeBank 2.0. *Proceedings of the 6<sup>th</sup> International Conference on Language Resources and Evaluation* (pp. 2961-2968).
- Sanders, T. J. M. et al. (2018). Unifying dimensions in coherence relations: How various annotation frameworks are related. *Corpus Linguistics and Linguistic Theory*.
- Spalek, K. & Zeldes, A. (2017). Converging evidence for the relevance of alternative sets: data from NPs with focus sensitive particles in German. *Language and Cognition*, 9, 24-51.