

Too strong to eliminate or negate is surely no illusion

Dario Paape (Universität Potsdam)
paape@uni-potsdam.de

The sentence *No head injury is too trivial to be ignored* is often interpreted to mean *Treat all head injuries, no matter how trivial*, even though the compositional meaning of the utterance is nonsensical (*Ignore all head injuries, no matter how trivial*). The pervasive misinterpretation was originally attributed to processing failure and subsequent superficial interpretation [1], but has more recently been argued to be due to a stored construction that licenses the non-compositional meaning [2,3].

We seek to answer two questions about the depth charge illusion:

1. Does the illusion still occur when word order is changed (*Too Y to Z is no X*; see page 3)?
2. Is the semantic representation of depth charge sentences detailed enough to be subjected to subsequent additional processing in the form of another negation operator?

If *No X is too Y to Z* expresses a grammatically conventionalized meaning, word order may not affect the depth charge effect. However, if the construction is conventionalized to such a degree that it is idiomatic, changing the word order may nullify the depth charge effect. Furthermore, if the grammaticalized meaning has the same mental status as a compositionally-derived semantic representation, it should be possible to apply formal-semantic operators such as negation to it. By contrast, if the semantic representation is underspecified, an additional negation should be difficult to apply.

Experiments 1 and 2 ($n = 20$; $n = 62$) presented German participants with depth charge sentences and transparently non-sensible control sentences. Word order was manipulated between canonical and inverted, as in (1). Experiment 1 used whole-sentence presentation while Experiment 2 used self-paced reading. Participants also judged the sentences' sensibleness on a scale from 1 to 7 (= completely sensible). If the illusion occurs, depth charge sentences should be judged as sensible.

We report nested effects in the presence of reliable interactions. In Experiment 1, sensibleness ratings were higher for depth charge versus transparently non-sensible control sentences with canonical word order (95% CrI of rating difference: [0.98, 3.12]), but were not reliably higher with inverted word order (CrI: [-0.18, 0.89]). This pattern was replicated in Experiment 2 (canonical CrI: [1.12, 2.54]; inverted CrI: [-0.15, 1.2]). For the self-paced reading data, we aligned regions by content as opposed to linear position in the sentence (see Figure 1). When word order was canonical, log reading times were faster in the spillover region following the end of the *No X is too Y to Z* construction in depth charge compared to control sentences (back-transformed 95% CrI of difference: [-315 ms, -77 ms]). With inverted word order, there was instead some indication of a slowdown in depth charge sentences relative to control sentences (back-transformed CrI: [-13 ms, 100 ms]).

Experiment 3 ($n = 20$) compared depth charge sentences against control sentences with fewer negations and a sensible meaning, as shown in (2). Participants were asked to give binary judgments of the sentences' sensibleness. For half of the trials, subjects were instructed to interpret the sentence in the NEG-world, where people always say the opposite of what they mean (example given: *I like Brussels sprouts* means *I don't like Brussels sprouts*). The intention was to introduce an additional sentential negation operator into the semantics. The remaining trials were to be interpreted in the normal world. Results showed that the meaning of control sentences was reliably inverted when moving from the normal world (95% CrI of endorsements: [90%, 99%]) to the NEG-world (CrI: [6%, 43%]). By contrast, depth charge sentences were endorsed about equally often in the normal world (CrI: [30%, 60%]) and the NEG-world (CrI: [31%, 60%]), suggesting that their meaning is underspecified. Interestingly, however, a subset of depth charge sentences *do* appear to show systematic reversibility: These items were always judged as sensible in the normal world but as not sensible in the NEG-world (see Figure 2). A higher-powered re-run of the study is currently underway (planned $n = 100$).

Experiments 1 and 2 suggest that the depth charge effect occurs only in the canonical *No X is too Y to Z* configuration, where it manifested in the form of improved sensibleness ratings and a speedup in on-line processing. This may either signal a highly idiomatic construction, or indicate that something about the canonical configuration causes processing failure, such as the global negation appearing before the implicitly negative element *too* and/or the negated adjective. Experiment 3 suggests that depth charge sentences mostly have underspecified semantic representations, but there may be exceptions.

(1) EXPERIMENTS 1 & 2

Depth charge, canonical (compositionally non-sensible)

Keine Kopfverletzung ist zu ungefährlich, um ignoriert zu werden,...

No head injury is too trivial to be ignored

Depth charge, inverted (compositionally non-sensible)

Zu ungefährlich, um ignoriert zu werden, ist sicher keine Kopfverletzung, ...

too trivial to be ignored is surely no head injury

Control, canonical (compositionally non-sensible)

So manche Kopfverletzung ist zu ungefährlich, um ignoriert zu werden,...

Some a head injury is too trivial to be ignored

Control, inverted (compositionally non-sensible)

Zu ungefährlich, um ignoriert zu werden, ist sicher so manche Kopfverletzung, ...

too trivial to be ignored is surely some a head injury

(2) EXPERIMENT 3

Control, canonical (compositionally sensible)

Manch eine Kopfverletzung ist zu gefährlich, um ignoriert zu werden,...

Some a head injury is too dangerous to be ignored

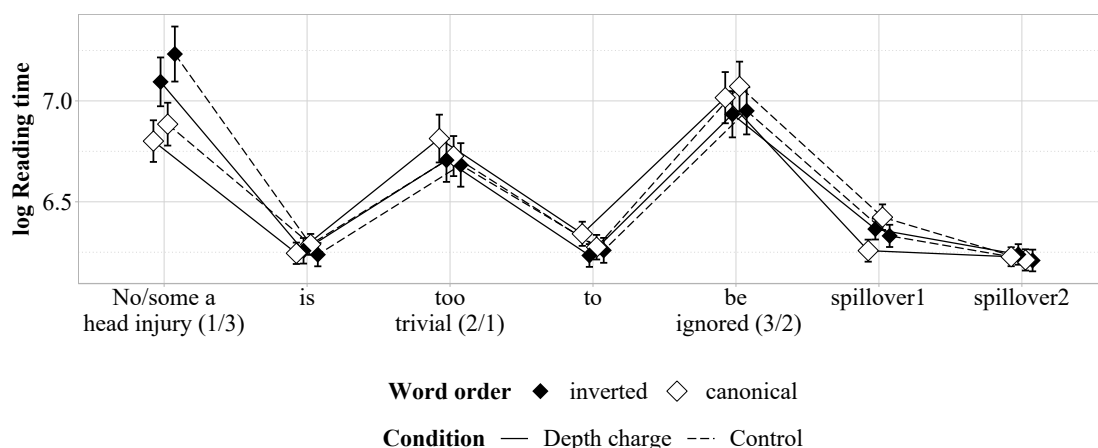


Figure 1: Self-paced reading results from Experiment 2.

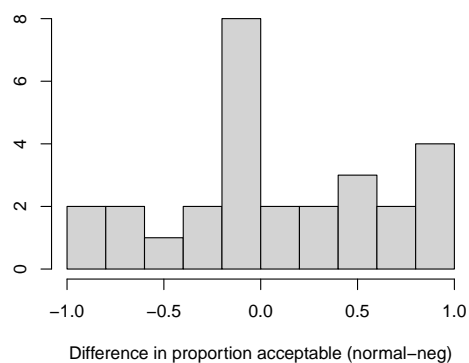


Figure 2: Consistency of meaning inversion by item in Experiment 3 (right = systematic).

References: [1] **Wason & Reich** (1979). The Quarterly Journal of Experimental Psychology, 31 (4), 591–597.
 [2] **Cook & Stevenson** (2010). Proceedings of the 2010 Workshop on NLP and Linguistics: Finding the Common Ground, 61–69. [3] **Fortuin** (2014). Cognitive Linguistics, 25 (2), 249–292.

A note on German word order: As observed by, e.g., **Müller** (Complex Predicates: Verbal Complexes, Resultative Constructions, and Particle Verbs in German, 2002; p. 71–72), German freely allows word order variation in copula constructions such as the ones used in the present studies (see also references therein):

- (3) Treu geblieben ist sich Dieter Kunzelmann also auf jeden Fall.
true stayed is self Dieter Kunzelmann so in any case

In Experiment 1, the canonical condition had a mean rating of 3.3, compared to 3.2 for the inverted condition. In Experiment 2, mean ratings were 3.6 for the canonical condition and 3.5 for the inverted condition. It is thus unlikely that the lack of a clear depth charge effect in inverted sentences is a floor effect.