Influences of word order and meaning reversal on the depth charge illusion

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 this meaning is not compositionally licensed. The non-compositional interpretation has been attributed to processing failure (Wason & Reich, 1979), but has also been argued to be due to a stored construction that licenses the non-compositional meaning (Cook & Stevenson, 2010; Fortuin, 2014).

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

1. Is the misinterpretation limited to the linear configuration *No X is too Y to Z* or does it also occur when the word order is changed (*Too Y to Z is surely no X*)?
2. Is the semantic representation of depth charge sentences detailed enough to be subjected to subsequent additional processing?

If the *No X is too Y to Z* construction is conventionalized to such a degree that it is idiomatic, changing the word order may nullify the depth charge effect. Furthermore, if the non-compositional meaning is grammaticalized and stable, it should be possible to productively apply negation to it. By contrast, if the interpretation is irregular and superficial, negation should be difficult or impossible to apply.

**Experiments 1 and 2** (*n* = 20; *n* = 60) presented German participants with depth charge sentences as well as non-sensible control sentences. We also manipulated the word order between canonical and inverted, as in (1). Experiment 1 used whole-sentence presentation while Experiment 2 was a self-paced reading study. Participants judged the sentences’ sensibleness on a scale from 1 to 7 (= completely sensible). Depth charge sentences but not control sentences should be judged as sensible when the illusion occurs.

Results from Experiment 1 showed that ratings for depth charge versus non-sensible control sentences improved more strongly with canonical word order (*Δ* = 1.88, CrI: [1.04, 2.83]) than with inverted word order, where there was no reliable evidence for the depth charge effect (*Δ* = 0.34, CrI: [−0.17, 0.81]). This pattern was replicated in the results of Experiment 2 (*Δ*<sub>can</sub> = 1.84, CrI: [1.17, 2.49]; *Δ*<sub>inv</sub> = 0.56, CrI: [−0.04, 1.2]). For the self-paced reading data, we aligned regions by content as opposed to linear position in the sentence (see Figure 1). In depth charge compared to control sentences, reading times were faster in the spillover region following the *No X is too Y to Z* sequence (*Δ* = −81 ms, CrI: [−130 ms, −32 ms]), suggesting superficial and/or non-compositional processing, but only when word order was canonical.

**Experiments 3 and 4** (*n* = 20; *n* = 89) 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. As a further manipulation, half of the trials stipulated that the utterance should be interpreted in the NEG-world, where people always say the opposite of what they mean (*I like Brussels sprouts* means *I don’t like Brussels sprouts*), while the other half was to be interpreted in the normal world. Results showed consistent inversion of meaning for control sentences, which received fewer positive acceptability judgments in the NEG-world compared to the normal world (*Δ*<sub>exp3</sub> = −67%, CrI: [−83%, −43%]; *Δ*<sub>exp4</sub> = −66%, CrI: [−75%, −56%]), but there was no indication of a difference for depth charge sentences, or even the opposite tendency (*Δ*<sub>exp3</sub> = 1%, CrI: [−26%, 27%]; *Δ*<sub>exp4</sub> = 12%, CrI: [−3%, 26%]).

Experiments 1 and 2 show that the depth charge effect is reduced when word order deviates from the canonical *No X is too Y to Z* configuration. This could either signal a highly idiomatic construction, or indicate that something about the canonical configuration in particular brings about processing failure, such as the global negation appearing before the implicitly negative element *too* and/or the negated adjective. Experiments 3 and 4 suggest that depth charge sentences mostly do not attain detailed semantic representations.
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

Experiments 3 & 4

**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.