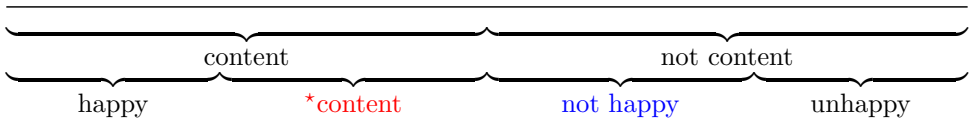
Scalar diversity and negative strengthening

Anton Benz, Carla Bombi Ferrer and Nicole Gotzner
Leibniz Centre General Linguistics

For more than a decade, scalar implicatures haven been a core topic of experimental pragmatics. However, it has been complained that psycholinguistic studies concentrate on a few scales only, most notably the scales ⟨*all, some*⟩ and ⟨*and, or*⟩. In [7] the authors provide an overview of 29 experimental studies from 2001 to 2014. Of them, only two studies consider scales other than ⟨*all, some*⟩ and ⟨*and, or*⟩. They speculate that the underlying reason for this bias is the belief that these scales are somehow representative for scales in general, such that findings on them can be generalised to all scales. This is the so–called uniformity hypothesis. This hypothesis has received some interest in recent years. The experimental studies in [2, 3, 7] addressed it in a special form: they tested the hypothesis that all scales show the same capacity for generating scalar implicature. This means, in this special form the hypothesis says that there is a constant percentage S such that for all scales i about $S\%$ of the subjects will draw an implicature for the weak scalar alternative. The most thorough and systematic study on this hypothesis is [7]. They tested 43 scales, among them 32 scales with adjectives, 6 with main verbs, 2 with auxiliary verbs, 2 with quantifiers, and 1 with adverbs. In their first experiment, they presented 25 subjects of an MTurk experiment with questions of the form: *John says: She is intelligent. Would you conclude from this that, according to John, she is not brilliant?* Subjects then had to choose between answers '*yes*' and '*no*'. The relevant scale is here ⟨*brilliant, intelligent*⟩. If subjects answer '*yes*', then they must have drawn the implicature '*intelligent*' +> '*not brilliant*'. The study showed that scales show considerable variance in their ability for generating scalar implicatures. In a post-analysis of their data, the authors of [7] found that boundedness of scales and perceived distance between strength of alternatives correlates with implicature rates, but none of the other parameters they considered. In particular, they briefly dismissed *negative strengthening* as a possible confounding parameter. In our talk, we present the results of a study based on [7] that shows that negative strengthening is (anti-)correlated with scalar implicatures, that a modified version of the uniformity hypothesis postulating a constant ratio between scalar implicature and negative strengthening can be maintained.

Negative strengthening [4, 6, 1, 5] is the phenomenon that the negation of the stronger scalar alternative is strengthened to an interpretation that also negates the weaker alternative. In (1) this is demonstrated for the scale ⟨*happy, content*⟩.

(1)



The second line shows the semantic extension of the adjective '*content*' and its negation, the third line the effect of scalar implicature and negative strengthening: the extension of '*content*' is shortened to '**content*', and that of '*not happy*' is strengthened such that it covers the area between '*content*' and '*not happy*'. Negative strengthening is variously explained as R-implicature [4], I–implicature [6], or blocking phenomenon [1, 5]. All authors agree that it arises differently from scalar implicature, which are special Q–implicature. The effect of negative strengthening in the experimental set up of [7] is the following: if subjects are asked *John says: He is content. Would you conclude from this that, according to John, he is not happy?*, then the '*no*'–answer may indicate a lack of scalar implicature, or the presence of negative strengthening. The scale ⟨*happy, content*⟩ shows particularly low levels

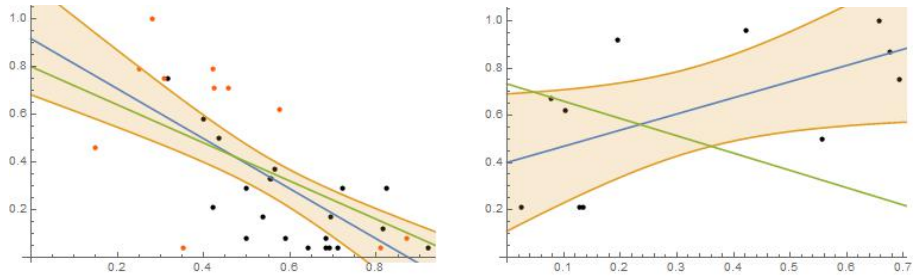


Table 1: Fit of modified uniformity hypothesis for M-scales (left) and L-scales (right)
orange: bounded scales, blue: unbounded scales, blue line: regression, green line: predicted by UH

of scalar implicature (4%). The authors dismiss this possibility, among other reasons, with the argument that their data show that scales with negative strong scale mates (henceforth *neg-scales*) generate high rates of implicature although these scales are known for showing a robust tendency towards negative strengthening. In order to test relevance of negative strengthening, we did an experiment with exactly the same items and fillers as [7] used in their experiment (Exp. 1), but modified the questions. For example, for ⟨*happy, content*⟩, we asked subjects *John says: He is not happy. Would you conclude from this that, according to John, he is not content?* If the answer '*yes*', this indicates that they negatively strengthened '*not happy*' to '*not content*'. On average, for all scales, 42.3% of the subjects answered '*yes*'. Selected results are shown in (2), with neg-scales on the right. Contrary to expectations, neg-scales are not particularly strong triggers of negative strengthening. Overall, we observe a correlation between $s_o(i)$, the observed % of SIs for scale i, and $(1 - n_o(i))$, with $n_o(i)$ the % of NS for i (Spearman's rank correlation: 0.463, $p < 0.002$).

(2) Results for selected scales: % of scalar implicature (SI) from [7], % of negative strengthening (NS) from our study

Scale	SI	NS
⟨*free, cheap*⟩:	100%	28%
⟨*all, some*⟩:	96%	42%
⟨*love, like*⟩:	50%	43%
⟨*finish, start*⟩:	21%	14%
⟨*exhausted, tired*⟩:	4%	69%
⟨*happy, content*⟩:	4%	92%

Scale	SI	NS
⟨*impossible, difficult*⟩:	79%	25%
⟨*none, few*⟩:	75%	31%
⟨*unsolvable, hard*⟩:	71%	43%
⟨*unavailable, scarce*⟩:	62%	58%
⟨*unforgettable, memorable*⟩:	50%	56%

The studies of [2, 3, 7] convincingly show that the uniformity hypothesis $s_0(i) = s$ for a constant s is false. However, we may ask whether there are other uniform constants that represent a similarity of behaviour for certain classes of scales. Given the correlation between $s_o(i)$ and $1 - n_o(i)$, the simplest reformulation of the uniformity hypothesis (UH) is to postulate a constant ratio between these values, i.e. there is a constant s such that for all i $s_o(i)/(1 - n_o(i)) = s$. The constant s can be fitted to the data. In Table 1, the results are shown for two classes of scales: M-scales for which the semantic extension of the scale mates covers a stretch from the middle to one extreme (e.g. ⟨*happy, content*⟩), and L-scales in which they cover the whole stretch except for the lower end (e.g. ⟨*all, some*⟩). As we can se, the two classes show a radically different behaviour. Whereas M-scales can be very well characterised by a fitted constant of $s = 0.80$ (SD 0.22), L-scales do not follow UH ($s = 0.73$, SD 0.45). We therefore propose the modified UH as a diagnostic that allows for more fine–grained differentiations between scale classes.

Conclusion. We show that a modified version of the uniformity hypothesis is consistent with the data presented by previous experimental studies. We present the first study that shows a numerical correlation between two different types of implicature: scalar implicature which are Q–based, and negative strengthening which is I– or M–based. The correlation between SI and NS may be sensitive to general scale structure. This shows that a more fine grained typology of scales can be motivated by numerical analysis.

References

[1] Blutner, R. (2004). Pragmatics and the lexicon. In Horn, L. and Ward, G., editors, *The Handbook of Pragmatics*, pages 488–514. Blackwell Publishing, Oxford.

[2] Doran, R., Baker, R. E., McNabb, Y., Larson, M., and Ward, G. (2009). On the non-unified nature of scalar implicature: an empirical investigation. *International Review of Pragmatics*, 1:1–38.

[3] Doran, R., Ward, G., McNabb, Y., Larson, M., and Baker, R. E. (2012). A novel paradigm for distinguishing between what is said and what is implicated. *Language*, 88:124–154.

[4] Horn, L. R. (1989). *A Natural History of Negation*. University of Chicago Press, Chicago.

[5] Krifka, M. (2007). Negated antonyms: Creating and filling the gap. In Sauerland, U. and Stateva, P., editors, *Presupposition and Implicature in Compositional Semantics*, pages 163–177. Palgrave Macmillan, Houndmill.

[6] Levinson, S. C. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicatures*. MIT Press, Cambridge, MA.

[7] van Tiel, B., van Miltenburg, E., Zevakhina, N., and Geurts, B. (2016). Scalar diversity. *Journal of Semantics*, 33(1):107–135.