Differences and similarities between scalar inferences and scalar modifiers

Speakers that utter a sentence like in (1) are usually understood to communicated that they didn't mean that all of Ann's friends shared her grim view of the world. In order to explain this scalar implicature, the scalar expression *a handful* is assumed to evoke a set of alternatives whose members are ordered in terms of informativeness along the lines of (2). A listener reasons that the speaker used the most informative expression, and so informationally-stronger alternatives like *most* or *all* don't hold (Horn 1972, Gazdar 1979, Levinson 2000, Geurts 2010, *inter alia*).

(1)   **A handful** of Ann's friends shared her view of the world as a dark place.

(2)   <some, a handful, many, most, all>

Many important questions have been addressed regarding scalar implicature, such as the processing of such inferences (e.g. Bott & Noveck 2004, Huang & Snedeker 2009) and whether they are part of a grammaticalized mechanism rather than pragmatic inferences (e.g. Chierchia 2006). What is striking about many of these studies is that the empirical investigation was confined mostly to *some* vs. *all*, neglecting the cross-categorial pervasiveness of the phenomena, as observed by, e.g., Horn (1972) and Hirschberg (1985). The few experimental studies that looked at a more diverse group of scalar implicatures have found that different scalar expressions differ in the extent to which they give rise to scalar implicature (Doran *et al*. 2009, 2012, van Tiel *et. al* ms.).

We present data from an experimental investigation of the extent to which implicature (in this case, upper-bound construal) arises in various scalar expressions, carried out in a new paradigm that differentiates between semantic upper-bound construal on the one hand and semantic blocking of the upper-bound construal with a constructions that are predicted to give rise to pragmatic scalar implicature. In order to do so, we draw upon the analyses of the modifiers *at least* and *at most*. Modification by *at least*, as in (3), is assumed to block the implicature, whereas *at most*, as in (4), would impose an upper-bound construal (Krifka 1999, Geurts & Nouwen 2007).

(3)   **At least handful** of Ann's friends shared her view of the world as a dark place.

(4)   **At most handful** of Ann's friends shared her view of the world as a dark place.

In addition to quantity expressions like *a handful*, other types of expressions have been examined, such as adjectives like *large*, ranked ordering like *lieutenant*, and non-scalar expressions like *saxophone* (6). 40 participants on Mechanical Turk rated pairs of claims and facts, as the one in (5) and were asked to rate their compatibility on a scale. Gradient data of acceptability judgements has been shown to help differentiate between semantic and pragmatic inferences (Cummin & Katsos 2010) or the likelihood of the latter (Chemla & Spector 2011). In addition to the three modification conditions (non-modification, *at least* and *at most*), there were also three discrepancy conditions, one in which the term given in the fact was contradictory to the one in the claim (e.g. *none*), entailed by it (*some*) or was informationally stronger than it (*all*).

(5)   **Claim:** $\left\{ \begin{array}{c} \emptyset \\ At\ least \\ At\ most \end{array} \right\}$ a handful of Ann's friends shared her view of the world as a dark place.

Fact: $\left\{ \begin{array}{c} None \\ Some \\ All \end{array} \right\}$ of Ann's friends shared her view of the world as a dark place.

How compatible is the claim with the fact?

-3   -2   -1   0   1   2   3

completely   completely
incompatible   compatible

Note that unlike the quantity expressions, adjectives, and non-scalar items, in which the stronger expression in the fact asymmetrically entails the one given in the claim, the stronger expression in ranked orderings (e.g. *general*) does not (*lieutenant*), and yet ranked-orderings give rise to scalar inferences (Hirschberg 1985).

Participants' ratings were analysed with mixed-effects ordered logit regression models, including slope and intercept random effects for subjects and items, respectively. In the strong discrepancy condition, the rates for *at least* were higher than those for the non-modification condition ($z=7.684, p<.0001$), and the rates for the non-modification condition were in turn higher than those for *at most*. In addition, the rates for the stronger condition were in aggregate lower in the weaker condition ($z=2.82, p<0.001$) and higher than those for the contradictory condition ($z=6.832, p<0.001$). This suggests that pragmatic inferences often, but not always, arise. Looking at individual expression types, the responses for quantity expressions in the *at most* and non modified conditions for the stronger condition weren't different ($z=-0.021, p=.0433$), whereas they were for the other expression types (all $p<.001$). This suggests that upper-bound construals are more likely to arise in quantity expressions than in adjectives, ranked orderings, or non-scalars. In the ad-hoc condition, the non-modification and *at least* conditions patterned together with respect to the weaker as well as the stronger condition, suggesting that even if participants exhaustify the expression at issue, thus ruling out any competing alternatives, as in *at most a saxophone* in (6), they clearly very readily suspend this exhaustification, and so a stronger expressions that includes more information is interpreted similarly to how it's interpreted when modified by *at least*, which ensures that no stronger alternatives are excluded.

(6)   **Claim:** The arrangement included a piano, a rhythm section, and $\left\{\begin{array}{c} \emptyset \\ at\,least \\ at\,most \end{array}\right\}$ a saxophone.

Fact: The arrangement included a piano, a rhythm section,
$\left\{\begin{array}{c} but\,no\,saxophone \\ and\,a\,wind\,instrument \\ a\,saxophone, a\,violin\,and\,a\,harp \end{array}\right\}.$

The results suggest a split between the quantity expressions and the other stimulus types with respect to the extent to which pragmatic upper-bound construal arises, whereby in the former the likelihood and strength of the implicature is greater than in the other expressions. The results from this study provide further support to findings from Doran *et al.* 2009, 2012 and van Tiel *et al.* ms., that not all scalar inferences are the same, and show that the comparison with *at most* is a successful and informative diagnosis in distinguishing between semantic and pragmatic upper-bound construal.

SELECTED REFERENCES: **Chierchia, Gennaro**. 2006. Broaden your views: Implicatures of domain widening and the "logicality" of language. **Doran et. al** 2009. On the non-unified nature of scalar implicature: an empirical investigation. **Geurts, Bart & Rick Nouwen**. 2007. At least et al. **Hirschberg, Julia**. 1985. A theory of scalar implicature. **Krifka, Manfred**. 1999. At least some determiners are not determiners. **van Tiel et. al**. ms. Scalar Diversity.