Title: Non-Doxastic Attitude Ascriptions and Semantic Meaning

Authors: Wojciech Rostworowski, Katarzyna Kuś, Bartosz Maćkiewicz (University of Warsaw)

Abstract: The aim of this talk is to provide new experimental evidence on (non-doxastic) attitude ascriptions and their entailment properties. We report two experiments using truth-value and acceptability judgement tasks, whose results suggest that the attitude verbs like 'want', 'fear' or 'glad' require a hyperintensional notion of meaning, including not only a truth conditional aspect but also the informational structure.

The problem under investigation emerges from the discussion on definite descriptions in attitude-verbs contexts. It has been observed that the statements ascribing a non-doxastic attitude to a subject (roughly, an attitude which does not involve *believing* in a proposition, e.g., 'hopes', or contains an extra component in addition to the belief, e.g., 'is glad') do not preserve their truth conditions once we substitute an embedded definite description with a corresponding 'there'-clause. For example, compare (1a) and (1b):

1. a. Hans wants the ghost in his attic to be quiet tonight.
    b. Hans wants *there to be* a (unique) ghost in his attic *and* for it to be quiet. (Elbourne 2010: 2)

Ascription (1a) seems to have different truth conditions than (1b) and it is possible for the latter to be intuitively false when the former is true (e.g., when Hans does not want to have any ghosts in his attic, but he actually believes that there is one and wants for that one ghost to be quiet). It is a matter of dispute whether there is a truth-conditional difference between the complement clauses in (1a) and (1b) (Russell 1905 vs Strawson 1950; for experimental findings see: Abrusán & Szendrői 2013, Schwarz 2016). A number of theorists (e.g., Heim 1991, Elbourne 2010, Schoubye 2013) has taken the contrast between (1a) and (1b) to be evidence against the Russellian interpretation of descriptions, thus explaining the contrast by positing a genuine semantic difference between the subclauses. One explanation appeals to the *presuppositional* status of definite descriptions – the subclause in (1a) presupposes the existence of a ghost while the subclause in (1b) does not, as the existence claim is a part of its assertoric content. However, further evidence suggests that contrastive ascriptions like (1) do not have to feature definites, but also various types of indefinite expressions (Schoubye 2013), and do not necessarily involve presuppositional differences (Blumberg 2017, Rostworowski 2018). For instance, (2a) is different from (2b):

2. a. Anne wonders whether the dictator has been assassinated.
    b. Anne wonders whether the dictator is dead and has been assassinated.

The subclauses in (2a) and (2b) do not, however, differ in terms of their presuppositions. In particular, 'being dead' is not presupposed by 'being assassinated', as it does exhibit typical projection behavior; instead, it seems to be an ordinary *entailment* (Rostworowski 2018: 1317-1323). Altogether, the theoretical literature indicates that the problem of substitutions in the scope of non-doxastic attitude verbs is more general and concerns the nature of these verbs rather than the issue of proper treatment of definite descriptions/presuppositions.

The aim of our first experiment (Study I) was to investigate to what extent the predictions of theoreticians about the contrast between a-type and b-type ascriptions are confirmed by evaluations of ordinary language speakers. Study I employed 2 (type of ascription: a-type vs b-type) x 2 (task: acceptability vs true value judgment) x 4 (non-doxastic attitude: fear, want, feel sorry, glad) mixed design. The first two factors were between-subject manipulation, the last one was within-subject. The study participants were presented with a set of simple contexts where a protagonist could be naturally ascribed a (non-doxastic) attitude of a-type, but not b-type. After each context, the participants were asked to evaluate an ascription of a given attitude (a-type or b-type ascription, depending on the experimental condition), that is, to say whether the ascription is true/acceptable in the context. Our informants were also requested to indicate how confident they are in their judgments. The main finding is a statistically significant effect of the type of ascription, with a-type rated much higher than b-type (p < 0.001).

The results of Study I indicate that there is a difference in both acceptability and truth conditions between a-type and b-type ascriptions, i.e., the former are more acceptable/regarded as 'true' in the contexts investigated in the study. This is in line with theoretical predictions. However, it is interesting that b-ascriptions are not fully rejected/regarded as 'false' in those contexts.

The aim of the second experiment (Study II) was to further explore the problem by investigating the nature of the discrepancy between a-type and b-type ascriptions. Roughly, there are two possible routes for the explanation to go: (i) we have a genuine semantic non-equivalence between the subclauses in a-type and b-type ascriptions, which goes beyond the mere presuppositional differences, and consequently generates non-equivalent readings of the ascriptions; (ii) a-type and b-type ascriptions are different for pragmatic reasons, in particular, in the contexts under investigation – where the a-type formulation is perfectly acceptable – b-type is 'infelicitous' as it violates the principle of 'contextual redundancy' (for details, see Blumberg 2017; cf. Fox 2008). The two approaches (i) and (ii) have different predictions about the status of Conjunction Elimination "under" attitude verbs (i.e., an inference to e.g. 'S wants p' based on that S wants p and q). According to (i), it may be a true semantic entailment (as the two ascriptions have different sets of entailments); for (ii) it must be derivable on pragmatic basis, i.e., it is akin to a conversational implicature. In Study (II), we test this particular prediction by appealing to 'cancelability' (Grice 1989), that is, we check whether our informants find it coherent to ascribe an attitude towards a conjunctive proposition to a person and to deny that the person has the attitude towards the conjuncts in isolation.

Study II employed a within-subject design (non-doxastic attitude vs semantic entailment vs implicature). In this study, the participants were presented with two-sentence discourses. The first sentence attributed a non-doxastic attitude to a subject (e.g., 'Anne feels sorry that she went to the forest and found no mushrooms'). The second sentence denied that the subject had the attitude towards a single conjunct alone (e.g., 'In fact, she doesn't feel sorry about being in the forest'). The participants were asked whether such discourses were coherent on the 7-point pseudo-Likert scale. The discourses with attitude ascriptions were contrasted with discourses with canceled implicatures on the one hand, and with canceled semantic entailments, on the other hand. The main finding of Study II is that the discourses with attitude ascriptions were judged as generally incoherent – similarly to the discourses with cancelled semantic entailments (the ratings significantly below the midpoint, $p < 0.001$) – and much different from those with canceled implicatures, which were deemed to be coherent (although weakly; the ratings significantly above midpoint, $p < 0.01$).

The results of Study II confirmed the prediction that Conjunction Elimination is supported by the considered attitude verbs and that the inference is semantic rather than pragmatic. This is a significant empirical result in light of the observation that non-doxastic attitude verbs do not generally support entailments (e.g., Asher 1987, Kaplan 2005: 985). More importantly, the results suggest that there is a genuine semantic difference between ascriptions like (2a)/(2b) and hence attitude verbs operate on the semantic content of the complement clauses taken to include not only truth conditions but also the information structure that goes beyond presuppositions.

References:
Abrusán, M. & Szendrői, K. 2013. Experimenting with the king of France: topics, verifiability, and definite descriptions. *Semantics & Pragmatics* 6, 1-43.
Asher, N. 1987. A typology for attitude verbs and their anaphoric properties. *Linguistics and Philosophy* 10, 125-198.
Blumberg, K. 2017. Ignorance implicatures and aon-doxastic attitude verbs. In A. Cremers, T. van Gessel, F. Roelofsen (eds.) *Proceedings of the 21st Amsterdam Colloquium*,135-144.
Elbourne, P. 2010. The existence entailments of definite descriptions. *Linguistics and Philosophy* 33(1), 1-10.
Fox, D. 2008. Two short notes on Schlenker's theory of presupposition projection. *Theoretical Linguistics*, 34(3), 237-252.
Grice, P. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
Heim, I. 1991. Artikel und Definitheit. In A. von Stechow, D. Wunderlich (eds.) *Semantik: ein internationales Handbuch der zeitgeňossischen Forschung*, 487–535. Berlin: Walter de Gruyter.
Kaplan, D. 2005. Reading 'On Denoting' on its centenary. *Mind* 114, 933–1003.
Rostworowski W. 2018. Descriptions and non-doxastic attitude ascriptions. *Philosophical Studies* 175(6), 1311-1331.
Russell, B. 1905. On denoting. *Mind*, 14, 479–493.
Schwarz, F. 2016. False but slow: evaluating statements with non-referring definites. *Journal of Semantics*, 33, 177-214.
Strawson, P. 1950. On referring. *Mind* 59(235) 320-344.

Schoubye, A. 2013. Ghosts, murderers, and the semantics of descriptions. *Noûs*, 47(3), 496-533.