

Desires, belief revisions, and polarity sensitivity

Julie Goncharov, University of Tromsø

This talk aims at providing a deeper understanding of the semantic/pragmatic contribution of desire attitudes like *want*. I take as a starting point the mysterious stipulation in Heim 1992 that in cases like (1), where intuitively John believes that he will go to the movies tonight, *want* is interpreted relative to a superset of John’s beliefs, which includes both going-to-the-movies worlds as well as not-going-to-the-movies worlds. I dub this stipulation *No Belief Revision*.

- (1) (John hired a babysitter because) he wants to go to the movies tonight.

In Heim’s system, *No Belief Revision* is needed to ensure that the presupposition of $want_x \phi$ that x does not believe ϕ nor $\neg\phi$ is satisfied. This latter presupposition, in turn, guarantees that the context update semantics of *want* avoids (by presupposition failure) Stalnaker’s fallacies such as *I want to get well* entailing *I want to have been sick*. **The main claim of this talk is that *No Belief Revision* doesn’t have to be stipulated.** On the contrary, I propose that belief revisions are obligatory and the lack of the revision in (1) is the result of an unsuccessful revision. This proposal will allow us to explain a number of grammatical phenomena that otherwise remain unaccounted for.

Setting up the scene The interaction between desire predicates and polarity sensitive items (PSIs) has been widely discussed in semantic literature (von Stechow 1999, Crnič 2011, Romoli 2012, a.o.). However, the phenomena I focus on is mostly disregarded (except for Szabolcsi 2004, 2010). The core observation is that (anti-)licensing of PSIs is affected by the interpretation of the embedded predicate. Weak Negative Polarity Items (NPIs) like *any* are insensitive to the interpretation of the verb, (2). Strong NPIs like *until* are deviant when the action expressed by the verb is accidental (in the sense specified below), (3). Positive Polarity Items (PPIs) like *some* show a mirror image, i.e. they are deviant when the action expressed by the verb is interpreted as controlled/non-accidental, (4).

- (2) a. I don’t want to call anyone/eat anything.
b. I don’t want to offend anyone/break anything.
- (3) a. I don’t want to leave until 10pm.
b. I don’t want to leave at the wrong moment ??until/✓before 10pm.
- (4) a. ??I don’t want to call someone/eat something. (where neg>some)
b. I don’t want to offend someone/break something. (where neg>some)

By the controlled/accidental distinction, I mean the following: an action is interpreted as controlled if the agent x of the action believes that if she acts so as to bring about ϕ , the state of affairs described by ϕ obtains and similarly for $\neg\phi$, (5a). An action is interpreted as accidental if it is not controlled, (5b), see also Pritchard 2005, 2016, Horst 2015. Accidental actions can be intentional or non-intentional. What distinguishes them from controlled actions is that the agent does not believe that if she performs (abstains from) the action, the result is guaranteed. For example, even if I intentionally try to offend Mary, she might not get offended, but if I try to call her, she will get called. We can view the conditions in (5) as conventional implicatures of corresponding types of verbs.

- (5) a. Controlled actions: $B_x(\psi \rightarrow \phi) \wedge B_x(\neg\psi \rightarrow \neg\phi)$ ($\psi = x$ acts so as to bring about ϕ)
b. Accidental actions: $\neg B_x(\psi \rightarrow \phi) \vee \neg B_x(\neg\psi \rightarrow \neg\phi)$

I put forward three linguistic tests that help to tell apart verbs that (stereotypically) express controlled vs. accidental actions, (6)-(8), and show how these contrasts can be explained by (5).

- (6) a. # I didn’t want to call Mary, but I did. (unless weakness of the will or accidentality)
b. I didn’t want to offend Mary, but I did. (accidental)
- (7) a. # I tried to call Mary, but she didn’t get called.
b. I tried to offend Mary, but she didn’t get offended.

- (8) a. # John was going to call Mary, but he didn't (actually) expect that he'd call her.
 b. John was going to offend M., but he didn't (actually) expect that he'd offend her.

About the presupposition of *want* Romoli 2012 argues that $want_x \phi$ has a weaker presupposition than the one proposed in Heim 1992, (9). His motivation is to make licensing of strong NPIs under *want*, as in *John doesn't want Mary to leave until Friday*, compatible with recent accounts which claim that strong NPIs are licensed when the conjunction of assertion and presuppositions is downward entailing (Gajewski 2011, Chierchia 2013). This licensing condition holds for (9b), but not (9a). In what follows, I assume that *want* has Romoli's presupposition in (9b).

- (9) a. Heim's presupposition of $want_x \phi : \hat{B}_x \phi \wedge \hat{B}_x \neg \phi$, where $\hat{B}\phi = \neg B\neg\phi$
 b. Romoli's presupposition of $want_x \phi : \hat{B}_x \phi \rightarrow \hat{B}_x \neg \phi$

Belief revisions I propose that belief revisions are obligatory and they follow the rules in (10):

- (10) a. $\mathcal{K} \otimes \phi \equiv (\mathcal{K} \ominus \neg \phi) \oplus \phi$ 'in order to revise a belief set \mathcal{K} with (possibly contradictory) ϕ , we first contract \mathcal{K} with $\neg \phi$ and then expand the resulting set with ϕ ' (Levi-identity)
 b. $\mathcal{K} \otimes \perp = \mathcal{K}$ 'do not revise with contradictions!' (unsuccessful revision)

To assess the doxastic modal base of $want_x \phi$ we start with $\mathcal{K} = \{\hat{B}_x \phi \wedge \hat{B}_x \neg \phi\}$ (x believes that both ϕ and $\neg \phi$ are possible) and revise \mathcal{K} with the conjunction of the following statements: a : the presupposition of *want* in (9b), projected under negation (Karttunen 1973), b : one of the conventional implicatures in (5) depending on the action, c : x 's beliefs about how she is going to act: $B_x \psi$ or $B_x \neg \psi$ ($\psi(\neg \psi) = x$ acts so as to bring about $\phi(\neg \phi)$). It is natural to assume that when x asserts $want_x \phi$, $B_x \psi$ holds, i.e. x believes that she will act so as to bring about ϕ . Similarly, when x asserts $\neg want_x \phi$, $B_x \neg \psi$ holds, i.e. x believes that she will act so as to bring about $\neg \phi$. (see e.g. Condoravdi and Lauer 2016 on effective preferences). The four possible revisions are as follows:

Rev.#1 *want* $\phi^{controlled}$

$\mathcal{K} : \hat{B}\phi \wedge \hat{B}\neg\phi$
 $a : \hat{B}\phi \rightarrow \hat{B}\neg\phi$
 $b : B(\psi \rightarrow \phi) \wedge B(\neg\psi \rightarrow \neg\phi)$
 $c : B\psi$
 $a \wedge b \wedge c = \perp; \mathcal{K} \otimes abc = \mathcal{K}$

Rev.#2 *want* $\phi^{accidental}$

$\mathcal{K} : \hat{B}\phi \wedge \hat{B}\neg\phi$
 $a : \hat{B}\phi \rightarrow \hat{B}\neg\phi$
 $b : \neg B(\psi \rightarrow \phi) \vee \neg B(\neg\psi \rightarrow \neg\phi)$
 $c : B\psi$
 $a \wedge b \wedge c = consist.; \mathcal{K} \otimes abc = \mathcal{K} \oplus B\psi$

Rev.#3 $\neg want$ $\phi^{controlled}$

$\mathcal{K} : \hat{B}\phi \wedge \hat{B}\neg\phi$
 $a : \hat{B}\phi \rightarrow \hat{B}\neg\phi$
 $b : B(\psi \rightarrow \phi) \wedge B(\neg\psi \rightarrow \neg\phi)$
 $c : B\neg\psi$
 $a \wedge b \wedge c = consist.; \mathcal{K} \otimes abc = B\neg\psi$

Rev.#4 $\neg want$ $\phi^{accidental}$

$\mathcal{K} : \hat{B}\phi \wedge \hat{B}\neg\phi$
 $a : \hat{B}\phi \rightarrow \hat{B}\neg\phi$
 $b : \neg B(\psi \rightarrow \phi) \vee \neg B(\neg\psi \rightarrow \neg\phi)$
 $c : B\neg\psi$
 $a \wedge b \wedge c = consist.; \mathcal{K} \otimes abc = \mathcal{K} \oplus B\neg\psi$

Results First note that Rev.#1 is an unsuccessful revision by (10b) since $b \wedge c \rightarrow B\phi$ (by Distribution: $(Bp \wedge B(p \rightarrow q)) \rightarrow Bq$) and $B\phi \wedge a = \perp$. Thus, we derive Heim's *No Belief Revision* stipulation for cases like (1). Rev.#2 describes cases in which the outcome of the action is not (fully) controlled by the agent. For example, *John wants to offend Mary* is interpreted against John's belief set such that he believes that he will act so as to offend Mary and he also believes that he might not succeed in it, which is a welcome result. More interesting are negated desires in Rev.#3 and Rev.#4. Observe that Rev.#3 results in strengthening, i.e. in all x 's belief-worlds $\neg \phi$ is true, whereas Rev.#4 keeps ϕ -worlds among x 's beliefs. I claim that this difference is responsible for the contrasts in (3)-(4). Strong NPIs require assertions as well as presuppositions to be downward entailing, therefore they are fully acceptable with controlled actions (3a) (Rev.#3 i.e. $B\neg\phi$), but become deviant with accidental actions (3b) (Rev.#4 where $\hat{B}\phi$ holds). The contrast in (4) can be accounted for if we assume that *some* under negation can be rescued if the presupposition creates an upward entailing environment (intervention effect), which is satisfied in Rev.#4, but not Rev.#3. Finally, as general consensus has it, weak NPIs are sensitive only to the assertive content, thus, no effect is seen in (2).

References

- Chierchia, Gennaro. 2013. *Logic in grammar: Polarity, free choice, and intervention*. Oxford: Oxford University Press.
- Condoravdi, Cleo, and Sven Lauer. 2016. Anankastic conditionals are just conditionals. *Semantics and Pragmatics* 9:1–61.
- von Stechow, Kai. 1999. NPI Licensing, Strawson Entailment, and Context Dependency. *Journal of Semantics* 16:97–148.
- Gajewski, Jon. 2011. Licensing strong NPIs. *Natural Language Semantics* 19:109–148.
- Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9:183–221.
- Horst, David. 2015. Actions and accidents. *Canadian Journal of Philosophy* 45:300–325.
- Karttunen, Lauri. 1973. Presuppositions of compound sentences. *Linguistic Inquiry* 4:169–193.
- Pritchard, Duncan. 2005. *Epistemic luck*. Oxford: Oxford University Press.
- Pritchard, Duncan. 2016. Safety-based epistemology: whither now? Ms. University of Edinburgh.
- Romoli, Jacopo. 2012. Soft but strong. neg-raising, soft triggers, and exhaustification. Doctoral Dissertation, Harvard University.
- Szabolcsi, Anna. 2004. Positive polarity – negative polarity. *Natural Language and Linguistic Theory* 22:409–452.
- Szabolcsi, Anna. 2010. Infinitives vs. subjunctives: What do we learn from obviation and from exemptions from obviation? URL <http://goo.gl/c16gIP> ms. New York University.