Julie Goncharov, University of Tromsø

Language and intentions

Language, as a cognitive system, interacts with other cognitive modules. In this paper, we model the interaction between language and a cognitive system responsible for intentions and decision making. As a study case, we investigate constructions with desire predicates like *want* and Polarity Sensitive Items (PSIs). We present the novel data that, we argue, can be accounted for if we (a) distinguish between *controlled* and *non-controlled* actions and (b) add the mechanism of belief revisions that can shape presuppositions.

 Setting up the scene The novel observation about PSIs is that their (anti-)licensing is affected by the interpretation of the embedded predicate (see also Szabolcsi 2004). Although weak Negative Polarity Items (NPIs) like *any* are insensitive to the interpretation of the verb, (1), strong NPIs like the punctual *until* are deviant when the action is not controlled by the agent (as specified below), (2). Positive Polarity Items (PPIs) like *some* show the mirror image, i.e. they are deviant when the action is controlled by the agent, (3). We show that the data in (1)-(3) are supported by a likert-scale felicity judgement experiment on the Amazon Mechanical Turk.

(1)   a.   I don't want to call anyone/eat anything.   (controlled)

       b.   I don't want to offend anyone/break anything.   (non-controlled)

(2)   a.   I don't want to leave until 10pm.   (controlled)

       b. ?? I don't want to leave at the wrong moment until 10pm.   (non-controlled)

(3)   a. ?? I don't want to call someone/eat something.   (neg > some, controlled)

       b.   I don't want to offend someone/break something.   (neg > some, non-controlled)

 (a) Controlled/non-controlled distinction We say that an action is interpreted as *controlled* when 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 $\neg\phi$, (4a). An action is interpreted as *non-controlled* when the negation of (4a) holds, (4b). **Notations used in this abstract:** $\phi =$ the proposition that describes the action, e.g. '$\exists x[\text{call}(\text{speaker,x})]$' in (1a); $\psi = $ 'the agent acts so as to bring about ϕ'; $K =$ a belief set, i.e. a set of propositions; if A is a consistent set, $[A] = \{w \in W \mid A \subseteq w\}$, otherwise $[A] = \emptyset$; $[p] = \{w \in W \mid p \in w\}$.

(4)   a.   Controlled actions: $[K] \subseteq [(\psi \to \phi) \wedge (\neg\psi \to \neg\phi)]$

       b.   Non-controlled actions: $[K] \cap [\neg((\psi \to \phi) \wedge (\neg\psi \to \neg\phi))] \neq \emptyset$

The condition in (4a) comes from two independent sources: i) the intentionality condition in philosophy, e.g. Pritchard 2016, Horst 2015, and ii) the so-called 'matching rule' in psychology, e.g. Shultz and Wells 1985, Miller and Aloise 1989. We argue that (4b) can be regarded as a conversational (scalar) implicature (Egré 2014). Circumstantial evidence for this comes from the fact that pre-school children over-attribute controllability. They acquire the non-controlled interpretation later on, at the same age as scalar implicatures like *some, but not all*, e.g. Guasti et al. 2005. Both controlled and non-controlled actions can be intentional or accidental (in the everyday sense). Only intentional controlled actions must satisfy (4a). Accidental controlled actions pattern together with non-controlled (intentional or accidental) actions, (4b). This classification correctly predicts that adding *accidentally/by mistake* to e.g. (3a) or interpreting *call* in the context of pocket dialing allows us to interpret *some* under negation.

 (b) Belief revisions and the presupposition of *want* To account for (1)-(3), we propose that the modal base of *want*, which is traditionally taken to be a belief state such that the attitude holder believes neither the prejacent nor its negation (Heim 1992, von Fintel 1999), is formed by a belief revision mechanism. This mechanism generates different presuppositions with controlled vs. non-controlled actions, which accounts for the contrasts in (1)-(3).

 To introduce belief revisions, we use a possible world model for AGM postulates (Grove 1988). We propose that the doxastic modal base of *want* is restricted by a selection function $\gamma : [K_0] \to [K_1]$ that takes a neutral belief state $[K_0]$ ($[K]$ is a neutral belief state iff $[K] \cap [p] \neq \emptyset$ and $[K] \cap [\neg p] \neq \emptyset$ for any relevant p) and returns a new (smallest) state $[K_1]$ updated with (i)-(iii) if (i)-(iii) are compatible (**Expand if compatible!**). Otherwise, γ returns the neutral state $[K_0]$ (**Do not revise with contradictions!**).

(i) *dec(ision)*, i.e. x's beliefs about how she is going to act:
  − $[K] \subseteq [\psi]$ when '$want_x\ \phi$' is asserted ($\approx$ 'effective preferences' in
  − $[K] \subseteq [\neg\psi]$ when '$\neg want_x\ \phi$' is asserted Condoravdi and Lauer 2016)
(ii) controlled/non-controlled inference in (4)
(iii) Romoli's (2012) presupposition of *want*: if $[K] \cap [\phi] \neq \emptyset$ then $[K] \cap [\neg\phi] \neq \emptyset$

Sentences with *want* can have four possible revisions:

Revision 1: want $\phi^{controlled}$
(i) $[K] \subseteq [\psi]$ *dec*
(ii) $[K] \subseteq [(\psi \to \phi) \wedge (\neg\psi \to \neg\phi)]$ (4a)
(iii) if $[K] \cap [\phi] \neq \emptyset$ then $[K] \cap [\neg\phi] \neq \emptyset$ psp

> $\gamma([K_0]) = [K_0]$ ((i)-(iii) are incompatible)
> **Do not revise with contradictions!**

Revision 2: want $\phi^{non-controlled}$
(i) $[K] \subseteq [\psi]$ *dec*
(ii) $[K] \cap [\neg((\psi \to \phi) \wedge (\neg\psi \to \neg\phi))] \neq \emptyset$ (4b)
(iii) if $[K] \cap [\phi] \neq \emptyset$ then $[K] \cap [\neg\phi] \neq \emptyset$ psp

> $\gamma([K_0]) = [K_1]$ s.t. $[K_1] \cap [\phi] \neq \emptyset$, $[K_1] \cap [\neg\phi] \neq \emptyset$, and $[K_1] \subseteq [\psi]$ **Expand!**

Revision 3: $\neg$ want $\phi^{controlled}$
(i) $[K] \subseteq [\neg\psi]$ *dec*
(ii) $[K] \subseteq [(\psi \to \phi) \wedge (\neg\psi \to \neg\phi)]$ (4a)
(iii) if $[K] \cap [\phi] \neq \emptyset$ then $[K] \cap [\neg\phi] \neq \emptyset$ psp

> $\gamma([K_0]) = [K_1]$ s.t. $[K_1] \subseteq [\neg\phi \wedge \neg\psi]$
> (strengthening)   **Expand!**

Revision 4: $\neg$ want $\phi^{non-controlled}$
(i) $[K] \subseteq [\neg\psi]$ *dec*
(ii) $[K] \cap [\neg((\psi \to \phi) \wedge (\neg\psi \to \neg\phi))] \neq \emptyset$ (4b)
(iii) if $[K] \cap [\phi] \neq \emptyset$ then $[K] \cap [\neg\phi] \neq \emptyset$ psp

> $\gamma([K_0]) = [K_1]$ s.t. $[K_1] \cap [\phi] \neq \emptyset$, $[K_1] \cap [\neg\phi] \neq \emptyset$ and $[K_1] \subseteq [\neg\psi]$ **Expand!**

Revision 1 is an unsuccessful revision. This is because (i)-(iii) are incompatible (we assume that a belief set is closed under modus ponens). By **Do not revise with contradictions!**, the revised belief state is the same as the neutral belief state in which x does not believe ϕ nor $\neg\phi$. We claim that this unsuccessful revision derives Heim's stipulation that in cases like *John hired a babysitter because he wants to go to the movies tonight*, the modal base of *want* is a superset of John's beliefs that does not take into account his decision about how to act (Dox*). When the action is non-controlled (Revisions 2 and 4), the result of the revision is expansion by ψ and $\neg\phi$ respectively. Hence, for (1b), (2b), and (3b), we have the felicity condition in (5b). However, when the action is controlled (Revision 3), the presupposition of *want* is strengthened. That is, for (1a), (2a), and (3a), we have the felicity condition in (5a).

(5) a. $\neg want_x\ \phi^{controlled}$ is felicitous iff $[K_1] \subseteq [\neg\phi \wedge \neg\psi]$ (strengthening)
  b. $\neg want_x\ \phi^{non-controlled}$ is felicitous iff $[K_1] \cap [\phi] \neq \emptyset$, $[K_1] \cap [\neg\phi] \neq \emptyset$, and $[K_1] \subseteq [\neg\psi]]$

<u>Accounting for (1)-(3)</u> Following Gajewski 2011 and Chierchia 2013, we assume that strong NPIs require assertions as well as presuppositions to be downward entailing. Therefore, strong NPIs in (2) are fully acceptable with controlled actions, (5a), but become degraded with non-controlled actions, (5b). The contrast in (3) 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 happens with non-controlled actions, (5b). Independent evidence for this assumption comes from the fact that *some* is not anti-licensed under *few, rarely,* and *at most*, i.e. the environments that are known to be presuppositional interveners for strong NPIs. Finally, as the general consensus has it, weak NPIs are sensitive only to the assertive content, thus, no effect is seen in (1).

<u>Extension</u> We show that the contrast in (3) is also found in Hungarian, Polish, Russian, and Hebrew (Szabolcsi 2010). We also discuss evidence that the controlled/non-controlled distinction is a promising candidate for a linguistic universal as it surfaces in different grammatical constructions across unrelated languages. For instance, in case assignment in Hindi/Urdu (Tuite et al. 1985) and Central Pomo ((Mithun 1991), aspect choice in Slavic (Forsyth 1970, Paducheva 2013), disjoint reference effect in Romance (Ruwet 1991, Costantini 2011), and verbal inflection in Newari (Zu 2018). We show that some of these phenomena can be explained by the same mechanism that we proposed here for PSIs under a negated desire predicate.