Although generics are studied mostly in formal semantics, recently they attracted the attention of cognitive psychologists and social philosophers as well. The reason is that generics play a core role in the way we *learn*, *represent* and *reason* about groups in the world (cf. Leslie, 2008). One of the important— and worrying (Hasslanger 2011)— findings in these areas is (i) that we *accept* generic statements of the form 'Gs are P' typically if relatively many Gs have property P, but (ii) that if we are unfamiliar with a group and we learn a generic statement about it, we still *interpret* it in a much stronger way: (almost) all Gs are P (Cimpian et al 2010). This paper provides an explanation of (i) and (ii) based on a new interpretation of Tverky & Kahneman's (1974) 'Heuristics and Biases'-program by analyzing *typicality* in terms of *contingency*.

Typicality is well-studied in cognitive psychology. Groups (or categories) are represented by typical, rather than by all and only all, members, or by typical features, rather than by necessary and sufficient features, because agents have limited attention and limited recall of examples. But what are a group's typical members or features? According to Rosch's (1973) *prototype* theory, it are the *average* members of the group, or the features *most* members have. Barsalou (1985) experimentally showed, however, that at least for goal-derived categories, the typical members are instead the category's *ideal* members; those that *best* satisfy the goal. More recent empirical findings (e.g. Ameel & Storms, 2006) show that extreme members of a group are also considered typical for many, if not most, other types of categories, namely if categorization is performed in a *contrastive* way. Typical members of a category have features that *distinguish* them from members of other categories; as such they highlight, or exaggerate, real differences between groups.[1]

Typical members have features that are *representative* for the group. On our *preliminary* proposal, we measure the 'representativines' of O for C in terms of what in Conditioning Psychology (Rescorla, 1968) is called the **contingency** between C and O. This contingency, or strength of association, between cue C (e.g. tone) and outcome O (e.g. shocks) is measured by $P(O/C) - P(O/\neg C)$, abbreviated by ΔP^O_C. On our preliminary proposal, a familiar feature f is representative for a rather unfamiliar group G iff there is no feature h with a significantly higher contingency with G than f, i.e., if $\neg\exists h \in Alt(f) : \Delta P^h_G \gg \Delta P^f_G$. Notice that a representative feature for group G doesn't have to be one that most, or even many, members of the group have. Instead, a representative feature is one that *distinguishes* group G from its alternative(s) (In the definition $\neg G$ is an abbreviation for $\bigcup Alt(G)$, where $G \notin Alt(G)$), which is exactly in line with the view on typicality discussed above.

In fear-conditioning experiments with animals it is found that the association between cue (e.g. tone) and outcome (e.g. shock) increases with the intensity of the outcome (i.e, the shock). Similarly, Slobic et al (2004) et al. found that events with high emotional impact are better remembered and give rise to stronger associations. To incorporate the insight of Slobic et al and of fear-conditioning, we will measure the **representativeness** of f for G by ∇P^f_G, where the latter is defined as follows: $\nabla P^f_G := \Delta P^f_G \times Value(f)$. $Value(f)$ measures the *intensitity*/importance of feature f.

Shanks (1995) shows that ΔP^f_G is the asymptotic value of the weight given to f when the learning task is modeled using the Rescorla-Wagner learning rule (Rescorla & Wagner, 1972), the most influential learning rule in associative learning which is equivalent to the delta rule used in connectionists models. Interestingly enough, Shanks also points out that associative models using the delta rule learning algorithm can immediately explain the finding that associative learning is influenced by the magnitude or *Value* of the outcome, motivating not only ΔP^f_G, but also ∇P^f_G.

[1] Important is that if a number of contrastive categories are represented by such extreme members, the extensions of the categories can still be recovered, just like when they are represented by prototypes.

Semantic claim: a generic sentence like 'Gs are f' is true iff $\neg\exists h \in Alt(f) : \nabla P^h_G \gg \nabla P^f_G$ and $\exists h \in Alt(f) : \nabla P^f_G > \nabla P^h_G$, i.e., if f is (among the relative alternative features $Alt(f)$) among the most representative feature of G.

This analysis makes the truth of a generic context-dependent. It depends on $Alt(f)$, on $Alt(G)$, and on $Value(f)$. I claim that this context-dependence allows us to account for various types of generics:

1. If $Alt(f) = \{f, \neg f\}$ and $Value(f) = Value(\neg f)$, the generic 'Gs are f' is true just in case $\Delta P^f_G > 0$, i.e., $P(f/G) > P(f/\neg G)$, i.e., Cohen's *relative reading* ('Dutchmen are good sailors')
2. If $\forall h \in Alt(f) : Value(h) = Value(f)$ and $P(f/G)$ is not high, 'Gs are f' is true just in case $P(f/\neg G)$ is very low, and thus f is very *distinctive* for Gs. ('Ducks lay eggs')
3. If $\forall h \in Alt(f) : Value(h) = Value(f)$ and $P(h) \approx P(f)$, 'Gs are f' is true only if $\forall h \in Alt(f) : P(f/G) \gg P(h/G)$, or if ΔP^f_G is only somewhat above 0 and $\forall h \in Alt(f)$, $P(h)$ is not low, $P(f/G)$ has to be (very) high ('*standard*' generics) ('Dogs have 4 legs')
4. If ΔP^f_G is only somewhat above 0, and $P(f/G)$ is not high, $Value(f)$ has to be high for 'Gs are f' to be true. (*striking* generics) ('Ticks carry the Lyme disease')
5. If $Alt(f) = \{f, \neg f\}$, $Value$ is irrelevant and $\bigcup Alt(G) \cap f = \emptyset$, then '$G$s are f' is true just in case $P(f/G) > 0$, i.e., the *existential* reading. ('Indians $[\text{do}]_F$ eat beaf')

Even if hearers *accept* generics due to our proposed weak truth conditions, hearers still *interpret* generics about a relatively unknown group in a much stronger way: (almost) all Gs are f. Why?

According to a natural **pragmatic proposal**, the strengthening of 'Gs are f' about a relatively unfamiliar group is because *typicality* (or perhaps better: **sterreotypicality**) is **confused** with *probability* (or perhaps **prototypicality**). There is a natural psychological reason why hearers confuse high ∇P^f_G, i.e. $(P(f/G) - P(f/\neg G)) \times Value(f)$, with high $P^*(f/G)$. Here, probability function P^* is more subjective, and $P^*(f/G)$ typically higher than $P(f/G)$. Why? (i) Because we forget about $Value(f)$ due to Tverky & Kahneman's (1974) *avalability heuristics* and (ii) we tend to think that if f is representative for G, then $P^*(f/\neg G) \approx 0$, due to the *causality* and *representativeness heuristics*. Notice that if $P^*(f/G)$ is high we have explained why the generic 'Gs are f' is interpreted meaning that most Gs are f.
Ad (i), according to T & K's (1974) *availability heuristics*, people assess the probability of an event by the ease with which instances or occurrences can be brought to mind. But this is certainly the case for events involving individuals with horrific or appalling features i.e., ones with high $Value$.
Ad (ii), Tverky & Kahneman (1974) show that if we see a correlation, we tend to interpret it in the preferred (strongest) way: as *causal*. (Note that somes philosophers of science, e.g. Suppes, use ΔP^f_G to measure the causal strength of G for f). Moreover, they show that an event is seen as more likely than it actually is, if it can be understood: if it can be causally explained. Furthermore, according to the Associative Theory of Probability Judgements judgements (Gluck & Bower, 1988; Shanks, 1990), people typically confuse high ΔP^f_G with high $P(f/G)$, which explains Tverky & Kahneman's conjunction fallacy and the fact that people many times neglect base rates.

If time permits, we will show how our *weak semantic* approach combined with our *strong pragmatic* analysis explains why generics are excellent tools for persuasion, and easy to (mis)use by gifted advertisers and populists ('Muslims are terrorists'). Also, we will suggest that our analysis seems promising as well for the analysis of other phenomena, such as habituals ('Hillary Clinton is a liar'), dispositions, and conditionals.

Representative references: Ameel & Storms (2006), 'From prototypes to caricatures: Geometrical models for concept typicality', *Journal of Memory and Language*; Asher & Morreau (1995), 'What some generic sentences mean', *The Generic Book*; Barsalou (1985), 'Ideals, central tendency and frequency of instantiation as determinants of graded structure in categories', *J. of Experimental Psychology*; Cimpian et al (2010), 'The power of generic statements', *Cognitive Science*; Cohen (1996), *Think generic!*, PhD diss.; Gluck & Bower (1988), 'From conditioning to category learning: An adaptive network model', *Journal of Experimental Psychology: General*; Haslanger (2011), 'Ideology, Generics, and Common Ground', in *Feminist Metaphysics*; Kahneman & Tversky (1972), 'Subjective Probability: A Judgment of Representativeness', *Cognitive Psychology*; Leslie (2008), 'Generics: Cognition and acquisition', *The Philosophical Review*; Rescorla (1968), 'Probability of shock in the presence and absence of CS in fear conditioning', *J. of Comp. and Phys. Psychology*; Shanks (1995), Shanks (1995), *The Psychology of Associative Learning*, Cambridge University Press, Cambridge; Slovic et al. (2004), 'Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality', *Risk Analysis*; Tversky & Kahneman (1974), 'Judgment under Uncertainty: Heuristics and Biases', *Science*.