Proceedings of SALT 34: 000–000, 2024

Sense as sampling propensity

Michael Goodale
Institut Jean-Nicod, École normale supérieure

1 Introduction

It is little news to any of us that the name that you call something can matter a great deal. An author may agonise over a choice between two words which ostensibly refer to the same thing but carry very different connotations. This extends well beyond writing; if I tell my friend that I know a place that serves great *beverages*, they will likely conclude that those beverages will be very different than if I had referred to the very same beverages as *drinks*. The former brings to mind cold refreshment, a Coca-Cola or an iced tea and the latter, perhaps a glass of Cabernet or a Guinness. Crucially though, there is no drink which is not a beverage nor any beverage which is not a drink.

This poses a problem to standard extensional semantics, since there is no way to distinguish the set of beverages from the identical set of drinks. It might be tempting to suggest that idioms are playing a role here, but while "going for drinks" may be idiomatic, the distinction between the two words can show up in decisively non-idiomatic contexts. One might think that intensions which map from worlds to sets might provide us a way out, but they can't solve the drink/beverage puzzle since both terms designate the same set in every possible world. As a result, under standard assumptions, they have identical interpretations, $[\![\text{drink}]\!] = [\![\text{beverage}]\!]$.

Despite the concepts' identical sets, (1) seems true while (2) seems somewhat inappropriate or false. Furthermore, this is far from an idiosyncrasy of "drink/beverage", the same phenomenon occurs with "food/comestible" in (3).

(1) a. Drinks are consumed in bars.
b. Beverages are consumed in fast-food restaurants.

(2) a. ? Beverages are consumed in bars.
b. ? Drinks are consumed in fast-food restaurants.

(3) a. French people love food.
b. ? French people love comestibles.

This problem doesn't only apply to common nouns. Proper names carry sometimes massively different connotations too. As Saul (1997) has pointed out, alter-egos can provide a difficult muddle to standard semantic views.

(4) a. Clark Kent went into the phone-booth and Superman came out.

b. Superman went into the phone-booth and Clark Kent came out.

(5) a. Clark Kent is mild-mannered journalist.

b. # Superman is a mild-mannered journalist.

(4a) and (4b) seemingly have different truth conditions and while (5a) seems true, (5b) seems false. This is not restricted to only alter-egos or fiction either: Augustus was the first emperor of Rome and Octavian was a member of the Second Triumvirate and yet they are one and the same individual. Any historian worth their salt will be exceedingly careful in whether they refer to that individual as Augustus or Octavian as they convey different connotations.

It's easy to see why these multiple modes of reference are troubling for traditional semantics: if $[\![\text{Superman}]\!] = s$ and $[\![\text{Clark Kent}]\!] = [\![\text{Superman}]\!]$, then we should be able to substitute Clark Kent for Superman and vice-versa *salva veritae*. Problems of this sort have been well-known since Frege (1892) and were later expanded on by Kripke (1979) and various solutions have been proposed to aspects of the problem (e.g. Aloni's (2001) Conceptual Covers). However, these solutions generally interpret the problem as a matter of mistaken beliefs, incomplete information or alternative possible worlds: we can handle the fact that Lois Lane thinks Clark Kent isn't Superman because of the opaque context. What makes Saul's problem so important, however, is that examples like (4) and (5) do not involve ignorance or opaque contexts—one can be perfectly aware that Superman *is* Clark Kent and yet distinguish (4a) and (4b). Since we cannot rely on opaque contexts, we need to take a clear, hard look at our core semantics.

This paper proposes a semantic theory which addresses the issues with both nouns like *drink/beverage* and proper names like *Superman/Clark Kent* with a shared mechanism. This mechanism is able to handle these problems of modes of reference and also give an account of some core aspects of genericity (namely, it handles how certain generic statements are evaluated—it is not a grammatical theory about the distribution of genericity). The theory uses an approach introduced by Icard (2016), *sampling propensity*.

The approach boils down to the idea that we *sample* from an underlying generative procedure rather than directly accessing objects of type e or $\langle e,t\rangle$ when applying a predicate. For example, when I say "beverages are nice on a hot day", I conjure up in my mind exemplars of cold beverages and use them to evaluate the underlying proposition, i.e. whether those sampled exemplars of beverages happen to be nice on a hot day. This procedure is not restricted to common nouns like "drink", but also applies to individuals. When I evaluate "Superman is a mild-mannered journalist", I don't simply evaluate the proposition for a single unitary Superman representation, rather, I consider a swathe of Superman exemplars and determine whether I think it is true on the basis of the Superman exemplars I happen to sample.

The power of this system comes from the fact that two concepts like *drink* and *beverage* may share the same extension *but are likely to sample different things*. When I think of a drink, I'll think of wine and when I think of a beverage, I'll think of Coca-Cola, because while the two concepts have the same set of members, different members are likely to be sampled. This means that concepts are (at least) bipartite; they consist of a set of members *and* a sampling propensity over those members.[1] Drink and beverage have the same set but happen to have different sampling propensities, giving rise to a different connotations.

Likewise, thinking of Superman will draw to mind leaping tall buildings in a single bound whereas Clark Kent will lead one to think of mild-mannered journalism. The concepts for individuals are also bipartite consisting of a *set* of that individual in different contexts or different open possibilities for that individual *and* a sampling propensity over those potential individuals. Nevertheless, since Superman *is* Clark Kent, any of these exemplars sampled from either Superman or Clark Kent must necessarily be in both the set of potential Supermen and the set of potential Clark-Kents. Superman and Clark Kent have the same set of members but different sampling propensities over those members, just like drink and beverage do.

Sampling propensities can be formalised using probability distributions but I want to make it quite clear that the view is not committed in any way to the idea that probability distributions themselves are represented in the mind. The idea is that the mind is a generative machine and is capable of coming up with examplars of a concept when necessary. Some members may come quickly to mind, and some may take a very long time and some may essentially never be considered. For example, some beverages come to mind quickly and immediately (e.g. Coca-Cola) and others perhaps less quickly (e.g. New England switchel). We can describe which members come to mind using a probability distribution, but this doesn't mean that the underlying generative procedure is directly represented as probabilities (or indeed if it involves randomness at all).

One concrete consequence of this view is that while our formalism makes use of probabilities, at no point in the semantics can we directly access probabilities. This is because a person has no means by which to evaluate their probability of sampling a given exemplar, they can only sample. In other words, we have a big button that says "sample" on it, but we have no way of knowing beforehand which things will be sampled and with what probability.

Sampling propensities also are *not* beliefs about frequency, they are simply what comes to mind. An archaeological dig might draw to mind the discoveries of great

1 Indeed, others have proposed that concepts need multiple independent parts (Putnam 1975; Del Pinal 2015, 2016). I give an account principally of connotation and extension, while leaving other major aspects of conceptual structure open (e.g. similarity).

treasures like King Tut's tomb, even if one is perfectly aware that the vast majority of archaeological digs find broken pottery shards at best.

This formalism has a number of potential applications in semantic puzzles where modes of reference play a role. Beyond Saul's problem, it can provide a straightforward analysis of slurs (i.e. many theories of slurs argue they are co-extensional with the neutral referents for the same thing), euphemisms (e.g. "urinate" and "piss" may draw to mind different scenarios) and circumlocutions (e.g. "kill" versus "cause to die" or "people with disabilities" versus "the handicapped"). While I will not expound on these applications, they show the potential power of the formalism at explaining many of the core issues of *sense* that remain outstanding in semantics.

The formalism also has other appealing properties. It is fully compositional and has a tight connection between idealised competence and the noisiness of performance thanks to the law of large numbers in repeated sampling. An observant reader may be concerned about how this approach handles that quantification but have no fear, no quantified babies are thrown out with the bathwater here. While this paper focuses exclusively on common nouns and individuals, the approach could be extended to other parts of language and thought. In particular, verbs (if analysed with a neo-Davidsonian event semantics) would have a very straightforward extension of the theory.

1.1 Sampling drinks and beverages

A given concept for a common noun like $[\![\text{beverage}]\!]$ is a tuple with two parts:

i) A set of *e*-type objects which belong to a concept, including potential (i.e. not necessarily existing) members of a concept. I shall henceforth refer to these objects as **atoms**.

ii) A propensity to sample atoms from that set.

I will give the precise formal details later on, but we can formalise the first part with our good old fashioned sets of atoms and the second with that ever-increasingly popular tool, probability distributions. These distinctions between the extension of a concept and its sampling propensity call for some notational clarification.

i) The extension, written in boldface: **beverage**

ii) The sampling propensity, written in sans-serif font: $\mathsf{beverage}$

So a concept like "beverage" is a tuple consisting of a set, **beverage**, and a probability distribution, $\mathsf{beverage}$.

$$[\![\text{beverage}]\!] = \langle \textbf{beverage}, \mathsf{beverage} \rangle$$

Compositional rules extract either the extension or sampling propensity in the appropriate contexts. For example, the predicate "is a beverage" simply consists of the following (completely standard) interpretation:

$$[\![\text{is a beverage}]\!] = \lambda x.\textbf{beverage}(x)$$

When used as a subject, "beverage" is a bit more complicated. When we say, "beverages are refreshing", we need to sample the appropriate beverages. The fact this is a generic sentence is not incidental—sampling will turn out to have the strongest effect in generic sentences. To do this, we need to introduce a kind of quantifier that can map from sampling propensities to sets of type e which can then be quantified over. While the precise details of quantification in general will be handled later on, the first and most vital quasi-quantifier is μ.

$$\mu_{x \sim \langle \textbf{beverage}, \mathsf{beverage} \rangle} [\![\text{are refreshing}]\!](x)$$

This operator takes an extension and a sampling propensity and samples from the extension according the propensity. It then checks whether the predicate is true for those samples and determines the average truth-value. The number of exemplars we sample, N, is variable and increasing N corresponds to increasing mental effort or consideration.

$$\mu_{x \sim \langle \textbf{beverage}, \mathsf{beverage} \rangle} [\![\text{are refreshing}]\!](x) = \frac{1}{N} \sum_{1}^{N} [\![\text{are refreshing}]\!](\textsc{Sample}(\langle \textbf{beverage}, \mathsf{beverage} \rangle))$$

As we increase N, the truth-value judgement will be more and more consistent as we converge, by the law of large numbers, on the expected value. This stochastic process of repeated sampling corresponds to an account of performance for μ which relies on our ability to sample members of a concept. Remember, the conceit of sampling propensities is that while we can sample, we cannot directly access probabilities. However, since the truth-value converges as N increases, we can also discuss competence for μ which is defined in terms of the probability distribution directly (where $P(x \sim \mathsf{y})$ is the probability of sampling x from y):

$$\mu_{x \sim \langle \textbf{beverage}, \mathsf{beverage} \rangle} [\![\text{are refreshing}]\!](x) = \sum_{x \in \textbf{beverage}} P(x \sim \mathsf{beverage}) [\![\text{are refreshing}]\!](x)$$

This gives us a tight connection between performance and competence for μ where competence can be characterised as sampling in the limit.

To return to our original problem of drinks and beverages, we can say that these two concepts have identical extensions while having different sampling propensities.

$$\mathbf{drink} = \mathbf{beverage} \quad \mathsf{drink} \neq \mathsf{beverage}$$

As a consequence, we can have the distinction in (1) and (2) by positing that $[\![\text{consumed in bars}]\!]$ is true for regions of high sampling-propensity for $\langle\mathbf{drink}, \mathsf{drink}\rangle$ whereas it is false for regions of high-sampling propensity for $\langle\mathbf{beverage}, \mathsf{beverage}\rangle$.

$$\mu_{x\sim\langle\mathbf{beverage},\mathsf{beverage}\rangle}[\![\text{consumed in bars}]\!](x) < \mu_{x\sim\langle\mathbf{drink},\mathsf{drink}\rangle}[\![\text{consumed in bars}]\!](x)$$

1.2 Fuzzy logic

Since we are averaging the truth value of the predicate across samples, we have a fuzzy logic, with continuum-many truth values between 0 and 1. Classical connectives are usually defined as follows in fuzzy logic:

$$[\![\neg\varphi]\!] = 1 - [\![\varphi]\!] \qquad [\![\varphi \wedge \psi]\!] = \min([\![\varphi]\!], [\![\psi]\!]) \qquad [\![\varphi \vee \psi]\!] = \max([\![\varphi]\!], [\![\psi]\!])$$

It's important to note that these are not the only possible definition of connectives for infinite-valued logic. There is a broad class of possible fuzzy logics, known as t-norm logics (Hájek 1998). One of these may turn out to be a better match for human cognition than the preceding definitions, but they are sufficient for our purposes here.

While in the early 70s, there was some interest in integrating fuzzy logic with natural language semantics (Lakoff 1973), criticisms from Kamp (1975) proved devastating to that endeavour. Despite aspects of his argument having later been criticised by Sauerland (2011), the conclusion against fuzzy logic is still widely-accepted. Kamp and Sauerland both criticise fuzzy logic from the perspective of it being a potential analysis of vagueness in natural language. Even though the logic in my approach is fuzzy, it is not meant as an analysis of vagueness and is instead an analysis of something akin to the degree of belief in a proposition. Furthermore, only quantification can introduce fuzziness as atomic predicates remain bivalent.

Kamp's argument against fuzzy logic is relatively straightforward. If φ has a truth-value of 0.5, then $\varphi \wedge \neg\varphi$ must have a truth value of 0.5 since $\min(0.5, 0.5) = 0.5$ which Kamp deems "absurd" because "how could a logical contradiction be true to *any* degree?" (Kamp 1975: 131). It does seem troubling for $\varphi \wedge \neg\varphi$ to not be a contradiction, but under the fuzzy logic I have defined, it can still be a contradiction under certain assumptions.[2]

2 Experimentally, people will often accept this kind of "contradiction" as acceptable if the predicate is vague and it is a borderline case (Alxatib & Pelletier 2011; Serchuk, Hargreaves & Zach 2011; Ripley 2011).

Take a sentence like “Drinks are refreshing” and let us assume it has a truth value of 0.5. Recall, μ here takes a concept, $\langle \textbf{drink}, \textsf{drink} \rangle$ and samples atoms from the extension, **drink** using the sampling propensity, drink and returns the proportion of sampled atoms which belong are in the set defined by $[\![\text{are refreshing}]\!]$.

$$\begin{aligned}
[\![\text{Drinks are refreshing}]\!] &= \mu_{x \sim \langle \textbf{drink}, \textsf{drink} \rangle} [\![\text{are refreshing}]\!](x) \\
&= \sum_{x \in \textbf{drink}} P(x \sim \textsf{drink}) [\![\text{are refreshing}]\!](x) \\
&= 0.5
\end{aligned}$$

If we take its negation, this will also have a truth value of 0.5.

$$\begin{aligned}
[\![\text{Drinks are not refreshing}]\!] &= \mu_{x \sim \langle \textbf{drink}, \textsf{drink} \rangle} \neg([\![\text{are refreshing}]\!](x)) \\
&= \sum_{x \in \textbf{drink}} P(x \sim \textsf{drink})(1 - [\![\text{are refreshing}]\!](x)) \\
&= \sum_{x \in \textbf{drink}} P(x \sim \textsf{drink}) - \sum_{x \in \textbf{drink}} (x \sim \textsf{drink}) [\![\text{are refreshing}]\!](x)) \\
&= 1 - \sum_{x \in \textbf{drink}} (x \sim \textsf{drink}) [\![\text{are refreshing}]\!](x)) \\
&= 0.5
\end{aligned}$$

However, the conjunction of the two predicates will lead to a contradiction:

$$\begin{aligned}
\left[\!\!\left[\begin{array}{r}\text{Drinks are refreshing} \\ \text{and not refreshing}\end{array}\right]\!\!\right] &= \mu_{x \sim \langle \textbf{drink}, \textsf{drink} \rangle} [\![\text{are refreshing}]\!](x) \wedge \neg [\![\text{are refreshing}]\!](x) \\
&= \sum_{x \in \textbf{drink}} P(x \sim \textsf{drink}) \min([\![\text{are refreshing}]\!](x), 1 - [\![\text{are refreshing}]\!](x)) \\
&= \sum_{x \in \textbf{drink}} P(x \sim \textsf{drink}) 0 \\
&= 0 \quad (\text{Since } [\![\text{are refreshing}]\!](x) \text{ is always 0 or 1})
\end{aligned}$$

Of course, this largely depends on the scope of negation. If we assume negation goes around the predicate under μ, we get a contradiction. However, if negation scopes over μ, the conjunction of the two statements would have a truth-value of 0.5, but this additionally requires sampling from the concept two separate times.

1.3 Prototypicality, sampling propensity and compositionality

My approach bears some relationship to proposals from psychology about concepts like prototype theory. My theory makes no claims as to how we classify things into concepts, merely that we have a set of atoms for a concept and a method to sample from that set. Sampling propensity, however, corresponds quite closely to prototypicality—beer is a prototypical drink whereas Coke is a prototypical beverage, just as robins are quickly categorised as birds whereas penguins are not (Rosch 1975). In other words, something with high sampling propensity for a concept tends to be prototypical for that concept.

The relationship between generic sentences and prototypicality is not a novel one (Heyer 1985; Geurts 1985), but my approach provides a formal cashing out of this approach. One classic criticism of prototypical theories comes from Osherson & Smith (1981) who show that prototypical theories cannot account for compositional concepts such as "pet fish" as there is no way to derive the prototype of pet fish from the distance to the prototypes of pets and fish nor is there a way to compose the "features" of both prototypes. Fodor & Lepore (1996); Fodor (2001, 2008) claim this and other arguments show that while extensions may compose, stereotypes do not.

The approach in this article preserves set extensions and so, at the very least, extensions are easy to put together into compositional concepts in a classical manner. Sampling propensities are more complicated. One simple possibility could be that the sampling propensity of a single concept is continually fed upwards as the extension is s the extension is further restricted. On this view, "pet fish" would consist of the intersection of "pet" and "fish" combined with the sampling propensity for "fish".[3] This would mean that we would sample from fish, but restricted to only the ones which are pets.

$$[\![\text{Pet fish}]\!] = \langle \lambda x.\mathbf{fish}(x) \wedge \mathbf{pet}(x), \mathsf{fish} \rangle$$

This would mean for example, that "pet that is a fish" and "fish that is a pet" would also have different sampling propensities despite their identical extensions because in the former "pet" is the base sampling propensity and in the later "fish" is the base sampling propensity.

$$[\![\text{pet that is a fish}]\!] = \langle \mathbf{fish}(x) \wedge \mathbf{pet}(x), \mathsf{pet} \rangle$$
$$[\![\text{fish that is a pet}]\!] = \langle \mathbf{fish}(x) \wedge \mathbf{pet}(x), \mathsf{fish} \rangle$$

While this simplifying assumption is sufficient to illustrate how compositional concepts are possible in the system, it is almost certainly incomplete. Under this view, the ordering of the atoms of a concept are preserved through compositon (i.e. we can only exclude members, we can't change which members are more probable than others). So, for goldfish to be a prototypical pet fish, would require that the ordering of pet fish (e.g. a pet goldfish is more likely than a pet salmon) to be already

3 Noun-noun compounds are another can of worms. For simplicity, here I assume there is just simple intersective modification

contained in the prototype of fish (e.g. a non-pet salmon is more likely than a pet goldfish which is in turn more likely than a pet salmon).

This might be able to explain some compositional effects of prototypicality but it seems nigh certain that the modifiers of a complex concept ought to influence the sampling propensity of a concept beyond simple restriction. For example, people are likely to assume that Harvard-educated carpenters are idealistic despite neither the Harvard-educated nor carpenters being considered idealistic (Connolly, Fodor, Gleitman & Gleitman 2007). Further work will need to look at how sampling propensities can be influenced by their modifiers. For example, intersective modification could be modeled as the multiplication of the probability distributions underlying the sampling propensities. Alternatively, it may require actually looking inside the generative structure underlying a sampling propensity rather than treating it as a probabilistic black-box.

2 Individual concepts

The most unorthodox aspect to this approach is my treatment of individuals. Rather than treating Superman, Augustus or John as simple atoms (e.g. *e*-type objects), I treat them with the same structure as category concepts: a tuple of an extension and a sampling propensity. While it is naturally odd for an individual to have multiple atoms, it is in some sense analogous to a Stalnakerian context but relative to an individual rather than entire possible worlds. That is, while the set for an individual can contain multiple atoms with different properties, they are all candidates for the *actual* individual, just as the set of possible worlds should contain the actual world. Under this view, each atom in the set, **Superman**, would consist of another possible Superman who is compatible with our current knowledge and context. In other words, as we learn about an individual, we reduce the space of possible atoms that can be associated with them by making the set of atoms that refers to them smaller and smaller.

Since we know that Superman *is* Clark Kent, they must have the same set of possible-atoms. When we're sampling from that individual, it means that we're considering the different possibilities associated with that individual. So, to sample from Clark Kent means to be more likely to think of the Clark-Kent-y possibilities, namely that he's working as a journalist or wearing glasses. Conversely, Superman is more likely to be saving the day or wearing a red-and-blue leotard. So, just as with drink/beverage, we have identical extensions but differing sampling propensities.

$$\mathbf{Superman} = \mathbf{ClarkKent} \quad \mathsf{Superman} \neq \mathsf{ClarkKent}$$

Likewise, we can predict the distinction between (5a) and (5b) by a difference in the sampling propensities for Clark Kent and Superman.

$$\mu_{x\sim\langle\mathbf{ClarkKent},\mathsf{ClarkKent}\rangle}[\![\text{mild-mannered}]\!](x) > \mu_{x\sim\langle\mathbf{Superman},\mathsf{Superman}\rangle}[\![\text{mild-mannered}]\!](x)$$

Since speakers are aware of the different atoms that different concepts draw to mind, they use different concepts to convey extra information. When I say "Clark Kent" did something, my interlocutor will likely sample moments that are Clark-Kent-y (e.g. wearing glasses). As a result, I can convey different implicatures by using different terms since I am aware of the kinds of atoms that they draw to mind in the semantics. So, if I were to explain that Clark Kent/Superman was flying in the distance to someone who doesn't seem him, I would be more apt to use "Clark Kent" if he were wearing a tie and glasses and more apt to use "Superman" if he were dressed in a leotard.

Likewise, different definite descriptions can also change connotations: "the villain" will draw to mind very different things than "the man". We can model this using an iota operator which takes the sampling propensity from its definite description and its extension from the individual concept which sastifies it (see 4.3 for full formal details).

$$\iota(\langle\mathbf{villain},\mathsf{villain}\rangle) = \langle\mathbf{John},\mathsf{villain}\rangle \quad \iota(\langle\mathbf{man},\mathsf{man}\rangle) = \langle\mathbf{John},\mathsf{man}\rangle$$

As a result, the different definite descriptions which pick out the same individual will draw to mind different atoms belonging to that individual.

3 Quantification

Our individuals and categories have identical ontological structure: they are sets of atoms paired with a sampling propensity. If we leave this as is, then we face a rather obvious problem because quantification would break down as we could quantify over the possible atoms of a single individual. This requires revamping how we handle quantification. While this is different from standard first-order logic approaches, the compositional story is largely compatible with traditional generalised quantifier approaches.

In essence, rather than quantifying over atoms (e-type objects), we quantify over *individual concepts*. This requires that we keep track of which concepts are category concepts and which are individual concepts. We can do this by "tagging" concepts as category- or individual-concepts.

This also means we have to think about how individuals and atoms are counted. Say we have a model with two lions, named Simba and Mufasa. They both could have several atoms consisting of different possibilities for those individuals and all of these different atoms would be lions. Therefore, the set of the concept for lion

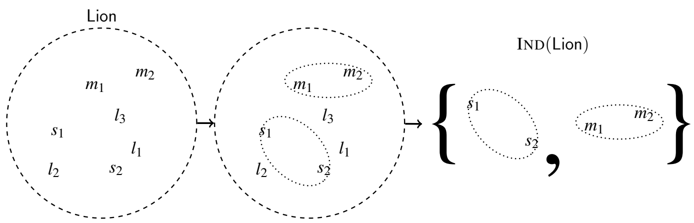


Figure 1 IND extracts the sampling propensities for two lion individual concepts, s and m, while ignoring l_1, l_2, and l_3 which are potential lion atoms that aren't a member of any individual concept.

would have many different lion-atoms but there would only be two lion individuals. So, when quantifying we would like to quantify over Simba and Mufasa, rather than all the different lion-atoms.

To do this, we need a new operator called IND which, given a category concept, extracts all individual concepts (defined as members of the set of individual concepts, $\mathscr{I}$) which are a subset of that concept.

$$[\![\textsc{Ind}(\langle\varphi,\mathsf{s}\rangle)]\!] = \{\langle\mathbf{p},\mathsf{p}\rangle \in \mathscr{I} \wedge \forall x\,(\mathbf{p}(x) \rightarrow \varphi(x))\}$$

For example, given the concept, $\langle\mathbf{lion},\mathsf{lion}\rangle$, $\textsc{Ind}(\langle\mathbf{lion},\mathsf{lion}\rangle)$ would give us the set $\{\langle\mathbf{s},\mathsf{s}\rangle,\langle\mathbf{m},\mathsf{m}\rangle\}$ consisting of the concepts for Simba and Mufasa (as shown in Figure 1). A generalised quantifier, then, would operate over that set of concepts, using μ on each individual concept to evaluate the predicate.

$$[\![\text{Five lions have a mane}]\!] = [\![\text{Five}]\!](\textsc{Ind}(\langle\mathbf{lion},\mathsf{lion}\rangle))(\lambda x.[\![\text{has a mane}]\!](x))$$
$$[\![\text{Every lion has a mane}]\!] = \forall_{\langle\mathbf{x},\mathsf{x}\rangle\in(\textsc{Ind}(\langle\mathbf{lion},\mathsf{lion}\rangle))}\left(\mu_{x\sim\langle\mathbf{x},\mathsf{x}\rangle}[\![\text{has a mane}]\!](x)\right)$$

While it is clearly necessary to quantify over each individual concept in a particular context, sometimes we want to make a quantificational statement about a concept in general. For example, the statement, "all men are created equal" is not merely a claim about whichever particular men (or women) who happen to exist right now; it is a claim about men (and women) who could ever exist. Indeed, many have remarked on the quasi-generic flavour of non-partitative "all" (Partee 1995; Gil 1995; Brisson 1998) which is missing from "every".

We can address this difference under this theory by positing two basic forms of quantification:

i. Individual-quantification (where we quantify over individual concepts using IND)

ii. Atom-quantification (where we quantify over atoms directly).

Atom-quantified sentences are quantificational statements that go over the concept in a generic manner. For example, μ is an atom-quantifier as well as non-partitative uses of all:

$$[\![\text{Lions have manes}]\!] = \mu_{x\sim\langle\textbf{lion},\mathsf{lion}\rangle}[\![\text{has a mane}]\!](x)$$
$$[\![\text{All lions have a mane}]\!] = \forall_{x\sim\langle\textbf{lion},\mathsf{lion}\rangle}[\![\text{has a mane}]\!](x)$$

Individually-quantified sentences (like every or five), on the other hand, quantify over the individual concepts, not their atoms.

Matthewson (2001) argues that non-partititive uses of English all have underlyingly a bare plural, unlike partitive uses or quantification with "every". My analysis is similar (see Figure 3 for a simple Montagovian treatment) but claims that the generic flavour arises from the difference between atom- and individual-quantification.

Children more understand generic quantifiers ontogenically earlier than other quantifiers, possibly due to generics being simpler cognitively (Hollander, Gelman & Star 2002; Leslie 2008). People also sometimes mistakenly think that universally quantified sentences are true if the corresponding generic sentence is true. For example, despite being aware that most lions are not adult male lions, people will often mistakenly agree with the statement that "all lions have manes" (Leslie, Khemlani & Glucksberg 2011).

This is compatible with a sampling-propensity approach as "all" merely samples a finite number of atoms (rather than individual concepts) and checks if the predicate is true for all atoms sampled. Crucially, the theory predicts that while people will do this with "all" (an atom-quantified quantifier), they shouldn't with "every" (an individually-quantified quantifier) because "every" operates over all (relevant) individual concepts rather than sampling atoms directly. As Leslie et al. (2011)'s experimental materials looked only at statements with "all", future work could see if this prediction is born out by evaluating differences in people's acceptance of "every" v.s. "all" in these kinds of statements. As a matter of intuitions, while both (7) and (6) are false, (6) has a significantly more appealing vibe than (7).

(6) All lions have manes.

(7) Every lion has a mane.

4 Formalism

4.1 Syntax

Signature The signature of this system is a tuple consisting of four parts:

i) $\mathscr{F}$: The set of function symbols

ii) $\mathscr{P}$: The set of predicate symbols

iii) $\mathscr{S}$: The set of sampling propensities.

iv) ar: An aritiy function such that $\mathscr{F} \cup \mathscr{P} \rightarrow \mathbb{N}$.

To help with clarity, each set has its own typographic conventions:

Set	Typographic convention	Examples	Use
$\mathscr{F}$	Normal serif font	f, g, h	Relational functions
$\mathscr{P}$	Boldface	$\mathbf{P}, \mathbf{Q}$	Predicates (e.g. **lion**(x) or **Superman**(x))
$\mathscr{S}$	Sans-serif	P, Q	Sampling propensities

Alphabet The alphabet, $\mathscr{A}^*$ consists of all members of $\mathscr{P} \cup \mathscr{F} \cup \mathscr{S}$ as well as the following:

- $\wedge$
- $\vee$
- $\neg$
- $\forall$
- $\exists$
- μ
- $\rangle$
- $\langle$
- $\sim$
- $=$
- $\bot$
- $\top$
- IND
- $\mathsf{x}, \mathsf{y}, \ldots$
- $\mathbf{x}, \mathbf{y}, \ldots$
- $x, y, \ldots$
- $\in$

Well-formed terms The set, $\mathscr{T} \subset \mathscr{A}^*$ of well-formed term can be defined inductively as follows:

i. If x is a variable, $x \in \mathscr{T}$.

ii. If $c \in \mathscr{F}$ and $\mathrm{ar}(c) = 0$, then $c \in \mathscr{T}$.

iii. If $f \in \mathscr{F}$, $\mathrm{ar}(f) \geq 1$ and $t_1, \ldots t_{\mathrm{ar}(f)} \in \mathscr{T}$ then $f(t_1, \ldots, t_{\mathrm{ar}(f)}) \in \mathscr{T}$.

Well-formed sampling terms The set of well-formed sampling-terms, $\mathscr{B} \subset \mathscr{A}^*$ is defined as follows:

i. If $\mathsf{s} \in \mathscr{S}$ and φ is a well-formed formula, then the concept for that pair, $\langle \varphi, \mathsf{s} \rangle \in \mathscr{B}$.

ii. If x is a sampling variable and $\mathbf{x}$ is a predicate variable, then $\langle \mathbf{x}, \mathsf{x} \rangle \in \mathscr{B}$.

Sampling terms include primitive concepts (e.g. $\langle \mathbf{drink}, \mathsf{drink}\rangle$), complex concepts (e.g. $\langle \lambda x.[\![\text{cold}]\!](x) \wedge \mathbf{drink}(x), \mathsf{drink}\rangle$) and concept variables (e.g. $\langle \mathbf{x}, \mathsf{x}\rangle$). Concept variables are used in individually-quantified statements like $[\![\text{Every lion has a mane}]\!] = \forall_{\langle \mathbf{x},\mathsf{x}\rangle \in (\textsc{Ind}(\langle \mathbf{lion}, \mathsf{lion}\rangle))} \left(\mu_{x\sim\langle \mathbf{x},\mathsf{x}\rangle} [\![\text{has a mane}]\!](x)\right)$. The predicate and sampling variables, $\mathbf{x}$ and x are used to access the predicate and sampling variable of the quantified individual concepts.

Well-formed formulae The set of well-formed formulae, $\mathscr{W} \subset \mathscr{A}^*$ can be defined inductively as follows:

i. $\top, \bot \in \mathscr{W}$

ii. If $P \in \mathscr{P}$ and $\text{ar}(P) = 0$, then $P \in \mathscr{W}$.

iii. If $P \in \mathscr{P}$ and $\text{ar}(P) \geq 1$ and $t_1, \ldots, t_{\text{ar}(P)} \in \mathscr{T}$, then $P\left(t_1, \ldots, t_{\text{ar}(P)}\right) \in \mathscr{W}$.

iv. If $\varphi \in \mathscr{W}$, then $(\neg\varphi) \in \mathscr{W}$.

v. If $\varphi, \psi \in \mathscr{W}$, then $(\varphi \wedge \psi) \in \mathscr{W}$ and $(\varphi \vee \psi) \in \mathscr{W}$.

vi. If $\varphi \in \mathscr{W}$, x is a variable and $\langle \mathbf{s}, \mathsf{s}\rangle \in \mathscr{B}$, then $\left(\forall_{x\sim\langle \mathbf{s},\mathsf{s}\rangle} \varphi\right) \in \mathscr{W}$.

vii. If $\varphi \in \mathscr{W}$, x is a variable and $\langle \mathbf{s}, \mathsf{s}\rangle \in \mathscr{B}$, then $\left(\exists_{x\sim\langle \mathbf{s},\mathsf{s}\rangle} \varphi\right) \in \mathscr{W}$.

viii. If $\varphi \in \mathscr{W}$, x is a variable and $\langle \mathbf{s}, \mathsf{s}\rangle \in \mathscr{B}$, then $\left(\mu_{x\sim\langle \mathbf{s},\mathsf{s}\rangle} \varphi\right) \in \mathscr{W}$.

ix. If $\varphi, \psi \in \mathscr{W}$, $\mathbf{x}$ is a predicate variable, and x is a sampling variable, then $\left(\forall_{\langle \mathbf{x},\mathsf{x}\rangle \in \textsc{Ind}(\psi)} \varphi\right) \in \mathscr{W}$

x. If $\varphi, \psi \in \mathscr{W}$, $\mathbf{x}$ is a predicate variable, and x is a sampling variable, then $\left(\exists_{\langle \mathbf{x},\mathsf{x}\rangle \in \textsc{Ind}(\psi)} \varphi\right) \in \mathscr{W}$

4.2 Semantics

A structure in the language is a triple, $(\mathbb{D}, \sigma, I, \mathscr{C}, \mathscr{I},)$, consisting of the domain, $\mathbb{D}$, the signature, σ, and the interpretation function, I, $\mathscr{C}$, and $\mathscr{I}$. $\mathscr{C}$ is the set of concepts and must meet the following condition.

$$\mathscr{C} = \{\langle \mathbf{p}, \mathsf{p}\rangle\} \text{ where for all } \langle \mathbf{p}, \mathsf{p}\rangle \in \mathscr{C},\ \mathsf{p} \in \mathscr{S} \wedge \mathbf{p} \in \mathscr{P} \wedge \text{ar}(\mathbf{p}) = 1$$

$\mathscr{I} \subseteq \mathscr{C}$ is the set of individual concepts in the model. These are concepts which refer to a single individual, e.g. Clark Kent.

Interpretation function

i. If $c \in \mathscr{F}$ and $\text{ar}(c) = 0$, then $[\![c]\!] \in \mathbb{D}$.

ii. If $f \in \mathscr{F}$ and $\text{ar}(c) \geq 1$, then $[\![f]\!]$ is a function with the same arity as f which maps from elements of $\mathbb{D}$ to an element of $\mathbb{D}$.

iii. If $p \in \mathscr{P}$ and $\text{ar}(P) = n$ then $[\![p]\!]$ is a function from n elements of $\mathbb{D}$ to $\{0,1\}$.

iv. If $[\![f]\!]$ is a function from n elements of $\mathbb{D}$, and $[\![t_1]\!], \ldots [\![t_n]\!] \in \mathbb{D}$, then $[\![f(t_1,\ldots,t_n)]\!]$ is $f([\![t_1]\!],\ldots,[\![t_n]\!])$.

v. $[\![\bot]\!] = 0$ and $[\![\top]\!] = 1$

Well-formed formulae without free variables have interpretations in the range $[0,1]$

Connectives The following definitions are here for completeness, however any t-norm based set of connectives would be appropriate. If φ and ψ are well-formed formulae, then:

$$[\![\neg\varphi]\!] = 1 - [\![\varphi]\!] \qquad [\![\varphi \wedge \psi]\!] = \min([\![\varphi]\!],[\![\psi]\!]) \qquad [\![\varphi \vee \psi]\!] = \max([\![\varphi]\!],[\![\psi]\!])$$

4.2.1 Sampling propensities and simple quantifiers

If $\mathsf{s} \in \mathscr{S}$, then $[\![\mathsf{s}]\!]$ is a function from $\mathbb{D}$ to $\mathbb{R}$ such that for all $x \in \mathbb{D}$, $s(x) > 0$[4] .

i. If φ is a well-formed formula and $\mathsf{s} \in \mathscr{S}$ then their associated sampling term, $[\![\langle\varphi,\mathsf{s}\rangle]\!] : \mathbb{D} \to [0,1]$ is the following function:

$$[\![\langle\varphi,\mathsf{s}\rangle]\!](x) = \begin{cases} \frac{[\![\mathsf{s}]\!](x)}{\sum_{x\in\{x\in\mathbb{D}\mid[\![\varphi(x)]\!]=1\}}[\![\mathsf{s}]\!](x)} & \text{if } x \in \{x \in \mathbb{D} \mid [\![\varphi(x)]\!] = 1\} \\ 0 & \text{otherwise} \end{cases}$$

Otherwise put, it is a probability distribution over the set of the domain where φ is true normalised according to s (other normalisation functions could be alternatives, e.g. soft-max with temperature).

ii. If φ is a well-formed formula, and $\langle\mathbf{t},\mathsf{t}\rangle \in \mathscr{B}$ is a well-formed sampling term, then:

$$[\![\forall_{x\sim\langle\mathbf{t},\mathsf{t}\rangle}\varphi]\!](\langle\mathbf{t},\mathsf{t}\rangle) = \begin{cases} 1 & \text{if } 1 = \sum_{x\in\mathbb{D}}[\![\varphi]\!](x)\cdot[\![\mathsf{t}]\!](x) \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

$$[\![\exists_{x\sim\langle\mathbf{t},\mathsf{t}\rangle}\varphi]\!](\langle\mathbf{t},\mathsf{t}\rangle) = \begin{cases} 1 & \text{if } 0 < \sum_{x\in\mathbb{D}}[\![\varphi]\!](x)\cdot[\![\mathsf{t}]\!](x) \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

$$[\![\mu_{x\sim\langle\mathbf{t},\mathsf{t}\rangle}\varphi]\!](\langle\mathbf{t},\mathsf{t}\rangle) = \sum_{x\in\mathbb{D}}[\![\varphi]\!](x)\cdot[\![\mathsf{t}]\!](x) \tag{10}$$

4 This means that all sampling propensities have a (small) probability of generating any atom in the domain. In a concept, $\langle\mathbf{p},\mathsf{p}\rangle$, the extension, $\mathbf{p}$, restricts the sampling propensity, p, to only sample from the extension. Since all sampling propensities are non-zero for the entire domain, no combination of a sampling propensity and a distribution can lead to an improper probability distribution.

Note that here, I give a characterisation of *competence* without stochasticity rather than *performance* where we expect actual sampling from $\langle \mathbf{t}, \mathsf{t} \rangle$.

4.3 Complex quantifiers and models

If φ and ψ are well-formed formulae and s is a sampling process, then:

$$\begin{aligned}
[\![\textsc{Ind}(\langle \varphi, \mathsf{s} \rangle)]\!] &= \{\langle \mathbf{p}, \mathsf{p} \rangle \in \mathscr{I} | \forall x (\mathbf{p}(x) \to \varphi(x))\} \\
[\![\forall_{\mathsf{x} \in \textsc{Ind}(\langle \psi, \mathsf{s} \rangle)} \varphi]\!] &= \bigwedge_{\mathsf{x} \in [\![\textsc{Ind}(\psi)]\!]} \varphi(\mathsf{x}) \\
[\![\exists_{\mathsf{x} \in \textsc{Ind}(\langle \psi, \mathsf{s} \rangle)} \varphi]\!] &= \bigvee_{\mathsf{x} \in [\![\textsc{Ind}(\psi)]\!]} \varphi(\mathsf{x})
\end{aligned}$$

We can also define the iota operator similarly:

$$[\![\iota(\langle \varphi, \mathsf{s} \rangle)]\!] = \begin{cases} \langle \mathbf{p}, \mathsf{s} \rangle \text{ if } \exists ! \mathbf{p} \text{ such that } \exists \mathsf{p} \text{ such that } \langle \mathbf{p}, \mathsf{p} \rangle \in \mathscr{I} \text{ and } \forall x \mathbf{p}(x) \to \varphi(x) \\ \text{otherwise undefined} \end{cases}$$

Note that this iota operator is defined in such a way so that we extract an individual concept and use its extension while using the sampling propensity of the input. In that way, “the man” or “the villain” could have different sampling propensities even when referring to the same individual.

The iota operator only works if there is a unique *extension* of an individual concept which is a subset of φ. To handle Superman/Clark Kent cases, there can be multiple individual concepts in $\mathscr{I}$ which are subsets of φ provided they have identical extensions (but possibly differing sampling propensities).

To access the predicate or sampling propensity of something retrieved by the iota operator, we add the following syntactic sugar. If $\iota(\langle \varphi, \mathsf{s} \rangle) = \langle \mathbf{p}, \mathsf{s} \rangle$, then:

$$\iota(\langle \varphi, \mathsf{s} \rangle)_{\text{pred}} = \mathbf{p} \qquad \iota(\langle \varphi, \mathsf{s} \rangle)_{\text{samp}} = \mathsf{s}$$

5 Montague grammar

The following fragment is not meant to provide a precise analysis on how languages uses the logical system described in this paper. Rather, it serves to provide a simple example which shows that compositional analyses are possible using the sampling propensity approach. Here, I am treating adjectives as a black-box which contribute only an extension. They likely also contribute sampling propensities, but I do not yet have such an analysis. I am also using quantifier-raising but an analysis with type-shifting would be equally possible.

- $[\![\text{John}]\!] = \langle \mathbf{j}, \mathsf{j} \rangle$
- $[\![\text{Mary}]\!] = \langle \mathbf{m}, \mathsf{m} \rangle$
- $[\![\text{man}]\!] = \langle \mathbf{man}, \mathsf{man} \rangle$
- $\mathscr{I} = \{\langle \mathbf{j}, \mathsf{j} \rangle, \langle \mathbf{m}, \mathsf{m} \rangle\}$
- $[\![\text{smart}]\!] = \lambda \langle \mathbf{t}, \mathsf{t} \rangle . \langle \mathbf{t}(x) \wedge \text{smart}(x), \mathsf{t} \rangle$
- $[\![\text{laughs}]\!] = \lambda x.\text{laugh}(x)$
- $[\![\text{saw}]\!] = \lambda yx.\text{saw}(x, y)$
- $[\![\varnothing_{\text{proper}}]\!] = \lambda \langle \mathbf{t}, \mathsf{t} \rangle \in \mathscr{I} . \langle \mathbf{t}, \mathsf{t} \rangle$
- $[\![\varnothing_{\text{GEN}}]\!] = \lambda \langle \mathbf{t}, \mathsf{t} \rangle P.\mu_{x \sim \langle \mathbf{t}, \mathsf{t} \rangle} P(x)$
- $[\![\text{all}]\!] = \lambda \langle \mathbf{t}, \mathsf{t} \rangle P.\forall_{x \sim \langle \mathbf{t}, \mathsf{t} \rangle} P(x)$
- $[\![\text{the}]\!] = \lambda \langle \mathbf{t}, \mathsf{t} \rangle . \iota(\langle \mathbf{t}, \mathsf{t} \rangle)$
- $[\![\text{a}]\!] = \lambda \langle \mathbf{t}, \mathsf{t} \rangle . \langle \mathbf{t}, \mathsf{t} \rangle$
- $[\![\text{every}]\!] = \lambda \langle \mathbf{t}, \mathsf{t} \rangle Q.\forall_{\langle \mathbf{x}, \mathsf{x} \rangle \in \text{IND}(\langle \mathbf{t}, \mathsf{t} \rangle)} \mu_{x \sim \langle \mathbf{x}, \mathsf{x} \rangle} Q(x)$
- $[\![\text{is}_{\text{copular}}]\!] = \lambda \langle \mathbf{t}, \mathsf{t} \rangle x.\mathbf{t}(x)$
- $[\![\text{is}_{\text{ident}}]\!] = \lambda \langle \mathbf{t}, \mathsf{t} \rangle \in \mathscr{I} \langle \mathbf{q}, \mathsf{q} \rangle \in \mathscr{I}.\mathbf{t} = \mathbf{q}$

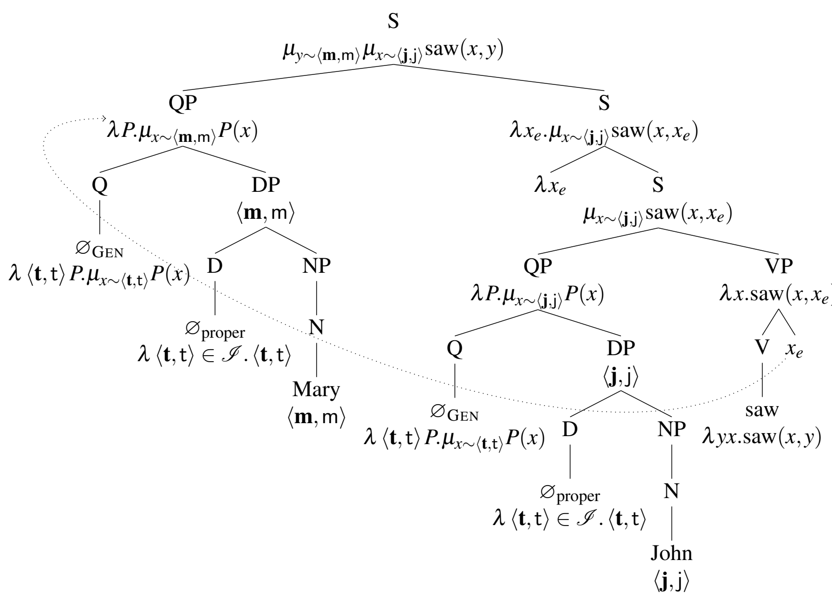
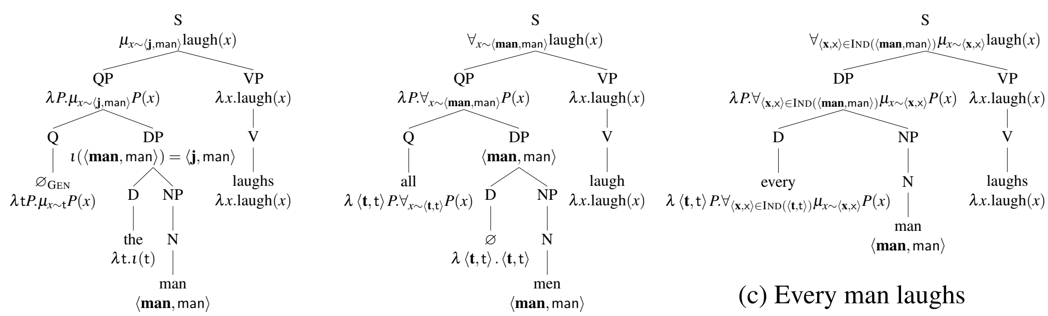


Figure 2 John saw Mary

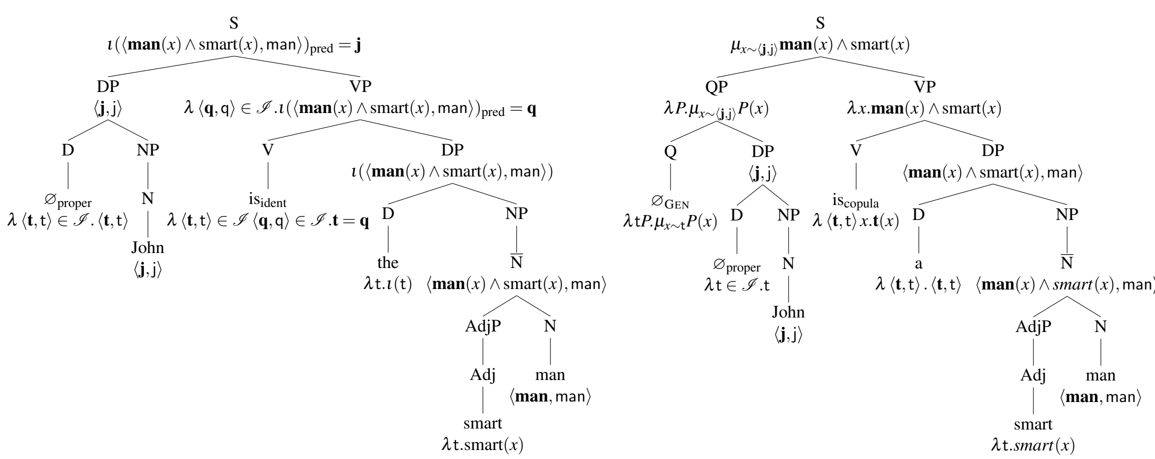


(a) The man laughs

(b) All men laugh

(c) Every man laughs

Figure 3 Different kinds of quantification with atom-quantified individual concepts (a), atom-quantified category concepts (b) and individually-quantified category concepts (c).



(a) John is the smart man

(b) John is a smart man

Figure 4 Differences between identificational and copular sentences

References

Aloni, Maria. 2001. *Quantification under Conceptual Covers*. Amsterdam, Netherlands: University of Amsterdam PhD dissertation.

Alxatib, Sam & Francis Jeffry Pelletier. 2011. The Psychology of Vagueness: Borderline Cases and Contradictions. *Mind and Language* 26(3). 287–326. doi:10.1111/j.1468-0017.2011.01419.x. Publisher: Blackwell.

Brisson, Christine D. 1998. *Distributivity, Maximality, and Floating Quantifiers*: Rutgers University PhD dissertation.

Connolly, Andrew C., Jerry A. Fodor, Lila R. Gleitman & Henry Gleitman. 2007. Why stereotypes don't even make good defaults. *Cognition* 103(1). 1–22. doi:10.1016/j.cognition.2006.02.005. https://www.sciencedirect.com/science/article/pii/S0010027706000564.

Del Pinal, Guillermo. 2015. Dual Content Semantics, privative adjectives, and dynamic compositionality. *Semantics and Pragmatics* 8. doi:10.3765/sp.8.7. http://semprag.org/article/view/sp.8.7.

Del Pinal, Guillermo. 2016. Prototypes as compositional components of concepts. *Synthese* 193(9). 2899–2927. https://www.jstor.org/stable/24897939. Publisher: Springer.

Fodor, Jerry A. 2001. Language, Thought and Compositionality. *Mind & Language* 16(1). 1–15. doi:10.1111/1468-0017.00153. https://onlinelibrary.wiley.com/doi/10.1111/1468-0017.00153.

Fodor, Jerry A. 2008. *LOT 2: The language of thought revisited*. Oxford : New York: Clarendon Press ; Oxford University Press. OCLC: ocn226360132.

Fodor, Jerry A. & Ernest Lepore. 1996. The red herring and the pet fish: why concepts still can't be prototypes. *Cognition* 58(2). 253–270. doi:10.1016/0010-0277(95)00694-X. https://www.sciencedirect.com/science/article/pii/001002779500694X. Number: 2.

Frege, Gottlob. 1892. Über Sinn Und Bedeutung. *Zeitschrift für Philosophie Und Philosophische Kritik* 100(1). 25–50. Number: 1 Publisher: Duke University Press.

Geurts, Bart. 1985. Generics. *Journal of Semantics* 4(3). 247–255. doi:10.1093/jos/4.3.247. https://doi.org/10.1093/jos/4.3.247.

Gil, David. 1995. Universal Quantifiers and Distributivity. In Emmon Bach, Eloise Jelinek, Angelika Kratzer & Barbara H. Partee (eds.), *Quantification in Natural Languages*, 321–362. Dordrecht: Springer Netherlands. doi:10.1007/978-94-017-2817-1_11. https://doi.org/10.1007/978-94-017-2817-1_11.

Heyer, Gerhard. 1985. Generic Descriptions, Default Reasoning, and Typicality. *Theoretical Linguistics* 12(1). 33–72. doi:10.1515/thli.1985.12.1.33. https://www.degruyter.com/document/doi/10.1515/thli.1985.12.1.33/html.

Hollander, Michelle A., Susan A. Gelman & Jon Star. 2002. Children's interpretation of generic noun phrases. *Developmental Psychology* 38(6). 883–894. doi:10.1037//0012-1649.38.6.883.

Hájek, Petr. 1998. *Metamathematics of fuzzy logic* (Trends in logic v. 4). Dordrecht ; Boston: Kluwer.

Icard, Thomas. 2016. Subjective Probability as Sampling Propensity. *Review of Philosophy and Psychology* 7(4). 863–903. doi:10.1007/s13164-015-0283-y. http://link.springer.com/10.1007/s13164-015-0283-y. Number: 4.

Kamp, J. A. W. 1975. Two theories about adjectives. In Edward L. Keenan (ed.), *Formal Semantics of Natural Language*, 123–155. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511897696.011. https://www.cambridge.org/core/books/formal-semantics-of-natural-language/two-theories-about-adjectives/DA68C05CF9AE0F0FC98CC73C79F844C0.

Kripke, Saul A. 1979. A Puzzle About Belief. In A. Margalit (ed.), *Meaning and Use*, 239–83. Reidel.

Lakoff, George. 1973. Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts. *Journal of Philosophical Logic* 2(4). 458–508. doi:10.1007/bf00262952. Publisher: Springer.

Leslie, Sarah-Jane. 2008. Generics: Cognition and Acquisition. *Philosophical Review* 117(1). 1–47. doi:10.1215/00318108-2007-023. Publisher: Duke University Press.

Leslie, Sarah-Jane, Sangeet Khemlani & Sam Glucksberg. 2011. Do all ducks lay eggs? The generic overgeneralization effect. *Journal of Memory and Language* 65(1). 15–31. doi:10.1016/j.jml.2010.12.005. https://www.sciencedirect.com/science/article/pii/S0749596X10001154. Number: 1.

Matthewson, Lisa. 2001. Quantification and the Nature of Crosslinguistic Variation. *Natural Language Semantics* 9(2). 145–189. https://www.jstor.org/stable/23748500. Publisher: Springer.

Osherson, Daniel N. & Edward E. Smith. 1981. On the adequacy of prototype theory as a theory of concepts. *Cognition* 9(1). 35–58. doi:10.1016/0010-0277(81)90013-5. https://linkinghub.elsevier.com/retrieve/pii/0010027781900135.

Partee, Barbara H. 1995. Quantificational Structures and Compositionality. In Emmon Bach, Eloise Jelinek, Angelika Kratzer & Barbara H. Partee (eds.), *Quantification in Natural Languages*, 541–601. Dordrecht: Springer Netherlands. doi:10.1007/978-94-017-2817-1_17. https://doi.org/10.1007/978-94-017-2817-1_17.

Putnam, Hilary. 1975. The Meaning of 'Meaning'. *Minnesota Studies in the Philosophy of Science* 7. 131–193. Publisher: University of Minnesota Press.

Ripley, David. 2011. Contradictions at the Borders. In Rick Nouwen, Robert van

Rooij, Uli Sauerland & Hans-Christian Schmitz (eds.), *Vagueness in Communication*, 169–188. Springer. https://philarchive.org/rec/RIPCAT.

Rosch, Eleanor. 1975. Cognitive representations of semantic categories. *Journal of Experimental Psychology: General* 104(3). 192–233. doi:10.1037/0096-3445.104.3.192. Place: US Publisher: American Psychological Association.

Sauerland, Uli. 2011. Vagueness in language : the case against fuzzy logic. In *Understanding vagueness : logical, philosophical, and linguistic perspectives, Studies in Logic, 36*, 185–198. College Publications. http://publikationen.ub.uni-frankfurt.de/frontdoor/index/index/docId/38575.

Saul, Jennifer M. 1997. Substitution and Simple Sentences. *Analysis* 57(2). 102–108. doi:10.1111/1467-8284.00060. https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-8284.00060. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-8284.00060.

Serchuk, Phil, Ian Hargreaves & Richard Zach. 2011. Vagueness, Logic and Use: Four Experimental Studies on Vagueness. *Mind and Language* 26(5). 540–573. doi:10.1111/j.1468-0017.2011.01430.x. https://philarchive.org/rec/SERVLA. Publisher: Blackwell.