

Genericity in similarity*

Janek Guerrini

February 2024

Abstract

In this paper, I give an account of constructions expressing similarity such as ‘like John’ and ‘like a lawyer’. The main point of the paper is that in ‘like a lawyer’, the indefinite receives a generic interpretation, which explains why under its most available reading, ‘John looks like a lawyer’ is equivalent to ‘John looks like a typical lawyer’. However, this indefinite is generic in a surprising way. Generic quantification is standardly thought to be brought about by a silent quantificational adverb, *Gen*, bearing a meaning akin to ‘generally’(cf. Krifka *et al.* 1995). It is therefore expected, on the standard picture, that an indefinite that can receive generic interpretations should also be bound by explicit quantificational adverbs, as for instance in ‘a bird flies’ \approx ‘typical birds fly’, parallel to ‘a bird rarely flies’ \approx ‘few birds fly’. However, indefinites embedded by ‘like’ escape this generalization: ‘John looks like a lawyer’ \approx ‘John looks like a typical lawyer’, but ‘John rarely looks like a lawyer’ $\not\approx$ ‘John looks like few lawyers’. To solve this puzzle, I propose that ‘like’ comes with a generic quantifier that is lexically hard-wired in its lexical entry, and show how this makes a number of surprising predictions which all turn out to be correct. Along the way, I also analyze properties of ‘like’ that are not necessarily linked to genericity, mainly: (i) it is a gradable expression over a closed scale, since it supports proportional modification such as in ‘the DNA of humans is 99% like that of chimps’. (ii) It can be modified both by scalar modifiers like ‘much’ and by ‘with respect to’ phrases like ‘with respect to size’, in similar but non-identical ways. (iii) It gives rise to homogeneity (cf. Križ 2015, a.o.), as ‘John is like Mary’ suggests they share all relevant properties, while ‘John isn’t like Mary’ suggests they share none of them.

*Acknowledgments to be added.

Contents

1	Introduction	3
2	‘Be like Mary’	6
2.1	Empirical landscape	6
2.1.1	Ellipsis?	6
2.1.2	Gradability, proportional modification, and ‘with respect to’ phrases	7
2.2	Previous approaches	11
2.2.1	Umbach & Gust (2014)	11
2.2.2	Alrenga (2010)	12
2.3	Analysis	13
2.4	Homogeneity	18
3	‘Look like Mary’	22
3.1	Distinguishing similarity from appearance	22
3.2	Similarity embedded by appearance: an illustration	23
3.3	Phenomenal versus epistemic embeddings of ‘like’ PPs	24
4	‘Be like a lawyer’	26
4.1	Similarities and differences between characterizing sentences and ‘like a <i>N</i> ’ constructions	26
4.2	Explaining the differences: the rigid Topic-Comment structure of sen- tences with ‘like’	29
4.3	Explaining the similarities: lexically inherent genericity	34
4.4	An extension of the analysis to non-copular occurrences of ‘like’	40
4.5	A note on the perceived asymmetry of sentences with ‘like’	41
5	Discussion and open issues	44
5.1	Other lexically generic constructions	44
5.2	Temporal arguments	47
6	Conclusion	49

1 Introduction

What do we mean when we say that someone is like a lawyer? An idea that underlies much work in psychology is that ‘like a *N*’ constructions correspond in a way or another to *assessments of similarity to a category*, by way of similarity to a prototype, to an exemplar, or so on (cf. Rips 1989; Gentner & Markman 1997; Hampton 1998, a.o.). Such an assumption leaves a number of questions open: why is it specifically the singular indefinite that triggers a ‘similarity to a prototype’ interpretation? In general, one would want to know how to articulate the connection between similarity to a category and similarity to an individual.

In this paper, I will argue that words expressing similarity do not inherently involve the comparison to a *category*. Instead, similarity itself is always understood to be similarity between *individuals*, and the indefinite in ‘like a lawyer’ receives a generic interpretation. This allows us to reduce ‘like a lawyer’ to ‘like John’. The presence of genericity explains why, under its most available reading, (1) implies that the object looks like a *typical* gun.

(1) This object looks like a gun.

Whether there is a connection to the structure of our concepts in terms of e.g. prototypes, then, depends solely on whether there is a connection between genericity and the structure of concepts.

Before delving into details, it is worth discussing why understanding talk about similarity matters. The main motivation for this work is that fleshing out the connection between similarity and genericity has the potential to illuminate larger issues that concern the interaction between the structure of psychological representations and the grammar, for a number of reasons.

First, it has been alternatively argued and assumed both in psychology (Gelman *et al.*, 2003) and in semantics (Leslie, 2008, 2015; van Rooij & Schulz, 2020) that genericity is itself a linguistic phenomenon that entertains a privileged link to the structure of psychological representations. Thus seeing it pop up in talk about similarity is not without interest, given the relevance of similarity for theories of concepts and categories.

Second, similarity is intuitively involved in a number of other phenomena that are both traditionally problematic for semantic theory and interesting from the perspective of the semantics-psychology interface. To give only one example, privative adjectives have been repeatedly claimed to provide a vantage point over the structure of lexical items or even of non-linguistic concepts (Franks, 1995; Pustejovsky, 1998; Coulson & Fauconnier, 1999). These adjectives are in many cases paraphrased in terms of similarity (Del Pinal, 2015, 2018; Guerrini, 2021). This means that the way one views privative adjectives may partly hinge on how one views constructions that express similarity.

Delving a bit more into details, similarity also provides an exceptional testbed to understand unattended facets of generic quantification. I will show that indefinites under

‘like’ display a puzzling combination of properties. On the one hand, they appear to be bound by the generic quantifier.

- (2) a. John looks like a lawyer.
 \approx John looks like a typical lawyer.
 b. $GEN_x[x \text{ is a lawyer}][SIMILAR(john, x)]$

On the other hand, they seem to feature a rather special kind of genericity. The generic quantifier *GEN*, present in the Logical Form of characterizing sentences such as ‘A bird flies’, is traditionally thought to be brought about by a silent quantificational adverb (Krifka *et al.*, 1995), call it *Gen*.¹ *Gen* adjoins to the VP as a part of verbal aspect (cf. Chierchia 1998, a.o.), and can be thought of, as a rough approximation, as a silent counterpart of ‘generally’, ‘typically’, or ‘usually’.

The main argument for thinking that genericity is generally adverbial is that it patterns exactly with overt quantificational adverbs in terms of what parts of the sentence go into its restrictor and which ones go into its scope. Consider for instance the traditional example in (3) (Krifka *et al.*, 1995). If the sentence is understood to be about computers, it receives the logical form in (3a). If it is understood to be about modern planes, it receives the logical form in (3b). The same goes for a counterpart of this sentence formulated with an explicit quantificational adverb like ‘often’, as in (4).

- (3) A computer routes a modern plane.
 a. $GEN_x[x \text{ is a computer}][x \text{ routes some modern plane}]$
 ‘If something is a computer, it routes some modern plane’
 ‘Typical computers route some modern plane’
 b. $GEN_x[x \text{ is a modern plane}][x \text{ is routed by some computer}]$
 ‘If something is a modern plane, it is routed by a computer’
 ‘Typical modern planes are routed by some computer’
- (4) A computer often routes a modern plane.
 a. $OFTEN_x[x \text{ is a computer}][x \text{ routes some modern plane}]$
 ‘Often, if something is a computer, it routes some modern plane’
 ‘Many computers route some modern plane’
 b. $OFTEN_x[x \text{ is a modern plane}][x \text{ is routed by some computer}]$
 ‘Often, if something is a modern plane, it is routed by a computer’
 ‘Many modern planes are routed by some computer’

‘Like a *N*’ constructions become extremely interesting in that, although the indefinite has the logical properties of genericity, it does not behave at all like the indefinites in (3). For instance, in the presence of an overt quantificational adverb, a sentence

¹I here distinguish between the lexical item *Gen*, a silent adverb, and the underspecified quantifier *GEN* that is part of the meaning of *Gen*.

involving ‘like’ cannot bear a logical form parallel to (4b):

(5) A judge is often like a lawyer.

Cannot mean:

OFTEN_x[*x* is a lawyer][there is a judge that is like *x*]

‘Often, if someone is a lawyer, there is a judge looks like them’

‘Many lawyers are such that there is a judge that looks like them’

Because of these contrasts, I will conclude that the generic quantifier GEN need not be systematically brought about by a silent quantificational adverb like *Gen*. Rather, looking at words expressing similarity allows us to see that GEN can in fact be hard-wired into the lexical entry of certain lexical items. I will argue that the argument of ‘like’ systematically goes into the restrictor of a generic operator that is lexically encoded by ‘like’.

(6) *like a lawyer* = $\lambda x.GEN_y[y \text{ is a lawyer}][SIMILAR(x,y)]$

This will allow us to explain a number of extremely surprising properties, starting from the near equivalence of the sentences in (2). Another surprising fact worth mentioning at the outset is that disjunction can receive quite strong readings:

(7) John looks like a judge or a lawyer.

≈ John looks like a judge and John looks like a lawyer.

This is easily explained if we assume that the disjunction is interpreted inside the restriction of the generic quantifier, and otherwise puzzling.

(8) *like a judge or a lawyer* = $\lambda x.GEN_y[y \text{ is a judge or } y \text{ is a lawyer}][SIMILAR(x,y)]$

This paper is structured in broadly two parts. In the first part, in section 2 I build up the first layer of the account, dealing with similarity between individuals and discussing all the properties of similarity that can be discussed without bringing genericity into the picture, in particular gradability and homogeneity. In section 3, I illustrate how this analysis can be embedded in other copular constructions such as ‘look like’.

In the second part, in section 4, I show how to reduce ‘like a lawyer’ to ‘like John’, by outlining my proposal based on lexically encoded genericity.

2 ‘Be like Mary’

In building up the first layer of the account, I first discard alternative theoretical possibilities: in 2.1.1 I show that we cannot capture ‘being like Mary’ as ‘being like Mary is’, as ‘like’ fails tests for ellipsis. I then show that previous theories of ‘like’ or related expressions fail to account for the full range of its properties. I then give an account, and discuss how it is compatible with (i) the homogeneous behaviour of ‘like’ and (ii) the perceived asymmetry of sentences involving ‘like’.

2.1 Empirical landscape

2.1.1 Ellipsis?

A quite simple theory of ‘like’ would view constructions as ‘like Mary’ as elliptical: ‘John is like Mary’ would be ‘John is like Mary is’, and ‘like’ always clausal. I think there are strong arguments to think that the complement of ‘like’ in fact does not behave like other elliptical constructions such as English clausal comparatives. In particular, English clausal comparatives give rise to the well-known ambiguity in (9).

- (9) John wants to write more articles than Mary.
a. John wants to write more articles than Mary wrote.
b. John wants to write more articles than Mary wants to write.
- (10) John wants to write more articles than books.
a. John wants to write [more articles than books].
b. John wants to write more articles than he wants to write books.

Not all comparative prepositions yield this ambiguity. Italian has two ways of expressing comparison: via the preposition ‘di’ and via the coordinator ‘che’ (see for instance Napoli & Nespors 1986). Italian sentence (11), featuring the clausal comparative ‘che’, is ambiguous between (11a) and (11b).²

- (11) Gianni vuole scrivere più articoli che libri.
Gianni wants write more articles than-CL books.
‘Gianni wants to write more articles than books’
a. Gianni wants to write [more articles than books].
b. Gianni wants to write more articles than he wants to write books.

A similar ambiguity does not obtain with the preposition ‘di’:

- (12) Gianni vuole scrivere più articoli di Maria.
Gianni wants write more articles than-N Maria.
‘Gianni wants to write more articles than Maria’

²I illustrate with an object-argument in (11) because ‘che’ displays a ban against subject-arguments (cf. Napoli & Nespors 1986). I illustrate the data in Italian because it provides a cleaner testbed than English in this respect. While English ‘than’ is either taken to be unambiguously clausal or ambiguously clausal or phrasal, Italian ‘che’ is standardly taken to be a coordinator, while ‘di’ is unambiguously phrasal.

- a. Gianni wants to write more articles than Maria has written.
- b. # Gianni wants to write more articles than Maria wants to write.

If ‘like a *N*’ constructions are systematically reduced clauses, then we expect them to behave more like Italian ‘che’ than like Italian ‘di’. Specifically, we expect a sentence like (13) to be ambiguous between reading (13a) and reading (13b); but we only observe (13a).

- (13) John wants to look like Mary.
- a. John wants to look like Mary looks.
 - b. *John wants to look like Mary wants to look.

One may think that ‘like’ has potential independent limitations, say, the movement of the ‘like’-phrase required to get reading (13b) may result in uninterpretability for independent reasons.³ However, this would not explain why the ambiguity is present in the case of ‘do’ ellipsis, as shown in (14):

- (14) John wants to look like Mary does.
- a. John wants to look like Mary looks.
 - b. John wants to look like Mary wants to look.

A second argument against a reduced clause view of ‘like’ constructions concerns wh-extraction: in (15), the trace cannot be extracted from the embedded clause, regardless of whether there is ‘do’ ellipsis or a clause with an explicit verb.

- (15) a. *Who does John look like [__ looks]?
- b. *Who does John look like [__ does]?

By contrast, extraction from phrasal ‘like’ is possible.⁴

- (16) Who does John look like __?

I take these facts to show that the complements of ‘like’ in constructions such as ‘like John’ and ‘like a lawyer’ are in fact nominals.

2.1.2 Gradability, proportional modification, and ‘with respect to’ phrases

The comparison provided by ‘like’ phrases is (i) gradable over a closed scale, and (ii) obtains based on specific similarity respects.

(i) Gradability over a closed scale

‘Like’ supports modification by scalar modifiers like ‘much’.

- (17) John is much like Mary.

³I thank the editor for bringing up this possibility.

⁴Of course, the sentences in (15) are probably bad because we are attempting to extract from a subject island. But if (16) were elliptical, we would be extracting out of a subject island there as well, and would expect (16) to be bad too.

The scale structure is closed, as both ‘like’ and its antonym ‘unlike’ support both proportional modifiers ((18), (20)) and maximal modifiers ((19)). Felicitous modification by proportional and maximal modifiers are standard tests for scale structure (cf. Kennedy & McNally 2005; Rotstein & Winter 2004).

- (18) She’s mostly {like, unlike} her mother.
- (19) Imagine a planet which is {perfectly, completely} {like, unlike} ours, except for its atmosphere.
- (20) a. The DNA of chimps is 99% like that of humans.
b. The DNA of chimps is 1% unlike that of humans.

Sentence (20a) clearly refers to the proportion of sequences shared by the two species: out of 100 nucleotides, chimps and humans share 99. I will take it to be a desideratum for a compositional theory of ‘like’ to ‘get the proportions right’ across uses of ‘like’.

Human +4	1	CCCAGTTCCTGTGGCTCCCTCGGAATGCTAAGGGGATCGGACATGAAAGGAC
Chimpanzee	1	CCCAGTTCCTGTGGCTCCCTCGGAATGCTAAGGGGATCGGACATGAAAGAAC

Figure 1: This table is taken from Durbin *et al.* (1997) (and used here by mere way of illustration).

(ii) Explicit or implicit similarity respects

The relevant similarity respects can be provided by context, as in (21a), or overtly specified with a ‘with respect to’ phrase (henceforth: WRT phrase), as in (22).

- (21) “Mary reacted well to the news. And given his personality, I expect to John to react well, too. You know...”
a. John is like Mary.
- (22) With respect to personality, John is like Mary.

(iii) Similarities between scalar modifiers (‘much like’) and WRT phrases (‘in many respects’)

When more than one criterion for similarity is relevant, (i) and (ii) can both bear on the proportion of relevant dimensions along which the two objects are close. To see this for scalar modifiers, consider (23).

- (23) **Context:** *we are comparing a number of exoplanets to Earth along four dimensions: atmosphere type, surface type, strength of gravity on the planet, and average elevation above sea level.*
a. (**Context:** *Planet P1 has the same atmosphere type as Earth, but different surface type, strength of gravity, and average elevation.*)
(i) P1 is a bit like Earth.
b. (**Context:** *Planet P1 has the same atmosphere type, surface type, strength of gravity, and average elevation as Earth.*)
(i) P1 is completely like Earth.

So, (23a)-(i) is compatible with a scenario in which the proportion of relevant dimensions along which P1 is close to Earth is little but not null, and (23b)-(i) with one where it is maximal. Similar results can be achieved with overt WRT phrases, too. Sentences (24a)-(i) and (24b)-(i) are compatible with the same scenarios as their gradable counterparts in (23a)-(i) and (23b)-(i):

- (24) a. (**Context:** Planet P1 has the same atmosphere type as Earth, but different surface type, strength of gravity, and average elevation.)
 (i) P1 is like Earth in a few respects.
- b. (**Context:** Planet P1 has the same atmosphere type, surface type, strength of gravity, and average elevation as Earth.)
 (i) P1 is like Earth in every respect.

(iv) Differences between scalar modifiers ('much like') and WRT phrases ('in many respects')

Scalar modifiers and WRT phrases differ, however, in that only scalar modifiers can bear on the similarity between two individuals *within* single dimensions. A sentence with a scalar modifier like (26a) is compatible with a scenario where P1 is *somewhat* close to Earth on *every* dimension, as for instance in a context like (25). By contrast, a counterpart with a WRT phrase such as sentence (27a) is not an adequate description of the facts described in (25).

(25) **Context:**

	Earth	P1
Atmosphere type	Terrestrial (N2, O2)	O2
Surface type	Mixed (land and water)	Water
Strength of gravity	9.81 m/s ²	5 m/s
Maximum oceanic depth	10,994 meters	15,000 meters

(26) (**Context:** as described in (25))

- a. P1 is a bit like Earth.

(27) (**Context:** as described in (25))

- a. P1 is like Earth in a few respects.

Additionally, scalar modifiers are felicitous when only one respect is relevant. For instance, in a context where only gravity strength is relevant, to state that P1 has a strength of gravity somewhat close to Earth's 9.8m/s², we can use a scalar modifier as in (28a). By contrast, a WRT phrase that is nearly equivalent to 'a bit' (at least for what concerns modification between respect) cannot be used, as shown in (28b).

(28) (*Two students are talking about P1, and realize that the strength of gravity of P1 is very relevant for their project.*)

"I don't really remember exactly what P1's strength of gravity is, but I know for sure that..."

- a. It is a bit like Earth's.
 b. # It is like Earth's in a few respects.

Sentence (28a) suggests that the strength of gravity of P1 is somewhat close to that of Earth. (28b) cannot bear this meaning, and would only be felicitous in a context in which strength of gravity could itself be broken down into some other dimensions.

Additionally, scalar modifiers and WRT phrases can be stacked. In this case, the scalar modifier acts within dimensions, and the WRT phrase between dimensions.

- (29) Context: as in (23).
- a. P1 is exactly like Earth in every respect.
 - b. P1 is like Earth in every respect.

Sentence (29a) differs from (29b) in the following way. In (29b), P1 is close *enough* to Earth along every dimension. In (29a), P1 is *maximally* close to Mary along every dimension. To see that this is truth-conditionally relevant, consider the contrast in felicity between (30b) and (31b):

- (30) a. P1 isn't like Earth in every respect.
 b. # ...For instance, with respect to the composition of its atmosphere, it is like Earth, but only moderately so.
- (31) a. P1 isn't completely like Earth in every respect.
 b. ...For instance, with respect to the composition of its atmosphere, it is like Earth, but only moderately so.

Summary

To summarize, we find that scalar modifiers are felicitous when only one respect is relevant, while WRT⁵ phrases aren't. Instead, both are felicitous when more than one respect is relevant, as summarized below.

	One relevant respect	Multiple relevant respects
Scalar modifiers	✓	✓
WRT Phrases	✗	✓

We also find that scalar modifiers seem to bear on single respects when only one respect is relevant, or when the respects are overtly specified. Otherwise, they can bear on the quantity of similarity respects along which the two individuals are close. WRT phrases, instead, always modify between respects, as summarized in the table below.

⁵For now I only consider plural WRT phrases such as 'in a few respects'. I will discuss expressions such as 'in one respect' when discussing my analysis.

	Within Respects	Between Respects
Scalar modifiers	Necessarily when only one respect is provided by context; possibly if more than one respect is provided by context; necessarily in presence of an overt WRT phrase.	Possibly when more than one respect is provided by context.
WRT Phrases	Never	Always

2.2 Previous approaches

2.2.1 Umbach & Gust (2014)

Umbach & Gust (2014), Gust & Umbach (2015), and Umbach & Gust (2021) provide an analysis of similarity demonstratives like ‘such a’ and extend it to similarity predicates like ‘similar’. They present a mathematically sophisticated framework to integrate referential semantics and conceptual structures. In a nutshell, in their framework individuals in a domain are mapped into multi-dimensional attribute spaces by generalized measure functions. Just like there are predicates that denote sets of individuals in the domain, there are predicates holding of sets of points on an attribute space that approximate predicates of individuals. In their system, a ‘representation’ is a tuple constituted by an attribute space, a generalized measure function, and a set of predicates over the attribute space. They then define similarity between two objects as indistinguishability with respect to predicates defined on the dimensions of the attribute space of a given representation: in their system, two objects are similar if the measure function of a given representation maps them onto points on the attribute space of the representation that are equivalent with respect to all predicates that are available in the representation.

Umbach and Gust implement gradable similarity in terms of the granularity of a given system of predicates over an attribute space of a representation. Essentially, a representation is more fine-grained than another one if the set of predicates of the former makes more distinctions between points on the attribute space than the set of predicates of the latter does.⁶ Scalar modification and comparatives are then defined in terms of representation granularity: *a* is more like *b* than like *c* if given a family of representations, there is a representation on which *a* and *b* are indiscernible and *a* and *c* are not indiscernible, and this representation is more fine-grained than any representation in this family making *a* and *c* indiscernible.

Their work opens new perspectives on the semantics-psychology interface that go well beyond the scope of this paper, and specifically on a more organic integration of theories about the way humans categorize objects with theories about how they talk about them. This can be seen as part of a larger body of the literature that has been calling to enrich compositional semantics to account for how natural languages encode

⁶To be more precise, in their system a Representation 1 is at least as fine-grained as Representation 2 if (i) the predicates of Representation 1 contain the predicates of Representation 2 and (ii) if for any two points on the attribute space, if they are equivalent with respect of the predicates of Representation 1, then they are equivalent with respect of the predicates of Representation 2 as well. So representation 1 is more fine-grained than representation 2 if 1 is at least as fine-grained as 2, but 2 is not at least as fine-grained as 1.

non-linguistic conceptual knowledge, such as for instance McNally (2005); Del Pinal (2015, 2018); Martin (2022).

Umbach and Gust do not discuss explicitly the generic flavour of indefinites embedded by similarity predicates ('like a lawyer' but also 'similar to a lawyer') presented in the introduction of this paper, and accounted for in section 4, which I take to be a desideratum for an explicit compositional treatment of 'like'. However, given the very direct link they establish between compositional semantics and conceptual thinking, it would be interesting for future research to explore how genericity can be captured in their framework, and how it interacts with similarity.

They do not explicitly discuss the difference between the two levels of modification supported by 'like' either (the facts are equivalent for 'similar'), i.e. scalar modifiers and WRT phrases, but there may be a way to capture those facts by assuming that scalar modifiers manipulate the granularity of representations, while WRT phrases bind different representations (this would be similar to the idea that I will present as part of my positive proposal, in fact). They also do not discuss modifiers such as '99%'. 'Getting the proportions right' when trying to capture such modifiers is not straightforward within their approach, as it is not obvious how one maps an ordering of representations in terms of fine-grainedness to proportions such as '99%'.

For this reason, I will resort to a simpler framework, which does not make use of attribute spaces and generalized measure functions, but only of sets of properties, intended in the standard way. This kind of approach has been already defended in work by Alrenga, which I now turn to discussing.

2.2.2 Alrenga (2010)

Alrenga (2010) presents a theory in which 'like' relates sets of properties:

$$(32) \quad \llbracket \text{like} \rrbracket = \lambda X_{(e,t),t}. \lambda Y_{(e,t),t}. \lambda P_{e,t}. P \in \mathbf{R} \wedge P \in X \wedge P \in Y$$

(With \mathbf{R} a set of contextually salient properties)

On this view, the set of (relevant) properties common to the arguments X or Y gets existentially closed when 'like' is unmodified, or fed to scalar modifiers like 'much', which impose a condition on the cardinalities of the relevant properties that are shared.

$$(33) \quad \llbracket \text{much} \rrbracket = \lambda X. |X| > n$$

An unwelcome prediction of this account is that 'like' in positive sentences, since it is existential, should not license exceptives. Licensing of exceptives is a well-established test for universal force, in particular concerning quantification over dimensions in multidimensional predicates (Sassoon, 2013), and 'like' seems to license them, as shown in (34).

$$(34) \quad \text{This planet is like ours, except for its atmosphere.}$$

By contrast, a welcome prediction is that negative sentences with 'like' should license

exceptives:

(35) This planet isn't like ours, except for its atmosphere.

This pattern, where positive sentences seem to behave like universals while negative sentences behave like negated existentials, is highly reminiscent of homogeneity, which we will discuss in section 2.4.

Secondly, Alrenga does not take a stance on whether the argument of 'like' is uniformly clausal (hence elliptical in cases like 'like John') or alternatively clausal and phrasal – but we have seen, in section 2.1.1 that a uniform analysis in terms of ellipsis of descriptive phrasal 'like' would present significant difficulties.

Finally, concerning the empirical generalizations outlined in the discussion in section 2.1.2 (i.e. (i) similarities and differences between scalar modifiers and (ii) proportional modification), Alrenga does not explicitly extend his account to them, but I think they can be accounted for in a theory not too dissimilar from his framework. In what follows, I present my analysis in which scalar modifiers receive interpretations similar⁷ to Alrenga's entry in (33), while WRT phrases *bind the sets of properties considered for the assessment of similarity*. We will see that this yields the correct results.

2.3 Analysis

I propose that, given a relevant dimension of comparison, 'like' phrases are gradable constructions just like regular gradable adjectives. The dimensions of comparison are sets of properties A . An expression such as 'like Earth' relates pairs of individuals to degrees of similarity, i.e. the ratio between the number of overlapping properties and the number of properties of the individual that has more of the relevant properties.^{8,9}

$$(36) \quad \llbracket \text{like} \rrbracket = \lambda y. \lambda d. \lambda \underline{A}_{\langle \langle e,t \rangle, t \rangle}. \lambda x. \frac{|\{P : P \in \underline{A} \wedge P(x)\} \cap \{P : P \in \underline{A} \wedge P(y)\}|}{\text{MAX}(|\{P : P \in \underline{A} \wedge P(x)\}|, |\{P : P \in \underline{A} \wedge P(y)\}|)} \geq d$$

This gives us a closed scale (from 0 to 1), as desired in view of our discussion from section 2.1.2. For ease of notation, I will write the fraction in (36) as simply $\text{SIM}(x, y, \underline{A})$.

⁷But not identical; we will see that this is important to 'get the proportions right'.

⁸The lexical entry in (36) refers to cardinalities of sets of properties, and never to specific properties.

From this follows the correct prediction that one need not now *which* properties two objects share to know that they are alike.

- (i) #I don't know which color x and y are, but I know they are red.
- (ii) I don't know which color x and y are, but I know that with respect to color x is like y .

See also:

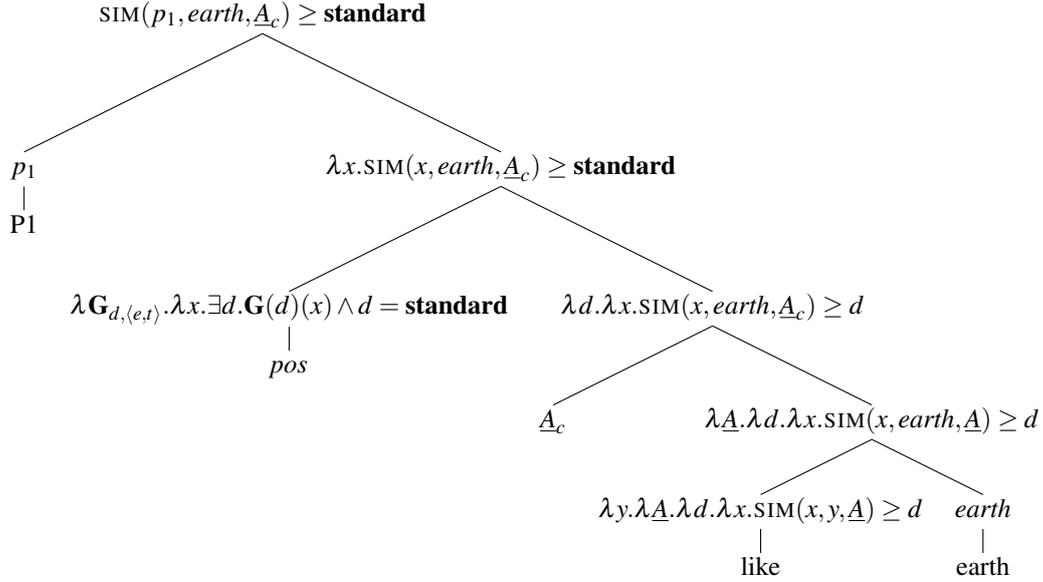
- (iii) Context: A secret agent has to find a suitcase in a house. Neither he nor his agency know the color of the suitcase. They also don't know the color of the entrance door. His boss says:
 - "We don't know the color of the suitcase, but we know from our sources that..."
 - a. "... with respect to color it is like the entrance door."

⁹For ease of readability, in (36) and similar formulas I will mix set notation, as in $P \in \underline{A}$, and functional notation, as in $P(x)$.

$$(37) \quad \text{SIM}(x, y, \underline{A}) \equiv \frac{|\{P : P \in \underline{A} \wedge P(x)\} \cap \{P : P \in \underline{A} \wedge P(y)\}|}{\text{MAX}(|\{P : P \in \underline{A} \wedge P(x)\}|, |\{P : P \in \underline{A} \wedge P(y)\}|)}$$

Compositionally, this takes the following form.¹⁰

(38) P1 is like Earth.



In the structure of (38), I treat \underline{A} in the same way as Von Stechow (1994) and Stanley (2002) treat Domain Restriction variables, i.e. I assume that \underline{A} is represented in the syntax. ‘Like’ comes with a covert ‘relevant attributes’ variable \underline{A} in the object language, whose value is determined by a (contextually determined) assignment function g , i.e. what I, for ease of notation, call \underline{A}_c is simply $g(\underline{A})$.¹¹

As usual, scalar modifiers will apply in lieu of *pos*.

- (39) a. $\llbracket \text{much} \rrbracket = \lambda \mathbf{G}_{d, \langle e, t \rangle}. \lambda x. \exists d. \mathbf{G}(x) \geq d \wedge d = \text{much}$
 b. $\llbracket \text{exactly} \rrbracket = \lambda \mathbf{G}_{d, \langle e, t \rangle}. \lambda x. \exists d. \mathbf{G}(x) \geq d \wedge d = 1$
 c. $\llbracket \text{nothing} \rrbracket = \lambda \mathbf{G}_{d, \langle e, t \rangle}. \lambda x. \exists d. \mathbf{G}(x) \geq d \wedge d = 0$

¹⁰In this derivation, at the node in which *pos* merges with ‘like Earth’, for simplicity I write, instead of (i), the equivalent formula in (ii), and will do so in similar cases throughout the paper.

(i) $\lambda x. \exists d. \text{SIM}(x, \text{earth}, \underline{A}_c) \geq d \wedge d = \text{standard}$

(ii) $\lambda x. \text{SIM}(x, \text{earth}, \underline{A}_c) \geq \text{standard}$

¹¹A variant of this approach captures cases of co-variation of the similarity respects with individuals. Consider the sentence below:

(i) Every student is like a lawyer (in one respect or another).

This sentence has a reading in which the respect in which the given individual is like a lawyer varies for each student. It is possible to accommodate such cases within my system, by assuming that the ‘relevant attributes’ variable is a *functional* variable that represents a function from individuals to such sets, along the lines of what has been proposed for functional domain restrictions in the scope of quantifiers (see, e.g., Partee 1989).

‘99%’ directly denotes a degree, in analogy with what happens in expressions such as ‘190cm tall’.

$$(40) \quad \llbracket 99\% \rrbracket = \frac{99}{100}$$

‘Getting the proportions right’

This makes the correct predictions concerning proportional modifiers. Let us start from a simple example, such as (41).

(41) Centaurs are half like humans (...and half like horses).

If we think, as is intuitive, that the relevant properties are the possible shapes of different parts of the body, then \underline{A} is as follows:

$$(42) \quad A_c = \{\lambda x.human-upper-body(x), \dots, \lambda x.dog-upper-body(x), \dots, \lambda x.horse-upper-body(x), \dots, \lambda x.human-lower-body(x), \dots, \lambda x.dog-lower-body(x), \dots, \lambda x.horse-lower-body(x)\}$$

Then:

$$(43) \quad \text{SIM}(centaur, human, \underline{A}_c) = \frac{|\{\lambda x.human-upper-body(x)\}|}{|\{\lambda x.human-upper-body(x), \lambda x.human-lower-body(x)\}|} = \frac{1}{2}$$

Similarly for our initial DNA example:

(44) The DNA of chimps is 99% like that of humans.

If we think, as is intuitive, that the relevant properties are the possible nucleobases that a species can have in a given nucleotide, then \underline{A} is as follows:

$$(45) \quad A_c = \{\lambda x.adenine-1^{st}-nucleotide(x), \lambda x.guanine-1^{st}-nucleotide(x), \dots, \lambda x.thymine-100^{th}-nucleotide(x)\}$$

Suppose for simplicity that we are looking at a sequence of 100 nucleotides in humans and chimps. Then since each nucleotide can take four possible values (one for each possible nucleobase), there are 400 relevant properties. Since both humans and chimps have each of the nucleotides in the sequence, both of them have 100 of the properties in A_c , so the denominator of $\text{SIM}(chimps-DNA, humans-DNA, \underline{A}_c)$ is 100. Say the two sequences differ in the 37th nucleotide in that humans have guanine and chimps have adenine in it. Then the intersection between $\{P : P \in \underline{A} \wedge P(x)\}$ and $\{P : P \in \underline{A} \wedge P(y)\}$ will contain 99 properties, and so $\text{SIM}(chimps-DNA, humans-DNA, \underline{A}_c) = \frac{99}{100}$, which is what we want.

Crucially, this entry also gives us correct predictions in cases in which one of the two individuals has *more* of the relevant properties with respect to another one. Say that we are comparing two individuals in terms of nationality, and that A is Greek, while B has both Greek and Finnish nationality.

(46) In terms of nationality, A is half like B .

- a. $\underline{A}_c = \{\lambda x.American(x), \dots, \lambda x.Greek(x), \dots, \lambda x.Finnish(x), \dots, \lambda x.Moroccan(x)\}$
b. $SIM(A, B, \underline{A}_c) = \frac{|\{\lambda x.Greek(x)\}|}{|\{\lambda x.Greek(x), \lambda x.Finnish(x)\}|} = \frac{1}{2}$

Capturing the two levels of modification: similarities and differences between scalar modifiers and WRT phrases

At this point, to see that the account makes the right predictions for what concerns scalar modifiers and WRT phrases, it is crucial to note that \underline{A} is an *unstructured* set of properties. This means that, intuitively, it can contain properties pertaining to different respects. For instance, consider a context in which we are comparing the exoplanet P1 to Earth along atmosphere composition, surface composition, and subsurface composition. \underline{A}_c will be as follows:

$$(47) \quad \underline{A}_c = \underbrace{atmosphere-composition-properties}_{\{\lambda x.contains-nitrogen(x), \dots, \lambda x.contains-methane(x)\} \cup \{\lambda x.contains-silicate-rocks(x), \dots, \lambda x.contains-water(x)\} \cup \{\lambda x.contains-iron(x), \dots, \lambda x.contains-nickel(x)\}} \cup \underbrace{surface-composition-properties}_{\{\lambda x.contains-silicate-rocks(x), \dots, \lambda x.contains-water(x)\} \cup \{\lambda x.contains-iron(x), \dots, \lambda x.contains-nickel(x)\}} \cup \underbrace{subsurface-composition-properties}_{\{\lambda x.contains-iron(x), \dots, \lambda x.contains-nickel(x)\}}$$

In this context, an expression with a scalar modifier such “a bit like Earth” will have the meaning in (48):

$$(48) \quad \llbracket \text{a bit like Earth} \rrbracket = \lambda x.SIM(x, earth, \underline{A}_c) \geq \mathbf{a-bit}$$

The formula in (48) only places a constraint on the number of properties for which a planet should have the same value as Earth (0 or 1) *across* \underline{A} . Therefore, importantly, it underspecifies whether modification proceeds within or between dimensions. We therefore correctly predict that (48) holds of the following:

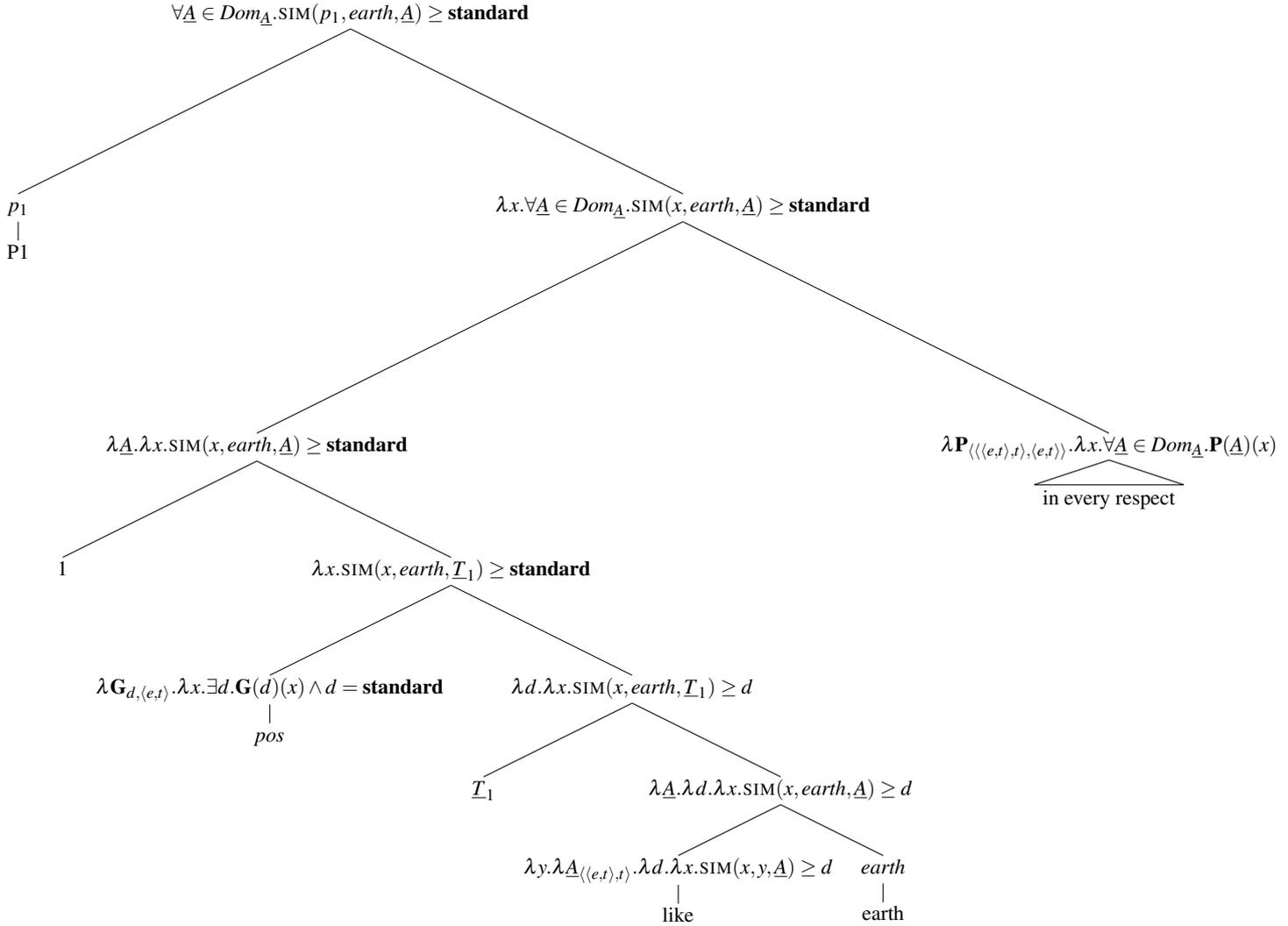
- (i) of planets that slightly overlap with Earth on every relevant dimension, as for instance P1 in the context in (49);
- (ii) of planets that fully overlap with Earth along a few dimensions, as for instance P2 in the context in (49).

(49) **Context:** we are comparing P1 and P2 to Earth, and we are looking at atmosphere composition, surface composition, subsurface composition.

	Earth	P1	P2
Atmosphere Composition	Nitrogen, Oxygen, Argon	Oxygen *partially overlapping*	Xenon, Methane *non-overlapping*
Surface Composition	Silicate rocks, Carbonate Rocks, Water	Water *partially overlapping*	Silicate rocks, Carbonate rocks, Water *fully overlapping*
Subsurface Composition	Iron, Nickel, silicate minerals	Nickel *partially overlapping*	Phosphates *non-overlapping*

Let me now turn to presenting the analysis of WRT phrases. I assume that WRT phrases bind \underline{A} , as in the structure below.

(50) P1 is like Earth in every respect.



Consequently, a ‘like’ phrase such as ‘like Earth in a few respects’ (which we here consider as the case parallel to ‘a bit like Earth’) can only hold of a planet that fully overlaps with Earth along some dimensions, and *not* of a planet that slightly overlaps with Earth along every dimension.¹²

(51) $\llbracket \text{like Earth in a few respects} \rrbracket =$
 $\lambda x. \text{FEW}_{\underline{A}}[\underline{A} \in \text{Dom}_{\underline{A}}][\text{SIM}(x, \text{earth}, \underline{A}) \geq \text{standard}]$

This straightforwardly derives the prediction that in a context like (49), the expression

¹²For simplicity, I assume that we are in a strict context, where it is not enough for a planet to share one out of three properties with Earth on a given dimension for it to be said to be like Earth with respect to that dimension. This does not affect the conclusion: if ‘in a few respects’ could bear on single attributes, it would be felicitous just like *a bit* is, even in a strict context.

in (52) will only be true of P2. For the same reason, this analysis predicts that we cannot use a WRT phrase like ‘in a few respects’ in contexts where only one respect is salient, just like it is hard to say *a few kids returned home* if only one did.^{13,14} Finally, we predict that in presence of a WRT phrase, modification by scalar modifiers should only bear on single respects, since it occurs *under* quantification over the respects by the WRT phrase. We can thus pin down the difference between sentences such as (52a) and (53a).

- (52) a. P1 is like Earth in every respect.
 b. $\forall \underline{A} \in \text{Dom}_{\underline{A}}.\text{SIM}(p_1, \text{earth}, \underline{A}) \geq \text{standard}$
- (53) a. P1 is {exactly, completely} like Earth in every respect.
 b. $\forall \underline{A} \in \text{Dom}_{\underline{A}}.\text{SIM}(p_1, \text{earth}, \underline{A}) = 1$

This captures the fact that in (53), there is necessarily complete overlap along all the dimensions quantified over. In (52a), instead, there is *enough* overlap along all the dimensions quantified over.

2.4 Homogeneity

Consider the context in (54).

(54)

	Earth	P1	P2	P3
Atmosphere Composition	Nitrogen, Oxygen, Argon	Nitrogen, Oxygen, Argon *completely overlapping with Earth*	Oxygen *partially overlapping with Earth*	Xenon, Methane *non-overlapping with Earth*

If we are comparing atmospheres, (55b) and (56a) are clearly false, while (55a) and (56b) are clearly true.

- (55) a. P1 is like Earth. TRUE
 b. P1 isn't like Earth. FALSE
- (56) a. P3 is like Earth. FALSE
 b. P3 isn't like Earth. TRUE

The pattern seems to be the following:

- (57) a. X is like Earth. \rightsquigarrow X and Earth share all relevant properties
 b. X isn't like Earth. \rightsquigarrow X and Earth share no relevant properties

This intuition is confirmed by the fact that exceptives are licensed in both positive and negative sentences involving ‘like’, as we have already observed in section 2.2.

¹³We also correctly predict, of course, that we can say that P2 is like Earth *in one respect*

¹⁴One crucial assumption is that the domain of respects \underline{A} is typically restricted to (unions of) psychologically coherent and salient attributes, like ‘eye color’ or ‘personality’. Of course, exotic similarity respects are possible (‘with respect to whether she goes to the gym on Tuesdays’), if explicitly specified, but will typically not be included in the domain restriction.

- (58) a. X is like Earth, except for its atmosphere.
 b. X isn't like Earth, except for its atmosphere.

Unlike what happens in (55)-(56), for the sentences in (59) it is not clear whether they are true or false, assuming the context described in (54).

- (59) a. P2 is like Earth.
 b. P2 isn't like Earth.

To sum up, for positive sentences to be clearly true, all relevant properties should be shared with Earth. For them to be clearly false, none of them should be. And vice versa for negative sentences. This polarity reversal is strongly reminiscent of the phenomenon known as homogeneity (cf. Löbner 2000). Homogeneity is a property of many expressions in natural language, most notably of definite plurals:

- (60) a. The presents are under the tree.
 ~> All of the presents are under the tree.
 b. The presents aren't under the tree.
 ~> None of the presents are under the tree

Homogenous sentences are known to tolerate exceptions or give rise to non-maximal interpretations (cf. Brisson 1998 and contributions on homogeneity thereafter, e.g. Križ 2015 a.o.). For example, (60a) may be acceptable as an utterance even when two presents aren't under the tree; (60b) may be acceptable even when one or two presents are under the tree. For this to happen, it must be irrelevant for current purposes whether all the presents or most of the presents are under the tree (cf. Malamud 2012, Križ 2015, Križ & Spector 2021, a.o.).

It has been observed that, although definite plurals tolerate exceptions (at least in some contexts), these exceptions cannot be overtly mentioned (irrespective of context) (Križ, 2015; Feinmann, 2020).¹⁵

- (61) a. #The presents are under the tree, but one of them isn't.
 b. #The presents aren't under the tree, but one of them is.

Non-maximality and homogeneity can be removed: there is no polarity reversal in (62), and no non-maximality either.

- (62) a. All of the presents are under the tree.
 b. Not all of the presents are under the tree.

All of these properties are shared by 'like'. It is subject to polarity reversals, as shown in (55) and (56). It displays non-maximality: the sentence '*a* is like *b*' is still acceptable if *a* and *b* do not share one relevant property among many shared ones. However, these exceptions cannot be overtly mentioned: we cannot utter sentences like (63).

- (63) a. #P2 is like Earth, but is a bit unlike Earth.

¹⁵Feinmann (2020) uses this empirical observation as a diagnosis for the presence of homogeneity, which he calls 'Strict Behavior in Contradiction Test'.

Lastly, the homogeneity of ‘like’ can be removed. If we modify ‘like’ with ‘completely’ or ‘exactly’, the ‘slack’ disappears. If we are comparing the atmosphere of P1, P2, and P3 to Earth, (64a) is clearly false, while (64b) and (64c) are clearly true.

- (64) a. P1 isn’t exactly like Earth. FALSE
 b. P2 isn’t exactly like Earth. TRUE
 c. P3 isn’t exactly like Earth. TRUE

This insight can be integrated into the account outlined in the previous section in a number of ways. One is to simply give a trivalent entry for *pos*, within Kriz’s trivalent framework and in analogy with Feinmann’s (2020) treatment of total and partial adjectives, so that we end up with a trivalent entry for the positive form of ‘like Earth’.¹⁶

$$(65) \quad \llbracket pos \text{ like Earth} \rrbracket = \lambda x. \begin{cases} 1 & \text{if } \text{SIM}(x, \text{earth}, \underline{A}_c) = 1 \\ 0 & \text{if } \text{SIM}(x, \text{earth}, \underline{A}_c) = 0 \\ \# & \text{otherwise} \end{cases}$$

This would contrast with the bivalent meaning of ‘completely like Earth’.¹⁷

$$(66) \quad \llbracket \text{exactly like Earth} \rrbracket = \lambda x. \begin{cases} 1 & \text{if } \text{SIM}(x, \text{earth}, \underline{A}_c) = 1 \\ 0 & \text{if } \text{SIM}(x, \text{earth}, \underline{A}_c) \neq 1 \\ \# & \text{never} \end{cases}$$

More recently, Bar-Lev (2021) has argued that homogeneity arises from an implicature. In a nutshell, the meaning of sentences with definite plural subjects is weak, i.e. existential. This straightforwardly derives negative sentences. Positive sentences get strengthened via innocent inclusion of sub-domain alternatives as a result of the application of an exhaustification operator to the sentence, in a way that is completely parallel to the account of free choice in Bar-Lev & Fox (2020). Non-maximality is achieved by assuming that some sub-domain alternatives can be pruned, and thus do not get innocently included. This is again doable here, if we take a weak positive form as in (67):

¹⁶This in principle requires there to be different *pos*’s for different gradable expressions. An alternative to this is to recouch these insights within Feinmann’s similarly trivalent framework, where there is no positive morpheme, and gradable expressions simply have lambda-abstraction over standards. This avoids having to stipulate different *pos*’s for different expressions.

¹⁷A reviewer observes that ‘the same as’ can be used to assess similarity, but seems to be stronger than ‘like’. I think that this is due to the fact that ‘the same as’ is essentially a non-homogenous version of ‘like’. ‘The same as’ can be relativized to relevant dimensions just like ‘like’, but the tests used in this section for homogeneity do not go through. There seems to be no polarity reversal under negation:

- (i) a. *X* is the same as Earth. \rightsquigarrow *X* and Earth share all relevant properties.
 b. *X* is not the same as Earth. $\not\rightsquigarrow$ *X* and Earth share no relevant properties.

Moreover, there seems to be no non-maximality: (ia) may not be acceptable as an utterance if *X* and Earth don’t share one or two of the relevant properties.

In this sense, *the same as* seems to be a bivalent version of ‘like’, as in (ii):

$$(ii) \quad \llbracket \text{the same as Earth} \rrbracket = \begin{cases} 1 & \text{if } \text{SIM}(x, \text{earth}, \underline{A}_c) = 1 \\ 0 & \text{otherwise} \end{cases}$$

This correctly predicts (a) that ‘like’ follows from ‘the same as’ (but not vice versa) and (b) that ‘exactly like *X*’ and ‘the same as *X*’ should be equivalent if we zero in on the same dimensions, which is intuitively the case.

$$(67) \quad \llbracket \text{pos like Earth} \rrbracket = \lambda x. \exists d \geq 0. \text{SIM}(x, \text{earth}, \underline{A}_c) \geq d$$

Including all subdomain alternatives among the degrees quantified over has the effect of strengthening (67) to mean that $\text{SIM}(x, \text{earth}, \underline{A}_c)$ is equal to 1. This would also predict that there can be non-maximality via pruning of some degrees in the domain.

To sum up, ‘like’ appears to induce homogeneity, as shown by the data I presented on polarity reversal, licensing of exceptives, non-maximality, unmentionability of exceptions, and slack removal. While I do not take a stance on which account of homogeneity should be adopted, I have shown that the theory developed in the previous section is in principle compatible with both Kriz’s trivalent framework and Bar-Lev’s implicature-based theory.

Before moving on, I would like to observe that the arguments given in this section and in section 2.1.2 also apply to clausal ‘like’:

- (68)
- a. **Proportional modification.**
The DNA of humans is 99% like that of chimps is.
 - b. **Two levels of modification.**
P1 is much like Earth is. {within, between}
P1 is like Earth is in many respects. {within, #between}
P1 is much^{within} like Earth is in many respects^{between}.
 - c. **Homogeneity.**
 - (i) Polarity reversal:
P1 is like Earth is. \rightsquigarrow P1 and Earth share all relevant properties
P1 isn’t like Earth is. \rightsquigarrow P1 and Earth share no relevant properties
 - (ii) Homogeneity removal:
P1 is exactly like Earth is. \rightsquigarrow P1 and Earth share all rel. prop.
P1 isn’t exactly like Earth is. $\not\rightsquigarrow$ P1 and Earth share no rel. prop.
 - (iii) Licensing of exceptives:
P1 is like Earth is, except for its atmosphere.
P1 isn’t like Earth is, except for its atmosphere.

As we have seen, in Alrenga’s uniform analysis, clausal ‘like’ compares the sets of properties Q such that San Francisco is Q with the set of properties P such that Palo Alto used to be P . Alrenga assumes the property-type gaps in these complements arise due to wh-movement of a null operator, whose semantic function is to yield the set of properties that satisfy the open proposition it combines with.

$$(69) \quad \text{San Francisco is [like } [Op_i \text{ Palo Alto used to (be } t_i)]]$$

- a. $\exists P [P \in \{Q : \text{San Francisco is } Q\} \& P \in \{Q : \text{Palo Alto used to be } Q\}]$

As stated, the existential force of (69a) is unlikely to capture the facts in (68), for the same reasons for which an extension of Alrenga’s analysis to phrasal ‘like’ fails. This paper is not concerned with clausal ‘like’, so what follows should not be seen as more than a speculation, but I think there are related ways to amend Alrenga’s analysis of clausal ‘like’.

We have seen in section 2.1.1 that a syntactically and semantically uniform analysis of

(descriptively) clausal ‘like’ and (descriptively) phrasal ‘like’ is unlikely to succeed. However, (68) suggests a common core to the meanings of preposition ‘like’ and coordinator ‘like’. The entry proposed in my analysis in (36) for phrasal ‘like’ compares the values of individuals with respect to a relevant set of properties. One way of looking at clausal ‘like’ may be that it *directly checks overlaps between its input sets of properties*:

$$(70) \quad \llbracket \text{like}_{\text{clausal}} \rrbracket = \lambda \underline{Q}_{\langle e,t \rangle,t} \cdot \lambda \underline{A}_{\langle e,t \rangle,t} \cdot \lambda d \cdot \lambda \underline{P}_{\langle e,t \rangle,t} \cdot \frac{|\underline{P} \cap \underline{Q} \cap \underline{A}|}{\text{MAX}(|\underline{P} \cap \underline{A}|, |\underline{Q} \cap \underline{A}|)} \geq d$$

- (71) *(Looking at body shape...)*
 a. This centaur is half like that man is.

$$\frac{|\{P_{e,t} : \text{this centaur has } P\} \cap \{Q_{e,t} : \text{that man has } Q\} \cap \underline{A}|}{\text{MAX}(|\{P_{e,t} : \text{this centaur has } P\} \cap \underline{A}|, |\{Q_{e,t} : \text{that man has } Q\} \cap \underline{A}|)} = \frac{|\{\lambda x.\text{human-upper-body}\}|}{|\{\lambda x.\text{human-upper-body}(x), \lambda x.\text{horse-lower-body}(x)\}|} \geq \frac{1}{2}$$

With $\underline{A} = \{\lambda x.\text{human-upper-body}(x), \lambda x.\text{dog-upper-body}(x), \lambda x.\text{horse-upper-body}(x), \dots, \lambda x.\text{horse-lower-body}(x)\}$

To integrate facts concerning homogeneity, we can give a trivalent Križ-style entry parallel to the one proposed in 2.4 for phrasal ‘like’.¹⁸

$$(72) \quad \llbracket \text{pro like that centaur is} \rrbracket = \lambda \underline{P} \cdot \begin{cases} 1 & \text{if } \text{SIM}_{\text{prop.}}(\underline{P}, \{Q_{e,t} : \text{that centaur is } Q\}, \underline{A}_c) = 1 \\ 0 & \text{if } \text{SIM}_{\text{prop.}}(\underline{P}, \{Q_{e,t} : \text{that centaur is } Q\}, \underline{A}_c) = 0 \\ \# & \text{otherwise} \end{cases}$$

3 ‘Look like Mary’

3.1 Distinguishing similarity from appearance

Notice that similarity talk in itself is distinct from talk about appearances or experiences: there is such a thing as objective property sharing. The sentences in (73) are parallel to those in (74).

- (73) a. With respect to personality, John seems like Sue, but he isn’t like Sue.
 b. Marriage seems like heaven, but once you’re in it, it isn’t like heaven.

- (74) a. John seems French, but he isn’t French.
 b. Marriage seems nice, but once you’re in it, it isn’t nice.

A point that militates in favor of the non-experiential nature of similarity is that the ‘to’-PP is an argument of subjective predicational copulas like ‘look’, but not of simple predicational copulas like ‘be’. For one, (75a) sounds more natural than (75b).¹⁹

¹⁸And similarly to the case of phrasal comparatives, for simplicity I call the ratio in (70) $\text{SIM}_{\text{prop.}}(\underline{P}, \underline{Q}, \underline{A})$

¹⁹Of course, relative truth is always an option. But notice that (75b) is good to the same extent as (i) is good:

(i) ? To John, Mary is a lawyer.

- (75) a. Mary looks like Tom to John.
b. ? Mary is like Tom to John.

A second marker of argumenthood is the selection of specific prepositions (Rudolph, 2019; Stephenson, 2007). ‘Proud’, for instance, requires ‘of’ PPs, while ‘pride’ requires ‘in’ PPs:

- (76) a. I am proud of / # in Mary.
b. I take pride # of / in Mary.

In this sense, ‘be like’ has a much less specific requirement than ‘look like’:

- (77) a. Mary looks like Tom to/#for John.
b. Mary is like Tom ? to / ? for John.

This suggests that, though ‘look like x_e ’ constructions are both modal and subjective, ‘like’ PPs in themselves do not introduce the subjectivity typical of experiential predicates.

Moreover, as pointed out by a reviewer, only appearance predicates can take clausal complements (as in ‘he looks/seems like he’s a rich guy’), whereas ‘be like’ cannot (as in ‘*he’s like he’s a rich guy’).

3.2 Similarity embedded by appearance: an illustration

What are the truth conditions of a ‘look like’ sentence? In her dissertation, Rudolph captures predicate-embedding appearance/experiential predicates, e.g. ‘look French’, with the semantics in (78).²⁰ While nothing in my analysis hinges on Rudolph’s specific analysis of appearance verbs, I will use her account to illustrate how ‘like’ PPs can be embedded in copular verbs other than ‘to be’.

$$(78) \quad \llbracket \text{seem} \rrbracket^w = \lambda P. \lambda x. \forall w' \in B_p(j, w). P(x)(w')$$

Where $B_p(x, w)$ are the best/most typical worlds compatible with x ’s perceptual evidence.²¹ Of course, different appearance verbs like ‘look’, ‘sound’, ‘smell’ (and so on) prompt different accessibility relations that refer to the worlds compatible with, respectively, visual, auditory, and olfactory evidence.

$$(79) \quad \llbracket \text{Mary looks French.} \rrbracket^w = \forall w' \in V(j, w). \text{French}(\text{Mary})(w')$$

Keeping in mind that we analyzed ‘like Mary’ as property-denoting, I argue that, as a first approximation, something entirely parallel to (79) happens in ‘look like x_e ’ constructions.

²⁰For ease of notation, instead of lambda-abstracting the judge argument, I leave it as a free variable here.

²¹More precisely:

- (i) a. **Typicality ordering:** For set of worlds X and set of “typical” propositions $T, \forall w, w' \in X, w <_T w'$ iff $\{p \in T : p(w') = 1\} \subset \{p \in T : p(w) = 1\}$
b. **Best visually accessible worlds:** Where $V(x, w)$ is the set of worlds left open by x ’s perceptual experience at $w, B_v(x, w) := \{w' \in V(x, w) : \neg \exists w'' \in v[w'' < T w']\}$

- (80) $\llbracket \text{Bob looks like Carl.} \rrbracket^w =$
 $\lambda j. \forall w'. w \in V(j, w). \text{SIM}(\text{bob}, \text{carl}, \underline{A}) \geq \text{standard}$

In words, (80) is true iff at all best/most typical worlds compatible with the judge j 's visual perception, Bob has the same value as Carl for relevant attributes. Rudolph's treatment of appearance verbs would also allow us to account for cases of similarity relativized to a specific judge:

- (81) a. $\llbracket \text{Bob looks like Mary to John.} \rrbracket^w =$
 $\forall w'. w' \in B(\text{john}, w). \text{SIM}(\text{bob}, \text{mary}, \underline{A}) \geq \text{standard}$
 b. 'at all best/most typical worlds compatible with the John's visual perception, Bob has the same value as Carl for relevant attributes'

3.3 Phenomenal versus epistemic embeddings of 'like' PPs

Rudolph captures experiential predicates like 'looks' as universally quantifying over worlds compatible with one's experience, viz. visual experience in this case, both when they embed a proposition and when they embed a predicate.

- (82) a. PREDICATE-EMBEDDING $\llbracket \text{look} \rrbracket^w = \lambda P. \lambda x. \forall w' \in V(j, w). P(x)(w')$
 b. PROPOSITION-EMBEDDING $\llbracket \text{look} \rrbracket^w = \lambda p_t. \forall w' \in V(j, w). p(w')$

This predicts that (80), repeated below in (83a), ends up having the same truth conditions as (83b).

- (83) a. Bob looks like Mary.
 b. Bob looks like he is like Mary.

This is an unwelcome prediction: while in (83a) Bob's appearance directly resembles Mary's appearance, (83b) is compatible with indirect visual evidence pointing to a similarity between Bob and Mary.

Rudolph notices a similar prediction of her account, namely that (84a) and (84b) are equivalent.

- (84) a. Mary looks French.
 b. Mary looks like she is French.

While Rudolph takes this to be an appropriate prediction, it seems that (84a) and (84b) are not entirely equivalent: in the context in (85), (85a) seems appropriate, while (85b) doesn't.

- (85) It is very clear that Bob is not French, and much indirect visual evidence points to him not being French: for instance, at some point he dropped his passport and we saw it was a German one. Suppose now that he puts on a Basque beret.
 a. Bob looks French.
 b. ?? Bob looks like he's French.

This seems to pertain to a distinction between *epistemic* and *phenomenal* uses of appearance verbs. Rudolph, while denying this distinction to hold for ‘look’, takes it to hold for verbs like ‘taste’ and ‘sound’. To illustrate, in the context in (86), the phenomenal use of ‘taste’ forced in (86a) is weird, while the epistemic use forced in (86b) is appropriate.

- (86) Context: Bob doesn’t like good quality wine, but he’s a very educated oenologist and can easily tell from taste when a wine is of good quality. He tastes a wine and says:
- a. ?? This wine tastes good.
 - b. This wine tastes like it’s good.

This difference carries over to sentences about similarity like (80) and (83b). To see this consider (87a) and (87b):

- (87) Context: As in (85).
- a. Bob looks like a Frenchman.
 - b. ?? Bob looks like he is a Frenchman.

Sentence (87b) is not felicitous because suggests that Bob may be a Frenchman; instead, (87a) implies that Bob isn’t a Frenchman.

Note that there are purely epistemic uses of ‘look like x_e ’ and ‘look like a N ’ constructions, as in (88a). And indeed, they end up being equivalent to their propositional versions: (88a) is felt to be equivalent to (88b):

- (88) Context: I left my wallet at Mary’s when I went to her party. The day after, she sends a friend, John, whom I don’t know, to give it back to me. She describes her friend to me on the phone. Someone rings my bell, I open the door and say:
- a. ‘You look like John!’
 - b. ‘You look like you are John!’

Then, it seems that while the epistemic use of ‘look like’ when embedding a ‘like’ PP can be treated on a par with sentence-embedding ‘look like’, its phenomenal use cannot.

This paper is about similarity, and not about appearance predicates; I will not offer a complete solution to this. But I want to sketch one simple way out, namely to cash out the two uses as simply deriving from two different accessibility relations, a possibility that Rudolph (2019) briefly mentions, too. I would suggest that epistemic uses of appearance verbs like ‘look’ feature an accessibility relation that returns the worlds (i) compatible with one’s perception that are (ii) *candidates* for the actual world. One may worry that this brings appearance verbs too close to belief attitudes. However, candidate worlds *compatible with one’s perception* are not necessarily candidates *tout court*, thus compatible with beliefs: beliefs are formed and held on other grounds beyond perception. I would suggest that such an epistemic-experiential accessibility relation is the only one available for sentence-embedding ‘look like’, but not for predicate-embedding ‘look like’, which is ambiguous between the epistemic and the phenomenal accessibility relation.

The phenomenal uses of appearance verbs would instead feature an accessibility relation that returns those worlds that make true most propositions compatible with one's perception, *regardless* of whether they are candidates for the actual world. This explains why someone can look French against all evidence of them actually *being* French. The same is not true for someone who looks like *they are* French, in which case there must be evidence pointing to them being French.

4 'Be like a lawyer'

4.1 Similarities and differences between characterizing sentences and 'like a *N*' constructions

Similarities

Similarity statements of the form 'be like a *N*' have two different readings. I will call them the specific and the general one.

- (89) John looks like a lawyer.
- a. SPECIFIC: There is a specific lawyer such that John looks like them.
 - b. GENERAL: John has the general appearance of a lawyer.

The general reading displays a number of logical and compositional properties that align with generic indefinites in characterizing sentences.

(I) Non-increasingness. Suppose we are speaking about John, who is a notary in France, and that you ask me what tasks his job involves concretely. I answer:

- (90) In many respects, he is like a British lawyer.

In this context, (91) does not follow from (90):

- (91) $\not\equiv$ In many respects, he is like a lawyer.

If, in the LF of (90), 'British lawyer' is interpreted in the restrictor of a generic quantifier, this behaviour would be explained.

- (92) $\text{GEN}_x[x \text{ is a British lawyer}][\text{John shares relevant properties with } x]$

I use GEN mostly as a black box, since its interpretation is a very broad and debated issue in itself.²² I will nevertheless refer to specific *desiderata* for an interpretation of GEN. For instance, the logical form in (92) predicts non-increasingness because the interpretation of GEN must ensure non-monotonicity independently (Krifka *et al.*, 1995; Asher & Morreau, 1995) This is because as is well-known, characterizing sentences are non-increasing: (93b) doesn't follow from (93a):

²²See for instance the introduction to Mari *et al.* 2012 for an extensive literature review.

- (93) a. A british judge wears a wig.
b. A judge wears a wig.

Sentence (90) can then be paraphrased as follows:

- (94) In general, if someone is a British lawyer, John shares relevant properties with them.

(II) Narrow readings of disjunction. It is known that generic sentences give rise to narrow scope readings of disjunction. There is a reading of the characterizing sentence in (95) which means that lawyers *and* judges generally wear a tie, suggesting that the disjunction is interpreted in the restriction of GEN.

- (95) A lawyer or judge wears a tie.

‘Like a *N*’ PPs behave similarly:

- (96) In terms of clothing, John is like a lawyer or a judge.

The narrow scope reading of disjunction, unlike the wide scope reading, implies that John dresses in a way in which *both* a lawyer and a judge dress - and the sentence is roughly equivalent to a conjunction of two similarity statements:

- (97) In terms of clothing, John looks like a lawyer and John looks like a judge.

This is straightforwardly explained if the argument of ‘like’ is interpreted in the restrictor of a generic quantifier:

- (98) GEN_x[*x* is a lawyer or *x* is a judge][John looks like *x*]

(III) Subtriggering pattern. In regular generic sentences, unmodified ‘someone’ can only be read existentially; modified ‘someone’ can be read both existentially and generically (Carlson, 1981; Dayal, 2004; Mascarenhas, 2012).

- (99) Someone is punctual.
a. Existential
b. *Generic
- (100) Someone who respects others is punctual
a. Existential
b. Generic

‘Like’ also licenses generic readings of ‘someone’ only in presence of modification:

- (101) He looks like someone.
a. Existential
b. *Generic
- (102) He looks like someone who respects others.
a. Existential

- There is someone who respects others such that he looks like them.*
- b. Generic
He looks like a typical others-respecting person.

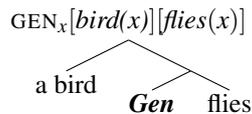
While there is no generally accepted explanation of subtriggering patterns, all proposals cash them out as deriving from specific constraints on what may or may not restrict GEN. If this is indeed the case, then it is completely expected that ‘like’ would behave the way it does if it is lexically generic.

(IV) ‘a *N*’ ≈ ‘a typical *N*’. We can substitute ‘a typical *N*’ for ‘a *N*’ without changing the meaning by much. This is a variation on a test commonly used since Krifka *et al.* (1995).

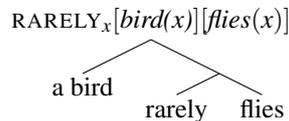
- (103) a. In terms of culinary taste, John is just like an Italian.
 b. ≈ In terms of culinary taste, John is just like a typical Italian.

Differences

Standard theories take GEN to be brought about by a silent quantificational adverb **Gen** that comes with verbal aspect. **Gen** puts material it c-commands in the scope of GEN, and material that c-commands it in the restriction of **Gen** (Chierchia, 1998).



When an overt quantificational adverb is present, it takes the place of **Gen**:



For this reason, the classical test to know whether an indefinite is in the restrictor of GEN in a sentence is to see whether it can be bound by an overt quantificational adverb in a parallel sentence. This is the case for ‘vanilla’ characterizing sentences, but not for sentences involving ‘like’.

- (104) A bird flies.
 ≈ Generally, if something is a bird, it flies.
- (105) A bird rarely flies.
 ≈ Rarely, if something is a bird, it flies.
- (106) John looks like a lawyer.
 ≈ Generally, if someone is a lawyer, John looks like them.
- (107) John rarely looks like a lawyer.
 ≠ Rarely, if someone is a lawyer, John looks like them.

Summing up, there are two apparently contradictory facts to explain:

- (108) (i) **(Differences)**
Overt Q-adverbs can bind indefinites found in characterizing sentences.
Overt Q-adverbs *cannot* bind arguments of ‘like’, despite their flexibility in terms of restrictor-scope splitting (as we have seen in the ‘computer-modern plane’ data in (3)-(4)).
- (ii) **(Similarities)**
GEN can bind indefinites found in characterizing sentences.
GEN can bind indefinites that are arguments of ‘like’.

In the remainder of the paper, I show how (i) and (ii) can be reconciled.

Concerning (108)-(i), I show that sentences involving a verb + a ‘like’ phrase have a rigid Topic-Comment structure, causing the argument of ‘like’ to be systematically a Comment. We know independently that material can enter the restriction of Q-adverbs only if it is a Topic (Krifka *et al.*, 1995; de Swart, 1996; Chierchia, 1995, 2009). As a result, we expect that quantificational adverbs should not bind arguments of ‘like’.

But then how come the indefinite in ‘like a lawyer’ can receive a generic interpretation? I will propose, to explain (108)-(ii), that GEN is not necessarily brought about by the silent quantificational adverb *Gen*, and can instead be hard-wired in the lexical entry of lexical items.

This fact does not only have interesting consequences for theories of genericity, but also generates an additional correct prediction pertaining to proper names. Narrow, conjunctive readings of disjunction appear, besides with indefinites, with proper names, too.

- (109) a. John looks like Bob or Mary.

This is once again predicted very straightforwardly if we suppose that the disjunction is interpreted in the restriction of GEN.

- (110) $GEN_x[x \text{ is Bob or Mary}][\text{John looks like } x]$

In what follows, I address (108)-(i) in 4.2, and (108)-(ii) in section 4.3.

4.2 Explaining the differences: the rigid Topic-Comment structure of sentences with ‘like’

Intuitively, a Topic amounts to the entity already introduced in discourse that constitutes the subject matter, and the Comment provides new information about it (see for instance Krifka 2008). I propose that the complement of ‘like’ is necessarily the Comment, and thus systematically provides new information about the Topic. This is quite evident in sentences such as (111): the complement of ‘like’ serves as an ‘anchor’ determining which properties are being compared.

- (111) Context: John has a red nose, Bob has yellow ears.
- a. John is like Bob. \rightsquigarrow John has yellow ears.
 - b. Bob is like John. \rightsquigarrow Bob has a red nose.
 - c. Bob and John are alike. \rightsquigarrow Bob and John have a red nose and yellow ears.

Discourse-linking

Despite lacking a systematic definition in the literature (Krifka, 2008), Topics are held to have two quite precise semantic properties (Jäger, 2001). First, they are discourse linked²³ in a broad sense, i.e. possibly via bridging, as shown in Jäger (2001).

- (112) John has a cottage. {The roof/It} is made from straw.

The idea is that the roof mentioned is not introduced before in discourse, so we cannot e.g. use an anaphoric pronoun to refer to it: ‘It is made from straw’ can only mean the whole cottage is made from straw, unlike what happens with ‘the roof is made from straw’. We can use the topical definite NP since we can establish a salient bridging relation with the cottage introduced in the previous sentence. Bridging works well with objects, too, as in (113).

- (113) John bought a run-down cottage. I’ve been working tirelessly on the roof.

Instead, if we look at sentences with ‘like’, the subject can be discourse-linked via bridging, but not the complement of ‘like’, as shown in (114).

- (114) a. John bought a run-down cottage. The roof looks like my garden canopy.
 b. John bought a run-down cottage. ?? My garden canopy looks like the roof.

Scrambling in German

Second, topics can scramble to the left of the left boundary of the VP in scrambling languages like German (Kratzer, 1995; Jäger, 2001).

Expressions that come to the left of discourse particles like ‘ja’, which marks the left boundary of the VP, are obligatorily interpreted as topics: for instance, (115a) is athetic statement, which does not require that the referent for the mountains has been introduced in discourse beforehand. Sentence (115b), instead, is a categorical statement (cf. Kuroda 1972), and concerns obligatorily a set of mountains previously introduced in discourse (Jäger, 2001).

- (115) a. **Categorical (subject topic)**
 (weil) die Berge ja sichtbar sind.
 (because) the mountains PRT visible are.
 ‘(because) the mountains are visible.’
- b. **Thetic**
 (weil) ja die Berge sichtbar sind.
 (because) PRT the mountains visible are.

²³Not in the modern sense of D-linking.

‘(because) the mountains are visible.’

Predictions vis-à-vis quantificational adverbs

Getting to Topics and quantificational adverbs, the minimal pair in (116b)-(116c) (Krifka, 2001) shows that scrambled indefinites must be bound by ‘gewöhnlich’ (‘often’), while indefinites that stay *in situ* can be closed existentially. Here ‘gewöhnlich’ plays the role of ‘ja’ in (115): it marks the left boundary of the VP since, like ‘ja’, it is adjoined to VP. If both occur to the right of ‘gewöhnlich’, the sentence is ambiguous, as shown in (116a). This constitutes a classical argument for the idea that quantificational adverbs put material in the Topic into their restrictor.

- (116) a. weil gewöhnlich einer alten Dame eine Katze gehört
because usually an old lady(DAT) a cat(NOM) belongs
(i) *‘(because) most old ladies own a cat’*
(ii) *‘(because) most cats belong to an old lady’*
b. weil einer alten Dame gewöhnlich eine Katze gehört
because an old lady(DAT) usually a cat(NOM) belongs
(because) most old ladies own a cat
c. weil eine Katze gewöhnlich einer alten Dame gehört
because a cat(NOM) usually an old lady(DAT) belongs
(because) most cats belong to an old lady

The hypothesis that expressions involving ‘like’/‘wie’ cannot be in the Topic leads to two correct predictions.

- (a) ‘Wie’ should not display meaning alternations like (116a). This is confirmed by (117a).
(b) ‘Wie’ should not be grammatical if it occurs on the left of ‘gewöhnlich’. This is confirmed by the ungrammaticality of (117c).
- (117) a. (weil) gewöhnlich eine alte Dame wie eine Katze aussieht.
(because) usually an old lady like a cat looks.
(i) *(because) most old ladies look like a cat.*
(ii) **(because some old lady looks like most cats)*
b. weil eine alte Dame gewöhnlich wie eine Katze aussieht
because an old lady usually like a cat looks
(because) most old ladies look like a cat’
c. **(weil) wie eine Katze gewöhnlich eine alte Dame aussieht.*
(because) like a cat usually an old lady looks.
(Intended): *(because) some old lady looks like most cats.*

And more generally, the German expression equivalent to ‘like’ cannot appear to the left of ‘ja’ if the sentence is uttered with neutral intonation.

- (118) a. **Thetic**
(weil) ja die alte Dame wie diese Katze aussieht.
(because) PRT the old lady like this cat looks.
(because) the old lady looks like this cat.
b. **Categorical**

(weil) die alte Dame ja wie diese Katze aussieht.
 (because) the old lady PRT like this cat looks.
(because) the old lady looks like this cat.

c. **Categorical (intended)**

*(weil) wie diese Katze ja die alte Dame aussieht.
 (because) like this cat PRT the old lady looks.
(because) the old lady looks like this cat.

English does not scramble, but we can test the ambiguity of sentences such as (116a) and (117a). Sentences like (119) and (120) show that English is no different from German in terms of the scope-splitting of Q-adverbs. If (119) is uttered with neutral intonation, depending on the Topic-Comment division we have meaning (119a) or (119b). This is not so in (120), where the subject, but not the complement of ‘like’ can be bound by ‘usually’.

- (119) An old lady usually owns a cat.
 a. Most old ladies own a cat (some cat).
 b. Most cats are owned by an old lady (some old lady).

- (120) An old lady usually looks like a cat.
 a. Most old ladies look like a cat (some cat).
 b. *An old lady (some old lady) looks like most cats.

A further correct prediction we make is that an item which is (i) similar in meaning to ‘like’ but which (ii) does not display the information-structural rigidity of ‘like’ should allow its arguments to be bound by overt quantificational adverbs. This is the case of ‘alike’: we have seen in (111c), repeated below in (121), that it does not display the inferential asymmetry of ‘like’:

- (121) Context: John has a red nose, Bob has yellow ears.
 Bob and John are alike. \rightsquigarrow Bob and John have a red nose and yellow ears.

And indeed, ‘often’ can bind arguments of ‘alike’.²⁴

- (122) A judge and a lawyer often look alike.
 OFTEN_{x,y}[lawyer(x) ∧ judge(y)][alike(x,y)]
‘Many judge-lawyer pairs are such that the judge and the lawyer look alike’

Topic/Comment is not Background/Focus

At this point, it is important to note that Topic/Comment and Focus/Background are

²⁴The fact that both are in the restriction of the Q-adverb is expected since they are both in the topic (and cannot be separated since they are conjoined). This can be seen by looking at German where a conjunction equivalent to ‘a judge and a lawyer’ can move to the left of ‘ja’, but cannot be split across ‘ja’.

- (i) weil sich ein Richter und ein Anwalt ja ähnlich sind.
 because REFL a judge and a lawyer . PRT alike are.
‘...because a judge and a lawyer are alike.’
- (ii) *weil sich ein Richter ja und ein Anwalt ähnlich sind.
 because REFL a judge PRT and a lawyer alike are.

two orthogonal distinctions.²⁵ Thus noting that the complement of ‘like’ can be both focused and backgrounded, as in (123), would not constitute evidence against the claim that it is not a Topic:

- (123) a. (Who is like John?)
 MARY is like John.
 b. (Who is Mary like?)
 Mary is like JOHN.

This is because Topic and Focus are orthogonal aspects of information structure, and neither can be defined in terms of the other. The main argument for this is the following. We know from examples such as (3) in the Introduction that quantificational adverbs put Topics in their restrictor. However, elements restricting a quantificational adverb can be focused.

- (124) - Who is altruistic?
 - [_{T/F}A fireman] is altruistic. (Jäger, 2001)

Therefore, we are led to conclude that the Topic and the Focus can coincide. From this it follows that Topic/Comment and Background/Focus are orthogonal.

Other arguments that Topic/Comment and Background/Focus are distinct notions has been given by both Jäger (2001) and Krifka (2008). Both argue that nearly all logically possible configurations between the topical and focused elements within a sentence are possible:

- (125) (From Jäger (2001), pp. 110-111:)
- a. *Complementary:*
 - What is this?
 - [_T This][_Fis a bike]
 - b. *Mutually exclusive but not complementary:*
 - Which languages do you know?
 - [_T I] know [_F Dutch]
 - c. *Coinciding:*
 - Who is altruistic?
 - [_{T/F} A fireman] is altruistic.
 - d. *Focus included in Topic:*
 - Which fireman is altruistic?
 - [_T The [_F old] fireman] is altruistic.

Krifka (2008) also mentions the fact that even Focus and Comment need not coincide.

- (126) (From Krifka (2008), p. 266:)
- a. - When did Aristotle Onassis marry Jacqueline Kennedy?
 - [He]_{Topic} [married her [in 1968]_{Focus}]]_{Comment}.

To sum up, in this discussion we have seen that a number of facts point to the fact that the ‘like’ phrase is systematically the Comment: the unavailability of discourse-

²⁵I thank an anonymous reviewer for encouraging me to clarify this.

linking and the ungrammaticality of scrambling of ‘wie’ phrases in German. This led us to correctly predict that sentences involving ‘like’ and ‘wie’ should not display alternations displayed by regular characterizing sentences involving two indefinites. More generally, this led us to correctly predict that quantificational adverbs should not bind arguments of ‘like’ (and ‘wie’). Finally, I have discussed Jäger and Krifka’s motivations to take the Topic-Comment distinction to be entirely orthogonal to the Focus-Background distinction.

4.3 Explaining the similarities: lexically inherent genericity

Let us now turn to explaining (108)-(ii), that is, why arguments of ‘like’ can be bound by GEN despite not being bound by overt quantificational adverbs. I propose that is the result of the lexical semantics of ‘like’.

$$(127) \quad \llbracket \text{like} \rrbracket = \lambda \mathbf{Q}_{\langle e,t \rangle, t} . \lambda A . \lambda d . \lambda x . \text{GEN}_y [\text{BE}(\mathbf{Q})(y)] [\text{SIM}(x, y, A) \geq d]$$

Where:

$$\text{BE} = \lambda \mathbf{Q}_{\langle e,t \rangle, t} . \lambda x . \mathbf{Q}(\lambda y . y = x) \quad \text{Partee (2002)}$$

Partee proposed BE as a natural type-shifting functor that ‘applies to a generalized quantifier, finds all the singletons therein, and collects their elements in a set’. If one analyzes indefinites as generalized quantifiers, BE can also be seen as the operation that the copular verb does to turn the generalized quantifier into a property. One nice property of this account, then, is that it captures the copular nature of similarity.

This correctly predicts generic readings of indefinite complements of ‘like’ without necessarily predicting that overt quantificational adverbs should bind them. Additionally, this view spontaneously derives a desirable prediction, namely that narrow readings of disjunction seem to arise not only with indefinites, but also with individuals:

- (128) With respect to personality, John is like Bob or Mary.
- a. With respect to personality, John is like Bob or with respect to personality, John is like Mary.
 - b. With respect to personality, John is like [Bob or Mary].
 \models With respect to personality, John is like Bob and Mary.
 \rightsquigarrow Bob and Mary have the same kind of personality.

This can only be accounted for by giving truth conditions that correspond to a rough paraphrase like (129), which corresponds to the proposed revision of the theory:

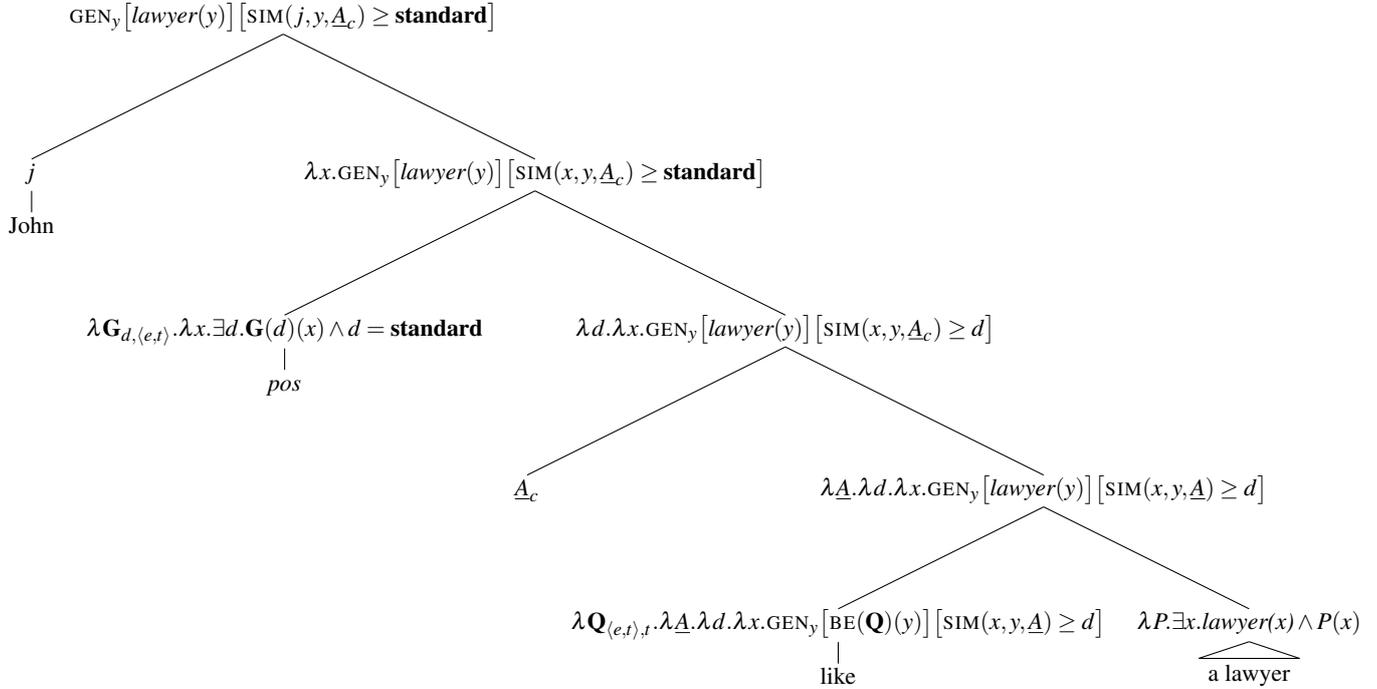
- (129) Typically, if someone has the property of being Mary or of being John, then John shares personality-relevant properties with them.

We will therefore treat proper names as generalized quantifiers, i.e. Montagovian individuals.

The general reading

To illustrate, let us start from the general reading of (130).

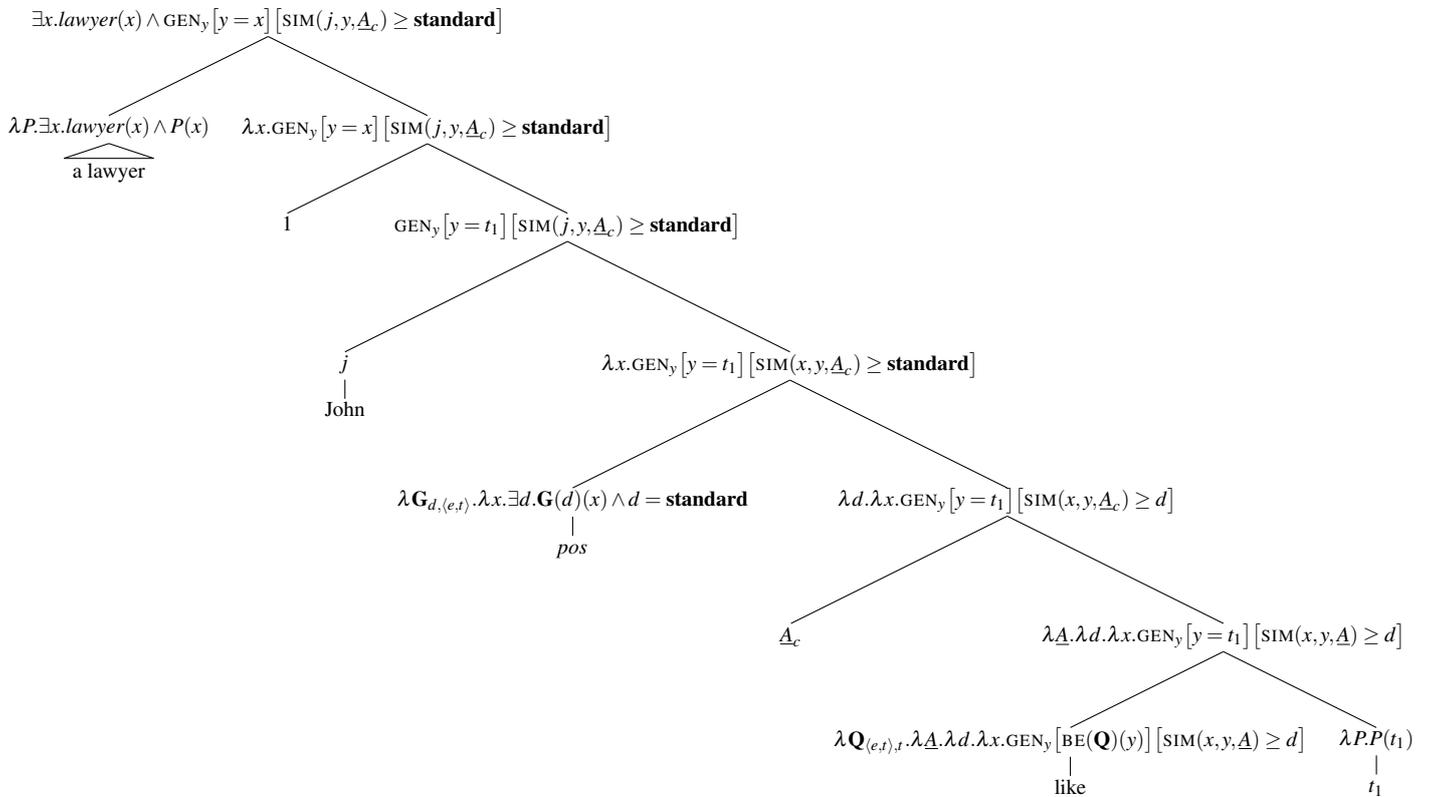
- (130) John is like a lawyer.
 a. General reading: John is like a typical lawyer.



The specific reading

The specific reading can be derived by taking the indefinite to scope out leaving a Montagovian trace.

- (131) John is like a lawyer.
 a. Specific reading: there is a lawyer such that John looks like them.



At this point, we should get back to the question of overt Q-adverbs.

- (132) John is often like a lawyer.
 $\not\approx$ John is like many lawyers.

One might think that the complement of ‘like’ cannot be bound by OFTEN because it is already bound by the lexically encoded GEN – and not because of the rigid information structure displayed by ‘like’.

However, traces can be interpreted inside the restriction of the GEN that comes lexically with ‘like’, for instance in the specific reading. illustrated above in (131). Thus an LF like the one in (133) would be in principle possible:

- (133) $OFTEN_x[lawyer(x)] [GEN_y[x = y] [SIM(john, y, \underline{A}_c) \geq \mathbf{standard}]]$

We should thus hold on to the conclusion of section 4.2 that what rules out the LF in (133) is that Q-adverbs bind Topics, while the complement of ‘like’ is systematically a Comment.

Disjunction

Taking disjunction to be type-flexible (cf. Winter 2002, a.o.), this approach also naturally predicts narrow readings of disjunction when the disjunction of generalized

quantifiers ‘a lawyer or a judge’ or ‘John or Mary’ is interpreted low, inside the scope of ‘like’. The wide reading is derived similarly to how it is derived in (133), by having the wide disjunction bind a Montagovian trace. This yields the wide reading that gives rise to a scalar implicature, ‘John looks like John or Mary, I don’t remember which’.

Other generalized quantifiers

How about other generalized quantifiers? Partee (2002) points out that for quantifiers containing no singleton sets BE will output ‘uninteresting’ properties insofar as they denote the empty set. McNally (1998) argues that because these outputs are of no communicative interest, they are not suitable inputs for BE. A quantifier that is not a singleton-set generator, then, will simply lack a property-type denotation, and systematically receive a wide scope interpretation in similarity statements. In particular, universal and proportional quantifiers lack property-type denotations because the NP they apply to will not contain, in the overwhelming majority of models, the singleton sets that BE looks for (Partee, 2002). The intuition behind these blocked type-shifts, as put by McNally (1998), is that ‘these NPs lack such denotations because their determiners are fundamentally relational and therefore cannot be treated as one-place properties of (atomic or sum) individuals; consequently, their descriptive content cannot be used to identify an individual.’²⁶ For instance, if we apply BE to ‘every’, we get the property of being every lawyer. This property is true of a lawyer only if there exists only one lawyer, or else is true of no lawyer at all:

$$(134) \quad \text{BE}(\llbracket \text{every lawyer} \rrbracket) = \lambda x. \forall y (\text{lawyer}(y)) \rightarrow (y = x)$$

This would lead to an unreasonable interpretation of the corresponding similarity statement, as in (135a); the wide scope reading in (135b) will thus be preferred.

- (135) John is like every lawyer.
- a. # Generally, if someone is every lawyer, then John shares relevant properties with them.
 - b. Every lawyer is such that generally, if someone is them, John shares relevant properties with them.

A similar reasoning applies to quantifiers like ‘most’: no individual has the property of being most lawyers, unless they’re the only existing lawyer:

$$(136) \quad \text{BE}(\llbracket \text{most lawyers} \rrbracket) = \lambda x. \text{MOST } y (\text{lawyer}(y)) (y = x)$$

Then, it is reasonable to assume that (137) will receive interpretation (137b) rather than (137a):

- (137) John is like most lawyers.
- a. # Generally, if someone is most lawyers, then John looks like them.
 - b. Most lawyers are such that generally, if someone is them, John shares relevant properties with them.

²⁶See pp. 371-375 of the same article for a comprehensive list of such quantifiers.

This perspective suggests a specific empirical generalization. If we suppose, with Partee (2002), that predicative uses of indefinites in ‘be’ copulas are yielded by application of the type shifter BE, then we expect a systematic parallelism between two phenomena:

- (i) unavailabilities of predicative readings of Generalized Quantifiers with ‘to be’ and
- (ii) systematic wide interpretations of quantifiers embedded in ‘like’, i.e. unavailabilities of generic readings with similarity statements.

And indeed, both (138a) and (138b) are ungrammatical:

- (138) a. * John is every lawyer.
b. * John is most lawyers.

Of course in models in which the predicate *lawyer* is itself a singleton set these readings should be available, but as noted by McNally (1998), the competition with more natural alternative sentences like (139) makes such sentences infelicitous nonetheless:

- (139) John is the only lawyer.

The parallelism between similarity statements and ‘be’ copulas, in fact, goes as far as to encompass facts about ‘some’, too:

- (140) John is some lawyer.
a. SPECIFIC
b. *PREDICATIVE
- (141) John looks like some lawyer.
a. SPECIFIC
b. *GENERIC

Turning to ‘someone’, note that a generic reading of ‘someone’ in similarity statements seems to be licensed when ‘someone’ is subtriggered:

- (142) John looks like someone.
a. SPECIFIC
b. # GENERIC
- (143) John looks like someone who just had an argument.
a. SPECIFIC
b. GENERIC

This is highly reminiscent of the subtriggering effects in characterizing sentences first discovered by Carlson (1981): (146), unlike (144) and (145), is acceptable under a generic reading.

- (144) Some lawyer should be punctual.
a. Existential

b. #Generic

(145) Someone should be punctual.

- a. Existential
- b. #Generic

(146) Someone who respects others should be punctual.

- a. Existential
- b. Generic

While this phenomenon remains quite mysterious, the accounts that have been proposed explain it as the result of constraints on what GEN can or cannot quantify over (Dayal, 2004; Mascarenhas, 2012): for one reason or other, the relative clause enables the DP headed by ‘someone’ to be in the restriction of GEN.

Mascarenhas (2012) further elaborates on this data, pointing out that while ‘someone’ can receive generic readings when subtriggered, ‘some’ cannot:

(147) Some lawyer who respects others should be punctual.

- a. Existential
- b. #Generic

Similarity, again, displays the same pattern:

(148) John looks like some lawyer who respects others.

- a. Existential
- b. #Generic

In sum, the availability of the general reading of similarity statements with indefinites (both ‘some *N*’ and ‘someone’) patterns *exactly* with the availability of generic readings with indefinites in other contexts. This constitutes a further argument that the general reading of similarity constructions is in fact a generic reading.²⁷

²⁷Of course an alternative to the solution outlined in this section would have been to take the complement of ‘like’ to denote a *property* instead of a Generalized Quantifier. Such an approach would be in line with much literature that proposes that English ‘a’ indefinites can act as predicates, in copulas (cf. Van Geenhoven 1998; Winter 2002 a.o.) but also in e.g. complements of locative prepositions (Mador-Haim & Winter, 2007). This is of course doable, but this approach would predict less naturally than the ‘Generalized Quantifiers’ approach the ungrammaticality of sentences such as (ia), where the adjective is not a grammatical input for the ‘like’ PP.

- (i) a. * John looks like French.
- b. John looks like a Frenchman.

This is in contrast with predicational copula, where both properties denoted by an adjective and (admittedly) by an indefinite are grammatical.

- (ii) a. John is French.
- b. John is a Frenchman.

In the ‘predicative indefinites’ approach, this pattern would have to be explained away by some possibly syntactic stipulations. In the ‘Generalized Quantifiers’ approach taken here, instead, the ungrammaticality of (ia) follows directly from the type mismatch between the unsaturated argument and the adjective.

Be that as it may, the issue comes down to whether indefinites directly denote properties and can be existentially closed via e.g. Semantic Incorporation (as in Van Geenhoven 1998, but also in McNally 1992; Zimmermann 1993), or whether they start out as an existential quantifier (or possibly an existentially closed choice function) that gets type-shifted into a property (in the spirit of Partee 2002). These are issues beyond the scope of this paper. Importantly, however, the ‘Predicative Indefinites’ approach would still be

4.4 An extension of the analysis to non-copular occurrences of 'like'

The entry given in the last section generalizes to any copular construction embedding 'like': 'sound like', 'taste like', etc, as illustrated in 3. At this point, we would like to also give an entry for the non-copular uses of 'like', such as the one occurring in (149):

(149) John walks like Mary.

The available paraphrases suggest that these uses are adverbial.

(150) John is like Mary.

- a. John is similar to Mary.
- b. *John is similarly to Mary.

(151) John walks like Mary.

- a. ??John walks similar to Mary.
- b. John walks similarly to Mary.

It is easy to check that this expression shares the core properties of its 'adjectival' counterpart.

(i) It is non-elliptical (cf. arguments in 2.1.1):

(152) John wants to walk like Mary.

- a. *'John wants to walk like Mary walks.'*
- b. **'John wants to walk like Mary wants to walk.'*

(153) a. *Who does John walk like [__ does]?

- b. Who does John walk like __?

(ii) It is gradable and sensitive to WRT phrases.

(154) a. John walks much like Mary.

- b. John walks like Mary in every respect.

(iii) It is inherently generic.

(155) John walks like Mary or Sue.

≈

John walks like Mary and John walks like Sue.

(156) John often walks like a lawyer.

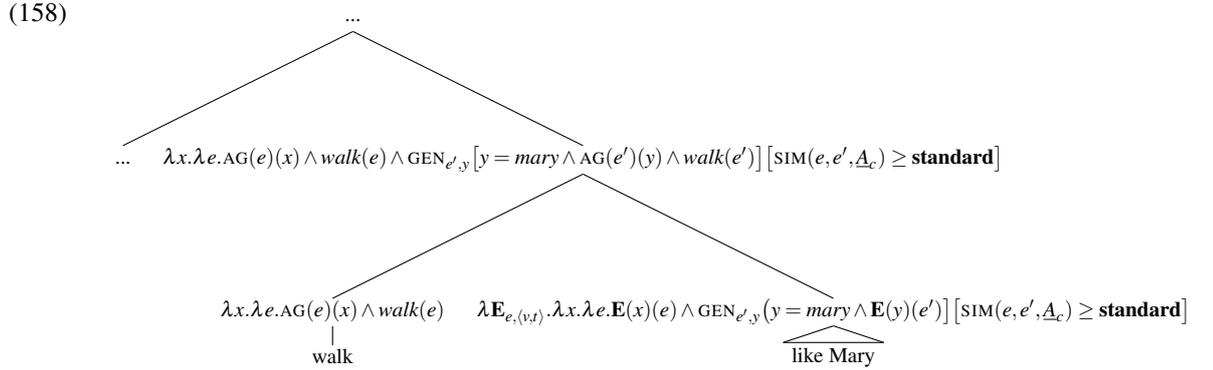
- ≈ John walks like many lawyers.

I propose to analyze it as a manner adverb related to the 'adjectival' use of 'like'. In analogy with the difference between 'quick' and 'quickly', while "adjectival" 'like' is compatible with the fact that 'like' is lexically generic.

relates individuals, “adverbial” ‘like’ relates events.

$$(157) \quad \llbracket \text{like}_{\text{adv.}} \rrbracket = \\ \lambda \mathbf{Q}_{\langle e,t \rangle,t} . \lambda \underline{A} . \lambda d . \lambda \mathbf{E}_{e,\langle v,t \rangle} . \lambda x . \lambda e . \mathbf{E}(x)(e) \wedge \\ \text{GEN}_{e',y} (\text{BE} [\mathbf{Q}(y) \wedge \mathbf{E}(y)(e')] [\text{SIM}(e, e', \underline{A}) \geq d])$$

\mathbf{E} essentially replicates the thematic relation holding between the verb and its subject within ‘like’. Thus, with (157), the expression ‘walk like Mary’ is true of an individual if they are agent of a walking event that shares relevant properties with typical walking events that have Mary as an agent, as shown in the structure in (158).



4.5 A note on the perceived asymmetry of sentences with ‘like’

Consider the contrast between (159a) and (159b), noticed by Tversky (1977).

- (159) a. The son is like the father.
b. The father is like the son.

Sentence (159a) conveys that the son takes on features of the father; (159b) that the father takes on features of the son, which is presumably why it is much less frequent. A similar example is the one we saw in section 4.2, which shows that the Comment serves as an anchor for the similarity respects.

- (160) Context: John has a red nose, Bob has yellow ears.
a. John is like Bob. \rightsquigarrow John has yellow ears.
b. Bob is like John. \rightsquigarrow Bob has a red nose.
c. Bob and John are alike. \rightsquigarrow Bob and John have a red nose and yellow ears.

Tversky accounted for this pattern by assuming that similarity is inherently asymmetrical, featuring a hard-wired salience imbalance between the features of the first and of the second argument, and presupposing that the second argument is more salient. Gleitman *et al.* (1996) resisted this hypothesis, arguing on the basis of a series of experimental findings that the perceived asymmetries are the result of fixed Figure/Ground effects in constructions with ‘like’, or ‘similar to’. We can see the

hypothesis of a rigid Figure/Ground structure as the psychological version of our hypothesis of a fixed Topic/Comment information structure. Gleitman observed that with ‘like’, for instance, the first argument is systematically the Figure, and the second is systematically the Ground; and this is not so in sentences involving ‘alike’.

(161) The father and the son are alike.

My analysis, in fact, expects sentences involving ‘like’ to be *both* semantically and information-structurally asymmetric.

- (i) The semantic asymmetry comes from the fact that the GEN lexically embedded by ‘like’ binds its argument, but not the subject of the sentence.
- (ii) The information-structural asymmetry is motivated by the inability of arguments of ‘like’ to be bound by overt quantificational adverbs, and in general by the discussion in section 4.2.

A way to show that this insight is correct consists in showing that the asymmetry is neutralized just in case we neutralize both (i) and (ii) – while other expressions like ‘alike’ are symmetrical, as they are neither inherently generic nor information-structurally rigid (cf. section 5.1).

Reducing the impact of (i) consists in making the subject of the sentence generic, too. Notice indeed that the two sentences in (162) feel more equivalent than the sentences in (163).

(162) a. A judge is like a lawyer.
b. A lawyer is like a judge.

(163) a. John is like a judge.
b. A judge is like John.

The same goes for sentences with two proper names: we should have a verb with habitual aspect so that the subject is interpreted in the restrictor of a generic quantifier, too.²⁸ Notice indeed that the meanings of the sentences in (164), which have habitual aspect, are closer than the meanings of the sentences in (165), which have progressive aspect.

(164) a. John walks like Bob.
↪ John habitually walks in the way Bob habitually walks.
b. Bob walks like John.
↪ Bob habitually walks in the way John habitually walks.

(165) a. John is walking like Bob.
↪ Right now, John is walking in the way Bob habitually walks.

²⁸I here refer to the theory of habituality defended in Chierchia (1995), which assumes that *Gen* is brought about by verbal aspect and gives rise to habitual interpretations:

(i) John smokes.
GEN_x[x = john][smokes(x)]

- b. Bob is walking like John.
 ↪ Right now, Bob is walking in the way John habitually walks.

This is not enough though, since in (163) we can still have two different inferences, as shown in (166). Similarly for proper names, in sentences like (160).

- (166) *(Say a judge wears a wig and a lawyer wears a tie)*
- a. A judge is like a lawyer. ↪ A judge wears a tie.
 - b. A lawyer is like a judge. ↪ A lawyer wears a wig.

In addition to reducing the impact of (i), we must reduce the impact of (ii). Reducing the impact of (ii) consists in looking at a context where nothing is known about either argument: in this case, the information-structural asymmetry should have a limited impact on meaning, since the inference that the Topic constitutes known information and the Comment ‘news’ can have no effect on the choice of the properties relevant for comparison. This involves, for instance, talking about things about which nothing is known except for the fact that they are alike. And indeed, the sentences in (167) feel equivalent, and so do those in (168).

- (167) *‘Look, I know absolutely nothing about ‘schmudges’ and ‘schmawyers’. The only thing I’ve been told is that...*
- a. A schmudge walks like a schmawyer.
 - b. A schmawyer walks like a schmudge.
- (168) *‘Look, I know absolutely nothing about suspect 1 and suspect 2. The only thing I know is that...*
- a. Suspect 1 walks like suspect 2.
 - b. Suspect 2 walks like suspect 1.

Notice in turn that only reducing the impact of (ii) is not enough either. Regardless of what we know about suspect 1 and suspect 2, the sentences in (169) are semantically distinct, perhaps most remarkably in the fact that the indefinite receives an existential interpretation in (169a) (because of the absence of habitual aspect), but a generic one in (169b).

- (169) *(I know nothing about ‘schmawyers’ and about suspect 1, but I see that...*
- a. A schmawyer is walking like suspect 1.
 ‘Right now, there is a schmawyer that is walking in the way suspect 1 habitually walks’.
 - b. Suspect 1 is walking like a schmawyer.
 ‘Right now, suspect 1 is walking like a typical schmawyer habitually walks’.

5 Discussion and open issues

5.1 Other lexically generic constructions

A number of other constructions align with ‘like’ in terms of non-adverbial genericity.

- (170) Similarity verbs without ‘like’:²⁹
- a. **Upward- and downward-non-monotonicity:**
 - (i) resemble a French lawyer $\not\approx$ resemble a lawyer
 - (ii) resemble a lawyer $\not\approx$ resemble a French lawyer
 - b. **Narrow readings of disjunction**
resemble a lawyer or a judge \approx resemble a lawyer and resemble a judge
 - c. **Subtriggering pattern**
 - (i) resemble someone { \exists , *GEN}
 - (ii) resemble someone who respects others { \exists , GEN}
 - d. ***a N \approx a typical N.***
resemble a lawyer \approx resemble a typical lawyer
 - e. **No binding by Q-adverbs.**
John often resembles a lawyer $\not\approx$ John resembles many lawyers
- (171) Comparatives:
- a. **Upward- and downward-non-monotonicity:**
 - (i) more competent than a French lawyer $\not\approx$ more competent than a lawyer
 - (ii) more competent than a lawyer $\not\approx$ more competent than a French lawyer
 - b. **Narrow readings of disjunction**
more competent than a lawyer or a judge \approx more competent than a lawyer and more competent than a judge
 - c. **Subtriggering pattern**
 - (i) more competent than someone { \exists , *GEN}
 - (ii) more competent than someone who respects others { \exists , GEN}
 - d. ***a N \approx a typical N.***
more competent than a lawyer \approx more competent than a typical lawyer
 - e. **No binding by Q-adverbs.**
John is often more competent than a lawyer $\not\approx$ John is more competent than many lawyers
- (172) Equatives:³⁰
- a. **Upward- and downward-non-monotonicity:**
 - (i) as competent as a French lawyer $\not\approx$ as competent as a lawyer
 - (ii) as competent as a lawyer $\not\approx$ as competent as a French lawyer
 - b. **Narrow readings of disjunction**
as competent as a lawyer or a judge \approx as competent as a lawyer and as competent as a judge

²⁹I thank an anonymous reviewer for pointing this out.

³⁰I thank an anonymous reviewer for pointing this out.

- c. **Subtriggering pattern**
 (i) as competent as someone { \exists , *GEN}
 (ii) as competent as someone who respects others { \exists , GEN}
- d. **$a N \approx a$ typical N .**
 as competent as a lawyer \approx as competent as a typical lawyer
- e. **No binding by Q-adverbs.**
 John is often as competent as a lawyer $\not\approx$ John is as competent as many lawyers
- (173) {*Similar to, different from, the same as*}.³¹
- a. **Upward- and downward-non-monotonicity:**
 (i) {Similar to, different from, the same as} a French lawyer $\not\approx$ {Similar to, different from, the same as} a lawyer
 (ii) {Similar to, different from, the same as} a lawyer $\not\approx$ {Similar to, different from, the same as} a French lawyer
- b. **Narrow readings of disjunction**
 {Similar to, different from, the same as} a lawyer or a judge \approx {Similar to, different from, the same as} a lawyer and {Similar to, different from, the same as} a judge
- c. **Subtriggering pattern**
 (i) {Similar to, different from, the same as} someone { \exists , *GEN}
 (ii) {Similar to, different from, the same as} someone who respects others { \exists , GEN}
- d. **$a N \approx a$ typical N .**
 {Similar to, different from, the same as} a lawyer \approx {Similar to, different from, the same as} a typical lawyer
- e. **No binding by Q-adverbs.**
 John is often {similar to, different from, the same as} a lawyer $\not\approx$ John is {similar to, different from, the same as} many lawyers
- (174) ‘the X of a Y’:
- a. **Upward- and downward-non-monotonicity:**
 (i) the charisma of a French lawyer $\not\approx$ the charisma of a lawyer
 (ii) the charisma of a lawyer $\not\approx$ the charisma of a French lawyer
- b. **Narrow readings of disjunction**
 having the charisma of a lawyer or a judge \approx having the charisma of a lawyer and having the charisma of a judge
- c. **Subtriggering pattern**
 (i) the charisma of someone { \exists , *GEN}
 (ii) the charisma of someone who respects others { \exists , GEN}
- d. **$a N \approx a$ typical N .**
 the charisma of a lawyer \approx the charisma of a typical lawyer
- e. **No binding by Q-adverbs.**
 John often has the charisma of a lawyer $\not\approx$ John has the charisma of many lawyers

This distinguishes these constructions from others whose complements can be interpreted *both* in the restriction of GEN and of overt quantificational adverbs, as happens

³¹I thank an anonymous reviewer for pointing this out.

in the ‘computer-modern plane’ examples in (3)-(3b), or in (119).

Are rigid information structure and lexically inherent genericity correlated? As we have seen in 4.2, ‘alike’, which is nearly equivalent to ‘like’, does not display a rigid Topic-Comment information structure. There are equivalent examples for other cases, as for instance the case of equatives. Such expressions seem to be flexible in terms of information structure, since explicit Q-adverbs can bind arguments embedded by them (cf. (175)) and (ii) their two arguments cannot be respectively Topic and Comment (cf. (176)-(177)).

- (175) a. A lawyer and a judge are often alike.
 $OFTEN_{x,y}[lawyer(x) \wedge judge(y)][alike(x,y)]$
 b. A lawyer and a judge are often equally competent.
 $OFTEN_{x,y}[lawyer(x) \wedge judge(y)][equally-competent(x,y)]$
- (176) *We are talking about John’s personality.*
 a. John is like Mary.
 b. #John and Mary are alike.
- (177) *We are talking about John’s competence.*
 a. John is as competent as Mary.
 b. #John and Mary are equally competent.

These expressions are infelicitous, at least out of the blue, if the second argument is generic, unlike what happens with their information-structurally rigid cousins:

- (178) a. John is like a lawyer.
 b. #John and a lawyer are alike.
- (179) How competent is John?
 a. John is as competent as a lawyer.
 b. #John and a lawyer are equally competent.

These facts would seem to suggest that there is a parallel between inherent genericity and a rigid Topic-Comment structure. ‘Like’ is rigid and inherently generic, ‘alike’ is flexible and not inherently generic.

However, there are other expressions that have at least been claimed to be inherently generic too, namely individual-level predicates (Chierchia, 1995). Chierchia’s hypothesis that these are inherently generic is aimed at explaining the infelicity of sentences such as (180):

- (180) #John is often intelligent.

These expressions do not have the same information structure as ‘like’, as shown by the fact that their arguments can be bound by overt quantificational adverbs.³²

³²And this does not rule out the possibility that they lexically embed GEN as the logical form of (i) could have a GEN nested inside OFTEN as in (ia).

- (i) A lawyer is often intelligent.
 a. $OFTEN_x[lawyer(x)][GEN_y[y=x][intelligent(y)]]$

- (181) A lawyer is often intelligent.
 \approx Many lawyers are intelligent.
- (182) a. John is intelligent.
 $\text{GEN}_x[x = \textit{john}][\textit{intelligent}(j)]$
 b. A lawyer is intelligent.
 $\text{GEN}_x[\textit{lawyer}(x)][\textit{intelligent}(j)]$

Consequently, the difference between individual-level predicates and ‘like’ phrases is that, because ‘like’ phrases prevent binding of their arguments by quantificational adverbs, they allow us to unambiguously prompt the presence of a lexical GEN.

Depending on whether one believes that individual-level predicates are inherently generic, there are two possible avenues to try and capture the relationship between rigid information structure and the data in (170)-(174).

- The first one is that there is a correlation between rigid information structure and lexically inherent genericity. This avenue would involve rejecting (a variation of) Chierchia’s (1995) hypothesis that individual-level predicates like ‘intelligent’ are lexically generic, and thus explain their behaviour in other terms. It would also have to explain why it is precisely broadly ‘comparative’ expressions that display this behavior.
- The second theoretical possibility is to say that lexically generic expressions might be anywhere in the grammar, but that information-structurally rigid expressions such as those in (170)-(174) allow us to *univocally prompt* inherent genericity. This is because we can show that expressions they embed have logical properties only explained by the presence of GEN, but at the same time compositionally diverge from explicit adverbial quantification. This avenue would have to precisely pin down the class of inherent generics that are not information-structurally rigid by proposing new tests to prompt inherent genericity.

5.2 Temporal arguments

A reviewer observes that given a strong context, ‘like’ is compatible with quite contingent statements. In the context in (183), it is acceptable to utter (183a), although Mary typically doesn’t soil her clothes.

- (183) **Context:** In the kindergarden, Mary is typically very well-behaved, and never soils her clothes while playing in the yard. One day, while picking up my son John, I see that both John and Mary are completely covered in dirt.
- a. John looks like Mary.

Until now, we have been working with a simplified version of GEN that only binds individual variables. However, we have, for independent grounds, reasons assume that GEN binds at least times too, for at least two reasons. First, GEN can be restricted by ‘when’-clauses.

- (184) a. When it rains, a lawyer wears a tie.

≈ Typical lawyer-raining time pairs $\langle x, t \rangle$ are such that x wears a tie at t .

- (185) a. When it rains, it doesn't snow.
 $GEN_t[rain(t)][\neg snow(t)]$
 b. When it rains, it usually doesn't snow.
 $USUALLY_t[rain(t)][\neg snow(t)]$

Second, standard theories of the silent Q-adverb *Gen* argue that it is brought about by aspect (see Chierchia 1998, a.o.). An argument for this is that the availability of generic readings seems to be modulated by aspect.

- (186) a. A bird flies. GENERIC
 b. A bird is flying. EXISTENTIAL

Additionally, GEN should come with a restriction on relevant times. (187) needs to be restricted to times where John is not under water, where he could not light a cigarette:

- (187) a. John smokes.
 b. $GEN_{x,t}[x = john \wedge C(t)][smoke(x,t)]$
'Take any relevant time and any person that is John, that person smokes at that time'

One possible avenue to explain the data point in (183a) is to assume that there is in fact generic quantification, but the contextual restriction simply restricts it to times around the time of the utterance. This results in trivial generic quantification. If this is so, why is it that we cannot say (188a) instead of (188b), when we are observing a current event in which John is walking? After all, (188a) involves generic quantification too.

- (188) **Context:** we observe John who is walking in front of us.
 a. * John walks.
 b. John is walking.

One possibility is simply to say that (188a) competes with the more restricted (188b), and therefore cannot bear the meaning that is restricted to a specific time. In other languages with arguably morphologically more complex ways of expressing progressive aspect, such as Italian and French, the two forms don't compete with each other, and thus the present simple can describe temporally circumscribed actions:

- (189) **Context:** we observe John, who is walking in front of us.
 a. John cammina.
 John walks.
'John is walking'
 b. John sta camminando.
 John is walking.
'John is walking'

To sum up, while this remains an open issue for my analysis, it is possible that in English genericity/habituality is necessarily limited to stable properties and actions only because habitual aspect competes with the progressive. This competition does not

concern the lexical genericity encoded in ‘like’, which therefore behaves like habitual aspect behaves in languages where the two verbal aspects don’t compete.

6 Conclusion

In this paper, I have presented a view on ‘like’ and related similarity constructions. My theory views ‘like’ as a gradable expression that denotes the proportion of properties that two individuals share from a set of relevant dimensions. This made it possible to capture a complex pattern of scalar modification. Mostly, it allowed us to (i) capture precisely the meaning of expressions such as 99% like *X*, (ii) capture the difference between WRT phrases and scalar modifiers. Though the main point of the paper concerns lexical genericity: there are many arguments to think that in addition, ‘like’ incorporates a hard-wired generic operator in its lexical meaning. We saw that this explains a range of phenomena that are otherwise very puzzling: strong readings of disjunction, a characteristic subtriggering pattern, and the non-monotonicity of these expressions. Crucially, this highlighted that there is a range of expressions in language sharing the same surprising properties: equatives and other broadly comparative expressions all (i) have a rigid Topic-Comment structure, (ii) display behavior consistent with (inherent) generic quantification. This raises the question of how to precisely predict and explain which constructions in language are lexically generic and which ones aren’t.

References

- Alrenga, P. (2010) Comparisons of similarity and difference. *Adjectives: Formal analyses in syntax and semantics* **153**, 155.
- Asher, N. & Morreau, M. (1995) What some generic sentences mean. *The generic book* pp. 300–338.
- Bar-Lev, M.E. (2021) An implicature account of homogeneity and non-maximality. *Linguistics and Philosophy* **44**, 1045–1097.
- Bar-Lev, M.E. & Fox, D. (2020) Free choice, simplification, and innocent inclusion. *Natural Language Semantics* **28**, 175–223.
- Brisson, C.M. (1998) *Distributivity, maximality, and floating quantifiers*. Rutgers The State University of New Jersey-New Brunswick.
- Carlson, G.N. (1981) Distribution of free-choice any. *Papers from the... Regional Meeting. Chicago Ling. Soc. Chicago, Ill*, 17, pp. 8–23.
- Chierchia, G. (1995) Individual-level predicates as inherent generics. *The Generic Book* **125**, 176–212.
- Chierchia, G. (1998) Reference to kinds across language. *Natural language semantics* **6**, 339–405.
- Chierchia, G. (2009) Dynamics of meaning. *Dynamics of Meaning*, University of Chicago Press.
- Coulson, S. & Fauconnier, G. (1999) Fake guns and stone lions: Conceptual blending and privative adjectives. *Cognition and function in language* pp. 143–158.
- Dayal, V. (2004) Licensing by modification. *Ilha do Desterro Florianópolis* **47**, 217–238.
- de Swart, H. (1996) (in) definites and genericity' in m. kanazawa, c. pinon, and h. de swart (eds.), quantifiers, deduction and context. center for the study of language and information, stanford.
- Del Pinal, G. (2015) Dual content semantics, privative adjectives, and dynamic compositionality. *Semantics and Pragmatics* **8**, 7–1–53
- Del Pinal, G. (2018) Meaning, modulation, and context: a multidimensional semantics for truth-conditional pragmatics. *Linguistics and Philosophy* **41**.
- Durbin, H., Novelli, M. & Bodmer, W. (1997) Genomic and cDNA sequence analysis of the cell matrix adhesion regulator gene. *Proceedings of the National Academy of Sciences* **94**, 14578–14583.
- Feinmann, D. (2020) *Lexical illusions, (non-) maximality, and invisible gaps*. Ph.D. thesis, University of Sheffield.
- Franks, B. (1995) Sense generation: A “quasi-classical” approach to concepts and concept combination. *Cognitive science* **19**, 441–505

- Gelman, S.A. et al. (2003) *The essential child: Origins of essentialism in everyday thought*. Oxford University Press, USA.
- Gentner, D. & Markman, A.B. (1997) Structure mapping in analogy and similarity. *American psychologist* **52**, 45.
- Gleitman, L.R., Gleitman, H., Miller, C. & Ostrin, R. (1996) Similar, and similar concepts. *Cognition* **58**, 321–376.
- Guerrini, J. (2021) Similarity statements, privative adjectives, and generics: a unified view. *Qualifying paper, Ecole Normale Supérieure* .
- Gust, H. & Umbach, C. (2015) Making use of similarity in referential semantics. *Modeling and Using Context: 9th International and Interdisciplinary Conference, CONTEXT 2015, Lanarca, Cyprus, November 2-6, 2015. Proceedings 9*, pp. 425–439, Springer.
- Hampton, J.A. (1998) Similarity-based categorization and fuzziness of natural categories. *Cognition* **65**, 137–165.
- Jäger, G. (2001) Topic-comment structure and the contrast between stage level and individual level predicates. *Journal of Semantics* **18**, 83–126.
- Kennedy, C. & McNally, L. (2005) Scale structure, degree modification, and the semantics of gradable predicates. *Language* pp. 345–381.
- Kratzer, A. (1995) Individual-level predicates. *The generic book* **125**.
- Krifka, M. (2001) Non-novel indefinites in adverbial quantification. *Logical Perspectives on Language and Information, CSLI Publications, Stanford* pp. 1–40.
- Krifka, M. (2008) Basic notions of information structure. *Acta Linguistica Hungarica* **55**, 243–276.
- Krifka, M., Pelletier, F.J., Carlson, G., Ter Meulen, A., Chierchia, G. & Link, G. (1995) Genericity: an introduction. *The Generic Book* .
- Križ, M. (2015) Aspects of homogeneity in the semantics of natural language: University of vienna dissertation.
- Križ, M. & Spector, B. (2021) Interpreting plural predication: Homogeneity and non-maximality. *Linguistics and Philosophy* **44**, 1131–1178.
- Kuroda, S.Y. (1972) The categorial and thethetic judgment: Evidence from Japanese syntax. *Foundations of language* pp. 153–185.
- Leslie, S.J. (2008) Generics: Cognition and acquisition. *Philosophical Review* **117**, 1–47.
- Leslie, S.J. (2015) Generics oversimplified. *Noûs* **49**, 28–54.
- Löbner, S. (2000) Polarity in natural language: Predication, quantification and negation in particular and characterizing sentences. *Linguistics and Philosophy* **23**, 213–308.

- Mador-Haim, S. & Winter, Y. (2007) Non-existential indefinites and semantic incorporation of pp complements. *Semantics and Linguistic Theory*, vol. 17, pp. 184–201.
- Malamud, S.A. (2012) The meaning of plural definites: A decision-theoretic approach. *Semantics and Pragmatics* **5**, 3–1.
- Mari, A., Beyssade, C. & Del Prete, F. (2012) *Genericity*, vol. 43. OUP Oxford.
- Martin, J. (2022) *Compositional routes to (non) intersectivity*. Ph.D. thesis, Harvard University.
- Mascarenhas, S. (2012) Licensing by modification: the case of positive polarity pronouns. *Proceedings of Sinn und Bedeutung*, vol. 16, pp. 417–430.
- McNally, L. (1998) Existential sentences without existential quantification. *Linguistics and Philosophy* pp. 353–392.
- McNally, L. (2005) Lexical representation and modification within the noun phrase. *Recherches linguistiques de Vincennes* pp. 191–206.
- McNally, L.E. (1992) *An interpretation for the English existential construction*. University of California, Santa Cruz.
- Napoli, D.J. & Nespors, M. (1986) Comparative structures in Italian. *Language* **62**, 622–653.
- Partee, B. (2002) Noun phrase interpretation and type-shifting principles. *Formal semantics: The essential readings* pp. 357–381.
- Partee, B.H. (1989) Binding implicit variables in quantified contexts. *University of Massachusetts Occasional Papers in Linguistics* **15**, 12.
- Pustejovsky, J. (1998) *The generative lexicon*. MIT press.
- Rips, L.J. (1989) Similarity, typicality, and categorization. *Similarity and analogical reasoning* **2159**.
- Rotstein, C. & Winter, Y. (2004) Total adjectives vs. partial adjectives: Scale structure and higher-order modifiers. *Natural language semantics* **12**, 259–288.
- Rudolph, R.E. (2019) *Talking about Appearances: Experience, Evaluation, and Evidence in Discourse*. Doctoral Dissertation, University of California, Berkeley.
- Sassoon, G.W. (2013) A typology of multidimensional adjectives. *Journal of semantics* **30**, 335–380.
- Stanley, J. (2002) Nominal restriction.
- Stephenson, T. (2007) Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and philosophy* **30**, 487–525.
- Tversky, A. (1977) Features of similarity. *Psychological review* **84**, 327.
- Umbach, C. & Gust, H. (2014) Similarity demonstratives. *Lingua* **149**, 74–93.

- Umbach, C. & Gust, H. (2021) Grading similarity. *Concepts, frames and cascades in semantics, cognition and ontology*, pp. 365–386, Springer International Publishing Cham.
- Van Geenhoven, V. (1998) *Semantic incorporation and indefinite descriptions: Semantic and syntactic aspects of noun incorporation in West Greenlandic*. CSLI publications.
- van Rooij, R. & Schulz, K. (2020) Generics and typicality: a bounded rationality approach. *Linguistics and Philosophy* **43**, 83–117.
- Von Stechow, P. (1994) *Restrictions on quantifier domains*. University of Massachusetts Amherst.
- Winter, Y. (2002) *Flexibility principles in Boolean semantics: The interpretation of coordination, plurality, and scope in natural language*. MIT press.
- Zimmermann, T.E. (1993) On the proper treatment of opacity in certain verbs. *Natural Language Semantics* **1**, 149–179.