Reference and Substitution: A New Perspective*

Ash Asudeh
University of Oxford & Carleton University
[Joint work with Gianluca Giorgolo]

Institute of Linguistics Colloquium
University of Minnesota

September 16, 2016

1 Introduction

- An important problem in the philosophy of language and the linguistic study of meaning (semantics and pragmatics) concerns co-referential terms and substitutability in different contexts.
- In the modern context, this problem is commonly associated with Frege (1892), and is often called *Frege's puzzle*.
- The puzzle can be presented in various ways, but its essence can be captured as follows: given two co-referential linguistic expressions, why is it that in certain linguistic contexts substitution of one expression for the other is truth-preserving, while in others it is not?
- For example, given that (1) is true, since *Hesperus* and *Phosphorus* are different names for the planet Venus, how can it be that (2) can be true while (3) is false?

  (1) Hesperus is Phosphorus.

  (2) Kim believes that Hesperus is a planet.

  (3) Kim believes that Phosphorus is a planet.

- Alternatively, we could characterize the puzzle by observing that a sentence like the following can be true without entailing that Kim does not believe a tautology:

  (4) Kim doesn't believe that Hesperus is Phosphorus.

- Frege's own solution was that in addition to a reference, nominals have a sense, or 'mode of presentation', and that in certain contexts, such as those involving propositional attitudes, it is these distinct senses that block substitutability.
- Frege's puzzle is thus clearly related to the problem of 'referential opacity' in the study of propositional attitudes (Quine 1953, 1956, 1960).
- Fregean senses are not the only way to construe modes of presentation (e.g., Schiffer 1990, Fiengo and May 1998) and the notion that names, in particular, have a mode of presentation or are interpreted differently under propositional attitude verbs is not universally accepted (among many others, Kripke 1972, 1979, 1980, Recanati 1997, Richard 1990).

*The paper that this talk is based on (Asudeh and Giorgolo 2016) is available here: `http://semprag.org/article/view/sp.9.21`

- We focus on linguistics aspects of substitutability/opacity.
- We take it for granted that there is an empirical phenomenon to be explained here — differing truth value judgements despite substitution of co-referential terms — and offer a formal mechanism for capturing and explaining it semantically.
- We follow Saul (1997) in observing that problems of substitutability also arise in 'simple sentences'. Our analysis captures some of these cases, too.
- Moreover, we also focus on cases of differential interpretation of the *same* expression (Kripke 1979, Fiengo and May 1998).
- Lastly, we briefly indicate how our analysis could give insight into cases other than referential expressions, as discussed by Carnap (1947), Mates (1950), and Kripke (1979), among others.
- Our proposal is not only formally well-founded on advances in formal logic and theoretical computer science, it also gives us the beginnings of a general semantics of what we might informally call perspective.[1]

2 Overview

Section 3 outlines the scope of the problem and set out the space of example types to be analyzed.

Section 4 introduces our key formal mechanism, *monads*.

Section 5 offers our analyses of particular examples.

Section 6 compares our approach to indices to that of Fiengo and May (1998) and our solutions to the puzzles to the sort that could be offered by a traditional approach based on scope in composition.

Section 7 concludes.

[1]This notion of perspective may well be related to notions of perspective or point of view in other phenomena, such as demonstratives and indexicals (e.g., Kaplan 1989, Schlenker 2003), *de se* attitudes (e.g., Lewis 1979, Chierchia 1989, Pearson 2013), logophoricity and related pronominal interpretations (e.g., Sells 1987, Hagège 1974, Kuno 1987, Oshima 2006, Sundaresan 2012), illocutionary adverbs (e.g., Austin 1975, Krifka 2001, 2014, Ernst 2009), expressivity (e.g., Potts 2005, 2007, Gutzmann 2012, Gutzmann and Gärtner 2013), and predicates of personal taste (e.g., Lasersohn 2005, Stephenson 2007a,b, MacFarlane 2014). We have not yet had the opportunity to systematically explore this, but we hope to do so in future work.

3 The Scope of the Problem

- The substitutability puzzle is standardly characterized as involving two factors:
  1. Embedding under a modal or propositional attitude expression, such as *believe*; and
  2. Co-referential but distinct terms, such as *Hesperus* and *Phosphorus*
- It has been shown in the literature that neither of these factors is necessary for the substitutability puzzle or related puzzles to arise, which yields a typology of four major classes of substitution puzzles.

3.1 Simple Sentences

- With respect to the first factor, Saul (1997) points out that the lack of substitutability can hold even in 'simple sentences' that 'contain no attitude, modal or quotational constructions' (Saul 1997: 102, fn.1).
  - Assuming it is common knowledge that Clark Kent is Superman's secret identity, she notes that if (5) is true, substitution of *Clark Kent* for *Superman* seems to render (6) false (Saul 1997: 102, (1) & (1*)):

    (5) Clark Kent went into the phone booth, and Superman came out.

    (6) Clark Kent went into the phone booth, and Clark Kent came out.
  - With respect to this pair, an obvious out presents itself: Why would someone *say* (6) if they *meant* (5)? That is, it seems that we could say these sentences are in fact semantically synonymous (so (6) is not false when (5) is true), but pragmatically distinct.
  - This is in fact basically what Saul argues (Braun and Saul 2002, Saul 2007): namely, what is mistaken is our intuition that (5) is true while (6) is false.
- We set this issue aside here and focus instead on a different simple sentence that makes the same point, but seems to us not to offer as much pragmatic wiggle room:[2]

  (7) Mary Jane loves Peter Parker, but she doesn't love Spider-Man.

  - Let's assume that the time of evaluation is a point in the stories before Mary Jane knows that Peter Parker is Spider-Man. There is a non-contradictory reading of this sentence.
  - According to the theory presented in Saul (2007), this sentence is simply false, but that seems to entirely set aside Mary Jane's say in the matter, which strikes us as problematic.
  - It would be very strange to insist that if Mary Jane loves Peter Parker, then she really does love Spider-Man. She certainly wouldn't agree to that.
  - In short, (7) shows a lack of substitutability in a grammatically very simple sentence and this lack of substitutability does not yield as readily to a pragmatic explanation as do Saul's examples.
  - Instead, (7) seems to crucially involve Mary Jane's *perspective*. This is an intuition we will build on in our analysis below.

[2]We prefer to use Spider-Man in our examples, because Superman is frankly kind of boring, but also because the Peter Parker/Spider-Man case involves a different (yet still familiar) set-up: it is not as clear which is whose secret identity, since Peter Parker is as much the "main character" in those stories as Spider-Man is. This avoids the problem of Pitt's (2001) concept of *primum egos*, as discussed by Saul (2007: 31–34,140). It also avoids the problem that both Superman and Clark Kent are in fact secret identities of a third identity, Kal-El, a Kryptonian refugee (Saul 2007: 31–34).

- This last observation also distinguishes our approach from one where a sentence like (7) is interpreted as simply saying that Mary Jane loves only one *guise* (Castañeda 1972, 1989, Heim 1998) of the entity that corresponds to Peter Parker but not another one.
  - First, one might object that people love entities not guises.[3]
  - Something like this seems at play in the criticism of MacColl (1905a,b) by Russell (1905: 491).
  - As linguists, we feel neither prepared nor compelled to enter that debate, but a simple guise-based theory in fact makes false empirical predictions, which we are prepared to discuss.
  - If it is indeed the case that different co-referring expressions simply pick out different guises of the same individual, then a sentence like (8) should have a non-contradictory reading, but this does not seem to be the case (assuming it is indeed Peter Parker who is Spider-Man at the time; i.e. there has been no passing of the mantle or any such thing):

    (8) #Dr. Octopus killed Spider-Man but he didn't kill Peter Parker.

- Moreover, the theory must capture the difference between, for example, *kill* and *murder*, since *murder* does involve intention and the minimal pair of (8) with *murder* substituted for *kill* is not contradictory (in the same circumstance):

  (9) Dr. Octopus murdered Spider-Man but he didn't murder Peter Parker.

- Although we characterize *murder* as involving intention, it is not a propositional attitude verb and there is no obvious evidence of embedding. We discuss this further in section 6.
- Lastly, whatever analysis we give must not lose sight of the fact that there are genuine cases of contradiction that must still be derivable, such as the following:

  (10) #Dr. Octopus punched Spider-Man but he didn't punch Spider-Man.

- What unites *murder* and *love* versus *kill* and *punch* is the fact that the subject/agent's *perspective* is part of the interpretation.[4]

[3] We thank Rob Stainton (p.c.) for this point and making us aware of its precursors in Russell (1905).

[4] A proponent of a guise-based theory may contend that our proposal is essentially similar to guises, but unless they can flesh out the comparison, with the same level of formal rigour as we provide, this is just a weak analogy that does not strike us as particularly insightful.

3.2 Non-Distinct Terms but Distinct Beliefs

- Kripke (1979) presents a puzzle that is closely related to the substitutability puzzle, but which relates to the second factor mentioned above: whether the terms involved must be distinct.

  - He considers the case of 'phonetically identical tokens of a single name'.
  - He provides the example of an individual, Peter, who has learned that *Paderewski* was the name of an accomplished Polish pianist. The following then seems true:

    (11) Peter believes that Paderewski had musical talent.

  - Peter then hears of a Polish politician named *Paderewski*, and concludes that this is a different person, since he has no reason to believe that politicians make good musicians. Given that in fact the same Paderewski was in fact both a politician and a pianist, is the following true or not?

    (12) Peter believes that Paderewski had no musical talent.

  - Kripke (1979) argues that this is a true paradox and we can neither conclude that (12) is true nor false, given the situation.

- Fiengo and May (1998) deny this conclusion on the basis of a theory of reference that crucially holds that names do not directly refer but only do so once part of linguistic expressions, which bear distinguishing indices, such as '$[_{\mathrm{NP}_1}$ Paderewski]'.

  - What the speaker believes is characterized by statements of the following form (Fiengo and May 1998: 388):

    (13) $\ulcorner$'$[_{\mathrm{NP}_i}\ X]$' has the value $\mathrm{NP}_i\urcorner$

  - They also propose the following principle:

    (14) **Singularity Principle**
    If cospelled expressions are covalued, they are coindexed.

  - For Fiengo and May, then, there are two distinct Paderewski indexations at play for Peter, which means that the two "cospelled" instances of Paderewski are not covalued, given the Singularity Principle.
  - Fiengo and May (1998: 399) ask us to consider a version of the Paderewski puzzle in which the speaker believes that John believes that there are two people named *Paderewski*, but the speaker herself believes that there is only one (contextually relevant) person named *Paderewski*.
  - The speaker may then say, without contradiction, (15a), which has the Fiengo and May logical form (15b), and (16a), which has the logical form (16b).[5] Thus, the beliefs of John are distinguished by the indexation.

    (15) a. John believes that Paderewski is a genius.
    b. John believes that [$\mathrm{Paderewski}_1$ is a genius and '$\mathrm{Paderewski}_1$' has the value $\mathrm{Paderewski}_1$]

    (16) a. John does not believe that Paderewski is a genius.
    b. John does not believe that [$\mathrm{Paderewski}_2$ is a genius and '$\mathrm{Paderewski}_2$' has the value $\mathrm{Paderewski}_2$]

  - Similarly, so long as Peter believes there are two Paderewskis, he can simultaneously believe that one had musical talent while the other did not.
  - If and once he realizes that these two are the same person, then the Singularity Principle requires that the two *Paderewski* expressions bear the same index and Peter could no longer believe both without contradiction.

[5] Fiengo and May's notation is unfortunately ambiguous; we have added the bracketing to make it clearer.

- The informal theory that Fiengo and May (1998) put forward seems to us to be closely related to the Interpreted Logical Form theory of Larson and Ludlow (1993), who Fiengo and May fail to cite.

  - Larson and Ludlow (1993: 336) provide the following cute and memorable, but ultimately unconvincing example:

    Context: Jason is from New York and does not know how the name *Harvard* is pronounced in a Boston accent.

    (17) Jason believes [Harvard is a fine school].

  - Using [harvard] to indicate Jason's pronunciation of *Harvard* and [hahvahd] to indicate the Boston pronunciation, Larson and Ludlow point out that, given this context, (18) is true, while (19), is false:

    (18) Jason believes that [[harvard] is a fine school].

    (19) Jason believes that [[hahvahd] is a fine school].

  - Why do we find this unconvincing? It seems to us that, for Jason, [harvard] and [hahvahd] are just different words.
  - The fact that they are different pronunciations of the same word is etymological knowledge that is irrelevant to Jason's synchronic knowledge of language.
  - Coincidence of spelling is similarly irrelevant — a criticism that applies to Fiengo and May's Singularity Principle, too (cf. 'cospelled expressions' in (14) above). Kripke in fact characterized things much more aptly when he wrote of 'phonetically identical tokens of a single name': homo*phony* is what's at stake, not homo*graphy*.

- In our opinion, a more satisfactory analysis of these kinds of linguistic puzzles rests on disentangling two different phenomena that seem at play in Paderewski puzzles.

  - Kripke's (1979) conclusion that we are dealing with a paradox seems to us motivated by the interplay between the perspectival dimension introduced by the verb *believe* together with the *ambiguous* nature of the name *Paderewski* in the context of Peter's lexicon.
  - In this case we not only have different perspectives regarding the interpretation of a term (the speaker's and Peter's), but the two interpretations also have different *cardinalities*.
  - Given that Peter can use the name Paderewski to refer to two different (from Peter's perspective) entities, in an example like (12) it is not possible to resolve whether we are talking about Peter's belief with regard to the pianist entity or the politician one.
  - Therefore (12) seems to lack a determinate truth value: It is true with respect to Paderewski the politician, but false with respect to Paderewski the musician.
  - We have competing interpretations, but each one is fully interpretable and can be assigned a truth value.
  - Of course, this move itself only makes sense if the two instances of the name *Paderewski* in fact do not refer to one and the same entity for Peter, which is not possible for Kripke.

3.3 Identity Statements: Delusions and Hucksters

- The observations that homophonous terms and simple sentences can likewise lead to the substitutability puzzle and related puzzles is thus established in the literature.

- But it seems to us that we can drive the point home in an even simpler way, by starting with basic identity statements involving two homophonous tokens of the same name, avoiding accents and bypassing Paderewskis.

- Statements such as the following are normally taken to be uninformative tautologies:

  (20) Sandy is Sandy.

  - If this is true, then a statement like the following should mean that Kim does not believe a tautology:

    (21) Kim doesn't believe Sandy is Sandy.

  - Let us call the reading where the Kim does not believe a tautology an *unsatisfiable* reading.
  - However, sentences like (21) also have satisfiable readings in the right context:

    (22) **Context:** Kim suffers from Capgras Syndrome, also known as the Capgras Delusion, a condition "in which a person holds a delusion that a friend, spouse, parent, or other close family member has been replaced by an identical-looking impostor."[6]

  - In this context, it is clear that one instance of Sandy is interpreted from the speaker's perspective, call this $Sandy_{\sigma}$ (where σ is the speaker index) and the other from Kim's, call this $Impostor_{kim}$. The speaker is then asserting that Kim does not believe that $Impostor_{kim} = Sandy_{\sigma}$.
  - In a sense, then, this is a simple, limiting case of the puzzles we have been looking at.

- These kinds of expressions are not restricted to pathological cases. We can even construct similar examples involving mathematical terms, a domain that we would not expect to be open to interpretation in the same way, because mathematical terms are paradigmatically assumed to denote constant intensions.

- Consider the following piece of American history:

  (23) **Context:** In 1897 Dr. Edwin J. Goodwin presented a bill to the Indiana General Assembly for "[...] introducing a new mathematical truth and offered as a contribution to education to be used only by the State of Indiana free of cost". He had copyrighted that $\pi = 3.2$ and offered this "new mathematical truth" for free use to the State of Indiana (but others would have to pay to use it).[7]

  - At the appropriate historical juncture, it is clear that the following sentence had a satisfiable reading:

    (24) Dr. Goodwin doesn't believe that pi is pi.

  - Dr. Goodwin was a huckster, but given the context, it seems that (24) accurately reported his beliefs.
  - It may be tempting to explain these facts in terms of a *de re*/*de dicto* distinction based on compositional scope, as in Montague Semantics (Montague 1973), but that kind of analysis is too permissive and generates readings that are unavailable.
  - For example, suppose Dr. Goodwin had had a rival — let's call him Dr. Badwin — in the Indiana General Assembly trying to push an alternative bill proposing that π equals 3.15.
  - A *de re*/*de dicto* analysis in this case would generate a reading for (24) such that there is at least one belief world of Dr. Goodwin's in which true π does not equal 3.15.

[6]Wikipedia: `http://en.wikipedia.org/wiki/Capgras_delusion`
[7]Wikipedia: `http://en.wikipedia.org/wiki/Indiana_Pi_Bill`

  - But this is too weak an interpretation for Dr. Goodwin's actual beliefs: *none* of his belief worlds are such that he believes π is 3.15 — that's Dr. Badwin's belief.
  - In other words, there is a stronger requirement on compositional interpretation than we would get, in the general case, by simply treating terms as ambiguous, where the ambiguity is a property of the term itself.
  - Rather, they are potentially ambiguous in different ways *for different speakers*.

- One way to capture this is our method of allowing interpretation to be anchored to different agents' potentially differing perspectives.

3.4 Summary

- Intuitively, what the Capgras and Indiana Pi Bill cases share is a mix of the speaker's perspective with some other perspective: that of the subject of the sentence.
- Thus, it seems to us that the key to these puzzles, as mentioned above, is a notion of perspective, which can also potentially explain the lack of substitutability in simple sentences involving verbs like *love* and *murder*.
- If we cross the factors of same/distinct terms with simple/embedded context, we obtain the space of explananda in Table 1, with cells filled by examples from the previous sections.

	Simple	**Embedded**
Same term	#Dr. Octopus punched Spider-Man but he didn't punch Spider-Man.	Kim doesn't believe Sandy is Sandy.
Distinct term	Mary Jane loves Peter Parker but she doesn't love Spider-Man. #Dr. Octopus killed Spider-Man but he didn't kill Peter Parker.	Kim doesn't believe Hesperus is Phosphorus.

Table 1: The problem space

4 Formalization: Monads

- Our formal proposal is a conservative extension of the simply-typed lambda calculus that allows us to model expressions that involve perspectives.
- For further details, see the appendix at the end of the handout, or better yet, the paper itself (for a link, see the footnote on the first page of the handout).
- Here I try to give only very little formal detail and focus instead on the main formal intuitions.
- Our extension is derived from previous work in the semantics of programming languages aimed at providing a mathematical characterization of computations that produce some kind of *side effect* (Moggi 1989), and is based on the notion of *monads*, which we have used in a number of previous papers to model analyses of natural language meaning (Giorgolo and Asudeh 2011, 2012a,b, 2014a,b), based on the pioneering work of Shan (2001); see also Charlow (2014).[8]
- Monads are a construction in category theory that defines a canonical way to map a set of objects and functions that we may consider simple in some sense into a more complex object and function space.
- Monads have been successfully used in the semantics of programming languages to characterise certain classes of computation (Moggi 1989, Wadler 1992, 1995).
- A monad is defined as a triple $\langle \Diamond, \eta, \star \rangle$.
    - $\Diamond$ is a *functor*.[9] In general a functor is a mapping between two categories that relates the objects that form the first category to the objects of the second one and also maps the morphisms that connects the objects of the first category into morphisms operating in the second category.
    - η takes a non-monadic value and maps it to a monadic value.
    - $\star$ (pronounced 'bind') is a polymorphic function and acts as a sort of enhanced functional application.
        - Using a computational metaphor, $\star$ takes care of combining the side effects of the argument and the function and returns the resulting computation.
- In our case we will use the monad that describes values that are made dependent on some external factor, commonly known in the functional programming literature as the **Reader** monad, following Shan (2001), who suggested the idea of using the Reader monad to model intensional phenomena in natural language.
- We will represent linguistic expressions that can be assigned potentially different interpretations as functions from perspective indices to values.[10]
- Effectively we will construct a kind of lexicon that not only represents the linguistic knowledge of a single speaker but also her (possibly partial) knowledge of the language of other speakers.
- In other words, we construe lexicons to be aspects of the knowledge of language of *individuals*, and take standard circumlocutions like the "lexicon of English" to be atheoretical folk talk, if not simply incoherent (Chomsky 1965, 1986, 2000, Jackendoff 1983, 1997, 2002, 2007).

[8] Monads are related to *continuations* (Wadler 1994), which are the formal tool used in a rich body of work by Chris Barker and Ken Shan (see Barker and Shan 2014 and references to their antecedent work therein). Charlow (2014) provides a particularly insightful study of the interplay of monads and continuations as applied to natural language semantics.

[9] A small note about notation: We use the symbol $\Diamond$ because it is the one used by Benton et al. (1998) in the logical system that is the inspiration for our formalization. The reader should not confuse this symbol with the more familiar possibility modality, although as Benton et al. note their logic does indeed define a notion of possibility, although one that is somewhat different from the usual one, as already observed by Curry (1952).

[10] Our indices should not be confused with those of Fiengo and May (1998) or with the kinds of indices that are commonly used in Logical Form semantics (Heim and Kratzer 1998) or in binding theory (Büring 2005). We return to a comparison of our indices to those of Fiengo and May in section 6.

- So, for example, we will claim that examples like the Capgras example (21) or the similar following example can be assigned a non-contradictory reading:

  (25) Reza doesn't believe Jesus is Jesus.

- This example is based on the controversy from the summer of 2013 in which the scholar Reza Aslan was taken to task by Fox News correspondent Lauren Green for his views about the historical figure of Jesus of Nazareth. It seems to us that (25) could have been said by Green in that context.[11]
- The speaker's lexicon also includes the information regarding Reza's interpretation of the name *Jesus* and therefore makes it possible for the speaker to use the same expression, in combination with a verb such as *believe*, to actually refer to two different entities.
- In one case we will argue that the name *Jesus* is interpreted using the speaker's interpretation while in the other case it is Reza's interpretation that is used.

[11] Wikipedia: `http://en.wikipedia.org/wiki/Reza_Aslan`

Word	Denotation	Type
Reza	$\mathbf{r}_\sigma$	e
Kim	$\mathbf{k}_\sigma$	e
Dr. Octopus	$\mathbf{o}_\sigma$	e
Mary Jane	$\mathbf{mj}_\sigma$	e
Peter Parker	$\mathbf{pp}_\sigma$	e
not	$\lambda p.\neg p$	$t \to t$
but	$\lambda p.\lambda q.p \wedge q$	$t \to t \to t$
is	$\lambda x.\lambda y.x = y$	$e \to e \to t$
punch	$\lambda o.\lambda s.\mathbf{punch}(s, o)$	$e \to e \to t$
believe	$\lambda c.\lambda s.\mathbf{B}(s, c(\kappa(s)))$	$\Diamond t \to e \to t$
love	$\lambda o.\lambda s.\mathbf{love}(s, o(\kappa(s)))$	$\Diamond e \to e \to t$
Hesperus	$\lambda i. \begin{cases} \mathbf{es}_{\mathtt{k}} & \text{if } i = \mathtt{k}, \\ \mathbf{v}_\sigma & \text{if } i = \sigma \end{cases}$	$\Diamond e$
Phosphorus	$\lambda i. \begin{cases} \mathbf{ms}_{\mathtt{k}} & \text{if } i = \mathtt{k}, \\ \mathbf{v}_\sigma & \text{if } i = \sigma \end{cases}$	$\Diamond e$
Spider-Man	$\lambda i. \begin{cases} \mathbf{sm}_i & \text{if } i = \mathtt{o} \text{ or } i = \mathtt{mj}, \\ \mathbf{pp}_\sigma & \text{if } i = \sigma \end{cases}$	$\Diamond e$
Jesus	$\lambda i. \begin{cases} \mathbf{j}_{\mathtt{r}} & \text{if } i = \mathtt{r}, \\ \mathbf{j}_\sigma & \text{if } i = \sigma \end{cases}$	$\Diamond e$
Sandy	$\lambda i. \begin{cases} \mathbf{imp}_{\mathtt{k}} & \text{if } i = \mathtt{k}, \\ \mathbf{s}_\sigma & \text{if } i = \sigma \end{cases}$	$\Diamond e$

Table 2: Speaker's lexicon.

5 Analysis

- We will exemplify our approach with analyses of a selection of the examples discussed above, repeated here for convenience:

(26) Kim doesn't believe Hesperus is Phosphorus.

(27) #Dr. Octopus punched Spider-Man but he didn't punch Spider-Man.

(28) Mary Jane loves Peter Parker but she doesn't love Spider-Man.

(29) Kim doesn't believe Sandy is Sandy.

- The starting point for our analysis of these examples is the lexicon in Table 2.
  - The lexicon represents the linguistic knowledge of the speaker, including her assumptions about other individuals' grammars.[12]
  - Most lexical entries are standard, since we do not have to generalize to the worst case.
  - So we do not need to change the type and denotation of lexical items that are not involved in the phenomena under discussion.
  - For instance, logical operators such as *not* and *but* are interpreted in the standard way, as is a verb like *punch* or *kill*.
  - Referring expressions that are possibly contentious, in the sense that they can be interpreted differently by the speaker and other individuals, instead have the monadic type $\Diamond e$. This is reflected in their denotation by the fact that their value varies according to an interpretation index.
  - We use a special index σ to refer to the speaker's own perspective, and assume that this is the default index used whenever no other index is specifically introduced. For example, in the case of the name *Spider-Man*, the speaker is aware of his secret identity and therefore interprets it as another name for the individual Peter Parker, while Mary Jane and Dr. Octopus consider Spider-Man a different entity from Peter Parker.
  - We assume an internalist/representationalist semantics such that sentences are interpreted in a model in which all entities are mental entities, i.e. that there is no direct reference to entities in the world, but only to mental representations.
  - Entities are therefore relativized with respect to the individual that mentally represents them, where entities that the speaker believes to be non-contentious are always relativized according to the speaker.
  - This allows us to represent the fact that different individuals may have distinct equivalencies between entities. For example, Kim in our model does not equate the evening star and the morning star, but the speaker equates them with each other and with Venus.
  - Therefore, the speaker's lexicon in Table 2 represents the fact that the speaker's epistemic model includes what the speaker knows about other individuals' models, e.g. that Kim has a distinct denotation (from the speaker) for Hesperus, that Mary Jane has a distinct representation for Spider-Man, that Kim has a distinct representation for Sandy, etc. [13]
  - We should stress that this stance is not a necessary stance for our formal theory, but we think it is a sensible one, despite its potentially controversial nature.
  - With respect to our formal theory, it does not matter what the model for interpretation is a model *of*, whether a mental representation or reality.

---

[12] We have simplified some entries in Table 2 by writing, e.g., '$\mathbf{ms}_{\mathtt{k}}$ if $i = \mathtt{k}$' instead of '$\mathbf{ms}_i$ if $i = \mathtt{k}$', where there are not multiple options for i. For example, contrast the entry of *Phosphorus* with that of *Spider-Man*.

[13] The speaker's Kim-denotation of *Sandy* is then not plausibly Kim's actual denotation — a mental representation that would seem privileged to Kim — but rather the speaker's representation of that representation.

- However, it is not immediately clear how to make sense of the notion of distinct denotations without a representational layer, especially in the Capgras or Jesus cases. For example, in the Reza Aslan case, Aslan and Lauren Green were not in disagreement about which historical figure they were referring to, but rather about which properties that very same person had; see page 10 above.
- The other special lexical entries in our lexicon are those for verbs like *believe* and *love*.
- The two entries are similar in the sense that they both take an already monadic resource and actively supply a specific interpretation index that corresponds to the subject of the verb.
- The function κ maps each entity to the corresponding interpretation index, i.e.:

$$\kappa : e \to i \tag{30}$$

- For example, in the lexical entries for *believe* and *love*, κ maps the subject to the interpretation index of the subject.
- Thus, the entry for *believe* uses the subject's point of view as the perspective used to evaluate its entire complement, while *love* changes the interpretation of its object relative to the perspective of its subject.
- However we will see that the interaction of these lexical entries and the evaluation order imposed by $\star$ will allow us to let the complement of a verb like *believe* and the object of a verb like *love* escape the specific effect of forcing the subject perspective, and instead we will be able to derive readings in which the arguments of the verb are interpreted using the speaker's perspective.

- We present a small fragment of our analysis in Figures 2, 3, and 4 in the appendix, but review the results a bit less formally here.

- Figure 2 in the appendix reports the four non-equivalent readings that we derive in our system for example (26), repeated here as (31).

(31) Kim doesn't believe that Hesperus is Phosphorus.

- The first reading assigns to both Hesperus and Phosphorus the subject interpretation and results, after contextualizing the sentence by applying it to the standard σ interpretation index, in the truth conditions in (32), i.e. that Kim does not believe that the evening star is the morning star.
- This reading would not be contradictory in an epistemic model (such as Kim's model) where the evening star and the morning star are not the same entity.

$$\neg\mathbf{B}(\mathbf{k})(\mathbf{es}_k = \mathbf{ms}_k) \tag{32}$$

- In the case of the second and third readings, we get a similar effect, although here we mix the epistemic models of the speaker and Kim: one of the referring expressions is interpreted under the speaker perspective while the other is again interpreted under Kim's perspective.
- For these two readings we obtain respectively the truth conditions in (33) and (34).

$$\neg\mathbf{B}(\mathbf{k})(\mathbf{v}_\sigma = \mathbf{ms}_k) \tag{33}$$
$$\neg\mathbf{B}(\mathbf{k})(\mathbf{v}_\sigma = \mathbf{es}_k) \tag{34}$$

- Finally for the fourth reading, (68), we get the contradictory reading that Kim does not believe that Venus is Venus, as both referring expressions are evaluated using the speaker's interpretation index.

$$\neg\mathbf{B}(\mathbf{k})(\mathbf{v}_\sigma = \mathbf{v}_\sigma) \tag{35}$$

- If we consider a case like sentence (27), repeated in (36), we ought to get only a contradictory reading as the statement is intuitively contradictory.

(36) #Dr. Octopus punched Spider-Man but he didn't punch Spider-Man.

- ∘ Our analysis produces a single reading that indeed corresponds to a contradictory interpretation:

$$[\![\text{Spider-Man}]\!] \star \lambda x. [\![\text{Spider-Man}]\!] \star \lambda y.\eta([\![\text{but}]\!]([\![\text{punch}]\!]([\![\text{Dr. Octopus}]\!])(x))([\![\text{not}]\!]([\![\text{punch}]\!]([\![\text{Dr. Octopus}]\!])(y)))) \tag{37}$$

- ∘ The verb *punch* is not a verb that can change the interpretation perspective and therefore the potentially controversial name *Spider-Man* is interpreted in both instances using the speaker's interpretation index. The result are unsatisfiable truth conditions, as expected:

$$\mathbf{punch}(\mathbf{o}_\sigma)(\mathbf{pp}_\sigma) \wedge \neg\mathbf{punch}(\mathbf{o}_\sigma)(\mathbf{pp}_\sigma) \tag{38}$$

- In contrast a verb like *love* is defined in the lexicon in Table 2 as possibly changing the interpretation perspective of its object to that of its subject.
- Therefore in the case of a sentence like (28), repeated in (39), we expect one reading where the potentially contentious name *Spider-Man* is interpreted according to the subject of *love*, Mary Jane.

(39) Mary Jane loves Peter Parker but she doesn't love Spider-Man.

- ∘ This is in fact the result we obtain. Figure 3 reports the two readings that our framework generates for (39).
- ∘ The first reading corresponds to the non-contradictory interpretation of sentence (39), where Spider-Man is interpreted according to Mary Jane's perspective and therefore is assigned an entity different from Peter Parker:

$$\mathbf{love}(\mathbf{mj}_\sigma)(\mathbf{pp}_\sigma) \wedge \neg\mathbf{love}(\mathbf{mj}_\sigma)(\mathbf{sm}_{mj}) \tag{40}$$

- ∘ The second reading instead generates unsatisfiable truth conditions, as Spider-Man is identified with Peter Parker according to the speaker's interpretation:

$$\mathbf{love}(\mathbf{mj}_\sigma)(\mathbf{pp}_\sigma) \wedge \neg\mathbf{love}(\mathbf{mj}_\sigma)(\mathbf{pp}_\sigma) \tag{41}$$

- Our last example, the Capgras example (29), repeated here as (42), is particularly interesting as the embedded clause is just a simple identity statement with two tokens of the same name. We are not aware of direct discussion of this kind of example in the literature.

- ∘ The non-contradictory reading that this sentence has seems to be connected specifically to two different interpretations of the same name, *Sandy*, both under the syntactic scope of the modal *believe*.

  (42) Kim doesn't believe Sandy is Sandy.
- ∘ Our system generates three non-equivalent readings, reported here in Figure 4.
- ∘ The first two readings correspond to two contradictory readings of the sentence: in the first case both instances of the name *Sandy* are interpreted from the subject perspective and therefore attribute to Kim the non-belief in a tautology, similarly in the second case, even though in this case the two names are interpreted from the perspective of the speaker.
- ∘ In contrast, the third reading corresponds to the interpretation that assigns two different referents to the two instances of the name Sandy, producing the truth conditions in (43) which are satisfiable in a suitable model.

$$\neg\mathbf{B}(\mathbf{k})(\mathbf{s}_\sigma = \mathbf{imp}_k) \tag{43}$$

  We use $\mathbf{imp}_k$ as the speaker's representation of the "impostor" that Kim thinks has taken the place of Sandy.

- The analysis of the Jesus example (25), repeated in (44), is equivalent; the non-contradictory reading is shown in (45).

(44) Reza doesn't believe Jesus is Jesus.

$$\neg \mathbf{B}(\mathbf{r})(\mathbf{j}_\sigma = \mathbf{j}_r) \tag{45}$$

  - More generally, there are again three non-equivalent readings, including the one above, which are just those in Figure 4, with ⟦Sandy⟧ replaced by ⟦Jesus⟧ and ⟦Kim⟧ replaced by ⟦Reza⟧.

6 Comparison with Previous Approaches

- Our approach seems superficially similar to Fiengo and May's, since it uses indices too, but it is in fact quite distinct.

1.
   - The use of indices here resides entirely in the model.
   - We do not require that the expression '[$_{\text{NP}}$ Sandy]' bear an index in order to address the puzzles above.[14]
2.
   - We are not forced to agree with Fiengo and May (1998: 381) that 'there is no value associated with the name 'Max' qua lexical item', despite agreeing with their claim that 'if we are to determine the identity conditions for words ... we should determine the identity conditions on words as *they appear in the lexicon of an individual*.' (emphasis in the original; Fiengo and May 1998: 379).[15]
   - Fiengo and May thus claim that the name *Max* in the lexicon does not refer, only the syntactic structure [$_{\text{NP}_i}$ Max] refers.
   - This claim is not unreasonable, although it's certainly hard to see how to test it empirically.
   - However, it does not seem to fit our folk understanding of names, which do seem referential in their own right, as words.
   - Now, we are perfectly comfortable with folk understanding giving way to theory, but if two theories are equal in all other respects, it would seem to us slightly capricious, if not outright perverse, to pick the one that rejects folk understanding of its phenomena.
   - On our theory, a name like *Max* does have its (range of) reference determined in the (speaker's) lexicon.
   - There is some sense, then, in which our theory is more natural.
3.
   - It is well known that these puzzles are not just about names, but also about natural kinds and other predicates.
   - It is hard to see how the kinds of indices that Fiengo and May (1998) use could be generalized to cover such cases; they seem too narrowly conceived to do the job, since they represent reference to individuals and this is not appropriate for predicates.
   - In contrast, a virtue of our analysis is that we can apply it to not just names and referring expressions, but to any natural language expression that may have different perspectival interpretations.
   - This means that we can extend our analysis to other cases, such as the standard examples involving synonymous natural kind terms like *groundhog* and *woodchuck* (see, e.g., Fox and Lappin 2005) or *furze* and *gorse* (Kripke 1979) or of synonymous verbs, such as *photocopy* and *xerox* (Larson and Ludlow 1993).

[14] This does not mean that there could not be other, independent reasons for expressions to bear indices, although we are most sympathetic to the spirit of David Beaver's question, 'What do those little numbers mean, and who put them there anyway?' (Beaver 1999).

[15] The full quote has 'in particular, names' where we have the ellipsis, but we are not in fact convinced that names are special in this regard, for reasons that will become clear shortly.

Word	Denotation	Type
dolphin	$\{\mathbf{flipper}_\sigma\}$	$e \to t$
seal	$\{\mathbf{hoover}_\sigma\}$	$e \to t$
marine mammal	$\lambda i.\begin{cases}\{\mathbf{hoover}_\sigma\} & \text{if } i = \mathbf{mj},\\ \{\mathbf{flipper}_\sigma, \mathbf{hoover}_\sigma\} & \text{if } i = \sigma\end{cases}$	$\Diamond(e \to t)$

Table 3: (Relevant portion of) speaker's lexicon for (46)

- As an illustration, consider the following example:

(46) Elena loves dolphins, but she doesn't love marine mammals.

  - Suppose Elena thinks that Flipper is a dolphin and Hoover is a seal, but she thinks only Hoover is a marine mammal; i.e., she thinks seals are marine mammals, but dolphins are not.
  - Suppose also that the speaker and Elena are in agreement about which entities the names *Flipper* and *Hoover* refer to, so the names are not controversial.
  - Table 3 sketches (the relevant part of) the lexicon for the speaker of (46).
  - We do not mean to imply that this extension of our approach is trivial, since matters of compositionality of, e.g., *marine mammal*, have not been addressed here, but the extension is at least a natural candidate for further exploration.

- Lastly, let us return to the matter of traditional approaches to substitutability/opacity based on *de re*/*de dicto* ambiguities derived from differential compositional scopings.

- In section 3.3, we discussed certain problems that such an approach may have for ordinary proper names in Capgras and Indiana Pi Bill examples, repeated here:

(47) Kim doesn't believe Sandy is Sandy.

(48) Dr. Goodwin doesn't believe pi is pi.

- Here we outline some further issues for composition-based scope approaches.

- First, substitutability puzzles in simple sentences, as discussed in section 3.1, pose a prima facie problem, since there does not seem to be a relevant scope operator present in the case of the composition of a simple predicate like *love* or *murder* with its object argument.

- Of course, many linguists would be perfectly willing to postulate null operators in such cases, but it is not clear this derives the right result.

- That would be tantamount to treating *love* or *murder* as an 'opaque transitive verb', but the former have existential entailments that the latter lack (Zimmermann 2006):

(49) Frodo owes Sam a horse.
$\not\to$ There is a horse that Frodo owes Sam.

(50) Saruman murdered a horse.
$\to$ There was a horse that Saruman murdered.

(51) Princess Carolyn loves a horse.
$\to$ There is a horse that Princess Carolyn loves.

- An account that postulates a hidden operator for *murder* or *love* would have to explain this contrast.
- Second, even in the case of embedded contexts, which offer a scopal operator in composition, in the form of a modal or propositional attitude verb, a scope-based *de re*/*de dicto* approach faces some problems, at least in the case of our most challenging examples, such as the Capgras example.
- To try to explain the two readings in the context of a standard possible worlds semantics, we could take (47) to be ambiguous with respect to a *de re*/*de dicto* reading.
- In the case of the *de dicto* reading (which corresponds to the unsatisfiable reading) the two names are evaluated under the scope of the doxastic operator *believe*, i.e. they both refer to the same entity that is assigned to the name *Sandy* in each accessible world.[16]
- Clearly this is always the case, and so (47) is not satisfiable. In the case of the *de re* reading, we assume that the two names are evaluated at different worlds that assign different referents to the two names.
- One of these two worlds will be the actual world and the other one of the accesible worlds.
- The reading is satisfiable if the doxastic modality links the actual world with one in which the name *Sandy* refers to a different entity.
- Notice that for this analysis to work we need to assume that names behave like quantifiers with respect to scoping both over and under modal and propositional attitude operators, as discussed in section 3.3.
- However, such an approach would nevertheless face some problems.

1. ◦ It has been argued that even if we model names as generalized quantifiers, they are scopeless (Zimmermann 1993).
   ◦ But this is problematic for a scopal approach to the Capgras example.
   ◦ It would predict that both instances of the name *Sandy* escape the scope of *believe*.
   ◦ The resulting reading would bind the quantified individual to the interpretation of *Sandy* in the actual world.
   ◦ This would capture only an unsatisfiable reading.
   ◦ To save the scopal approach, we would need to assume that names in fact are sometimes interpreted in the scope of operators.
2. ◦ Even assuming that we find a satisfactory solution for these problems, the scopal approach cannot really capture the intuitions behind opacity in all contexts.
   ◦ Consider again our Jesus example, repeated here:
   (52) Reza doesn't believe Jesus is Jesus.
   ◦ Assume that there are two views about Jesus: Jesus as a divine being and Jesus as a human being. Assume that Jesus is a human being in the actual world and that Reza is an atheist,[17] then the only possible reading is the unsatisfiable one, as the referent for Jesus will be the same in the actual world and all accessible Reza-belief-worlds.
   ◦ The problem is that the scopal approach assumes a single modal model, while in this case it seems that there are two doxastic belief models necessary, Reza's model and the speaker's model.
   ◦ In contrast, in our approach the relevant part of Reza's model is embedded inside the speaker's model and interpretation indices indicate which interpretation belongs to Reza and which to the speaker.

[16]In order to put the discussion on firmer footing, we have adopted the language of a Hintikka-like analysis of propositional attitudes (Hintikka 1969), since this view is currently influential in linguistics, but our points should follow without a substantive commitment to that sort of analysis, so long as the propositional attitude verb is assumed to provide a scope point, as is commonly assumed.

[17]Alternatively, assume that Jesus is a divine being and Reza is a Christian.

7 Conclusion

- We have offered a semantics of perspective that offers a solution to the substitutability puzzle in both simple and embedded contexts.
  - Our solution extends to cases of distinct interpretations of tokens of the same name, which gives rise to a related puzzle. We exemplified this case with respect to simple identity cases, as in the Capgras, Indiana Pi Bill, and Aslan/Jesus examples.
  - Our solution to these puzzles rests on an analysis in terms of a combination of different perspectives.
  - We have claimed that the switch to a different perspective is triggered by specific lexical items, such as propositional attitude verbs, but also verbs like *love* and *murder* which express some kind of perspective on the part of the subject of the verb towards its object, but which nevertheless cannot easily be argued to be opaque in their object position.
  - The context switch is not obligatory, as witnessed by the multiple readings that the sentences discussed seem to have.
- The formalization of our analysis is based on monads. The main idea of our formal implementation is that referring expressions that have a potential perspectival dependency can be implemented as functions from perspective indices to fully interpreted values.
  - Similarly, the linguistic triggers for context switch are implemented in the lexicon as functions that can modify the interpretation context of their arguments.
  - Monads allow us to freely combine these "enriched" meanings with standard ones, avoiding unilluminating generalization to the worst case.
- We have also seen how more traditional approaches, while capable of dealing with some of the examples we discuss, are not capable of providing a general explanation of the totality of observed phenomena.
- We briefly explored how our approach could be extended to other types of natural language expressions, such as natural kind terms, nominal predicates, and verbs. Careful exploration of these extensions remains part of future work.
- We have inevitably had to take positions on some issues that are far from settled, but we do not mean these positions themselves to be our main contribution.
- Rather, it seems to us that philosophers and linguists are in broad agreement that in some linguistic contexts there seems to be an "extra something" involved in interpreting names, and other expressions; we have made a formal proposal about what that extra something could be: perspectives.

References

Ades, Anthony, and Mark Steedman. 1982. On the order of words. *Linguistics and Philosophy* 4: 517–558.

Asudeh, Ash. 2012. *The Logic of Pronominal Resumption*. Oxford: Oxford University Press.

Asudeh, Ash, and Richard Crouch. 2002. Glue Semantics for HPSG. In Frank van Eynde, Lars Hellan, and Dorothee Beermann, eds., *Proceedings of the 8th International HPSG Conference*, 1–19. Stanford, CA: CSLI Publications.

Asudeh, Ash, and Gianluca Giorgolo. 2016. Perspectives. *Semantics and Pragmatics* 9(21).

Austin, John L. 1975. *How to Do Things with Words*. Oxford: Oxford University Press, 2nd edn.

Barker, Chris, and Chung-Chieh Shan. 2014. *Continuations and Natural Language*. Oxford: Oxford University Press.

Beaver, David. 1999. Pragmatics (to a First Approximation). In Jelle Gerbrandy, Maarten Marx, Maarten de Rijke, and Yde Venema, eds., *JFAK — Essays Dedicated to Johan van Benthem on the Occasion of his 50th Birthday*. Amsterdam University Press.

Benton, Nick, Gavin M. Bierman, and Valeria de Paiva. 1998. Computational Types from a Logical Perspective. *Journal of Functional Programming* 8(2): 177–193.

Braun, David, and Jennifer Saul. 2002. Simple Sentences, Substition, and Mistaken Evaluations. *Philosophical Studies* 111(1): 1–41.

Bresnan, Joan, Ash Asudeh, Ida Toivonen, and Stephen Wechsler. 2016. *Lexical-Functional Syntax*. Malden, MA: Wiley-Blackwell., 2nd edn.

Büring, Daniel. 2005. *Binding Theory*. Cambridge: Cambridge University Press.

Burke, Luke Edward. 2015. Glue TAG Semantics for Binary Branching Syntactic Structures. Master's thesis, Birkbeck College.

Carnap, Rudolf. 1947. *Meaning and Necessity: A Study in Semantics and Modal Logic*. Chicago, IL: University of Chicago Press.

Carpenter, Bob. 1998. *Type-Logical Semantics*. Cambridge, MA: MIT Press.

Castañeda, Héctor-Neri. 1972. Thinking and the Structure of the World: *Discours d'ontologie*. *Crítica: Revista Hispanoamericana de Filosofía* 6(18): 43–86.

—. 1989. *Thinking, Language, and Experience*. Minneapolis: University of Minnesota Press.

Charlow, Simon. 2014. On the Semantics of Exceptional Scope. Ph.D. thesis, New York University.

Chierchia, Gennaro. 1989. Anaphora and Attitudes *De Se*. In Renate Bartsch, Johan van Benthem, and Peter van Emde Boas, eds., *Language and Contextual Expressions*, 1–31. Dordrecht: Foris.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

—. 1986. *Knowledge of Language: Its nature, origin, and use*. New York, NY: Praeger.

—. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

—. 2000. *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

Curry, Haskell B. 1952. The Elimination Theorem when Modality is Present. *Journal of Symbolic Logic* 4(17): 249–265.

Curry, Haskell B., and Robert Feys. 1958. *Combinatory Logic*, vol. 1. Amsterdam: North-Holland.

Dalrymple, Mary, ed. 1999. *Semantics and Syntax in Lexical Functional Grammar: The Resource Logic Approach*. Cambridge, MA: MIT Press.

Dalrymple, Mary. 2001. *Lexical Functional Grammar*. San Diego, CA: Academic Press.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell III, and Annie Zaenen, eds. 1995. *Formal Issues in Lexical-Functional Grammar*. Stanford, CA: CSLI Publications.

de Groote, Philippe, ed. 1995. *The Curry-Howard Isomorphism*, vol. 8 of *Cahiers du Centre de Logique*. Louvain-la-neuve, Belgium: Academia.

Ernst, Thomas. 2009. Speaker-Oriented Adverbs. *Natural Language and Linguistic Theory* 27(3): 497–544.

Fiengo, Robert, and Robert May. 1998. Names and Expressions. *Journal of Philosophy* 95(8): 377–409.

Fox, Chris, and Shalom Lappin. 2005. *Foundations of Intensional Semantics*. Oxford: Blackwell.

Frege, Gottlob. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100: 25–50.

Giorgolo, Gianluca, and Ash Asudeh. 2011. Multidimensional Semantics with Unidimensional Glue Logic. In Miriam Butt and Tracy Halloway King, eds., *Proceedings of the LFG11 Conference*, 236–256. CSLI Publications.

—. 2012a. $\langle M, \eta, \star \rangle$ Monads for Conventional Implicatures. In Ana Aguilar Guevara, Anna Chernilovskaya, and Rick Nouwen, eds., *Proceedings of Sinn und Bedeutung 16*, vol. 1, 265–278. MIT Working Papers in Linguistics.

—. 2012b. Missing Resources in a Resource-Sensitive Semantics. In Miriam Butt and Tracy Holloway King, eds., *Proceedings of the LFG12 Conference*, 219–239. Stanford, CA: CSLI Publications.

—. 2014a. Monads as a Solution for Generalized Opacity. In *Proceedings of the EACL 2014 Workshop on Type Theory and Natural Language Semantics (TTNLS)*, 19–27. Association for Computational Linguistics, Gothenburg.

—. 2014b. One Semiring to Rule Them All. In *Proceedings of the 36th Annual Cognitive Science Society Confer-*

ence, 116–121.

Girard, Jean-Yves. 1987. Linear Logic. *Theoretical Computer Science* 50(1): 1–102.

Gutzmann, Daniel. 2012. Use-Conditional Meaning: Studies in Multidimensional Semantics. Ph.D. thesis, University of Frankfurt.

Gutzmann, Daniel, and Hans-Martin Gärtner, eds. 2013. *Beyond Expressives: Explorations in Use-Conditional Meaning*. Leiden: Brill.

Hagège, Claude. 1974. Les pronoms logophoriques. *Bulletin de la Société de Linguistique de Paris* 69: 287–310.

Heim, Irene. 1998. Anaphora and Semantic Interpretation: A Reinterpretation of Reinhart's Approach. In Uli Sauerland and Orin Percus, eds., *The Interpretive Tract*, vol. 25 of *MIT Working Papers in Linguistics*, 205–246. Cambridge, MA: MITWPL.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.

Hintikka, Jaakko. 1969. Semantics for Propositional Attitudes. In J.W. Davis, D.J. Hockney, and W.K. Wilson, eds., *Philosophical Logic*, 21–45. Dordrecht: Reidel.

Howard, William A. 1980. The Formulae-as-Types Notion of Construction. In Jonathan P. Seldin and J. Roger Hindley, eds., *To H.B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism*, 479–490. London: Academic press. Circulated in unpublished form from 1969. Reprinted in de Groote (1995: 15–26).

Jackendoff, Ray. 1983. *Semantics and Cognition*. Cambridge, MA: MIT Press.

—. 1997. *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.

—. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

—. 2007. *Language, Consciousness, Culture: Essays on Mental Structure*. Cambridge, MA: MIT Press.

Joshi, Aravind K., Leon S. Levy, and Masako Takahashi. 1975. Tree Adjunct Grammars. *Journal of Computer and Systems Sciences* 10(1): 136–163.

Kaplan, David. 1989. Demonstratives. In Joseph Almog, John Perry, and Harvey Wettstein, eds., *Themes from Kaplan*, 481–563. Oxford: Oxford University Press.

Kaplan, Ronald M., and Joan Bresnan. 1982. Lexical-Functional Grammar: A Formal System for Grammatical Representation. In Joan Bresnan, ed., *The Mental Representation of Grammatical Relations*, 173–281. Cambridge, MA: MIT Press. Reprinted in Dalrymple et al. (1995: 29–135).

Krifka, Manfred. 2001. Quantifying into Question Acts. *Natural Language Semantics* 9(1): 1–40.

—. 2014. Embedding Illocutionary Acts. In Thomas Roeper and Margaret Speas, eds., *Recursion: Complexity in Cognition*, 59–87. Springer International Publishing.

Kripke, Saul. 1972. Naming and Necessity. In Donald Davidson and Gilbert Harman, eds., *Semantics of Natural Language*, 253–355. Reidel.

—. 1979. A Puzzle about Belief. In Avishai Margalit, ed., *Meaning and Use*, 239–283. Dordrecht: Reidel.

—. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Kuno, Susumo. 1987. *Functional Syntax: Anaphora, Discourse and Empathy*. Chicago: University of Chicago Press.

Larson, Richard K., and Peter Ludlow. 1993. Interpreted Logical Forms. *Synthese* 95(3): 305–355.

Lasersohn, Peter. 2005. Context Dependence, Disagreement, and Predicates of Personal Taste. *Linguistics and Philosophy* 28(6): 643–686.

Lewis, David. 1979. Attitudes *De Dicto* and *De Se*. *Philosophical Review* 88(4): 513–543.

MacColl, Hugh. 1905a. Existential Import of Propositions. *Mind* 14: 401–402.

—. 1905b. Symbolic Reasoning. *Mind* 14: 490–397.

MacFarlane, John. 2014. *Assessment Sensitivity: Relative Truth and its Applications*. Oxford: Clarendon Press.

Mates, Benson. 1950. Synonymity. In Donald S. MacKay, George P. Adams, and William R. Dennes, eds., *Meaning and Interpretation: Lectures Delivered Before the Philosophical Union of the University of California, 1948–1949*, 201–226. Berkeley: University of California Press.

Moggi, Eugenio. 1989. Computational Lambda-Calculus and Monads. In *LICS*, 14–23. IEEE Computer Society Press.

Montague, Richard. 1973. The Proper Treatment of Quantification in Ordinary English. In Jaakko Hintikka, Julian Moravcsik, and Patrick Suppes, eds., *Approaches to Language*, 221–242. Dordrecht: Reidel. Reprinted

in Montague (1974: 247–270).
—. 1974. *Formal Philosophy: Selected Papers of Richard Montague*. New Haven, CT: Yale University Press. Edited and with an introduction by Richmond H. Thomason.
Moortgat, Michael. 1999. Constants of Grammatical Reasoning. In Gosse Bouma, Erhard W. Hinrichs, Geert-Jan M. Kruijff, and Richard T. Oehrle, eds., *Constraints and Resources in Natural Language Syntax and Semantics*, 199–219. Stanford, CA: CSLI Publications.
—. 2011. Categorial Type Logics. In Johan van Benthem and Alice ter Meulen, eds., *Handbook of Logic and Language*, 95–179. Elsevier, 2nd edn.
Morrill, Glyn V. 2011. *Categorial Grammar: Logical Syntax, Semantics, and Processing*. Oxford: Oxford University Press.
Oshima, David Y. 2006. Perspectives in Reported Discourse. Ph.D. thesis, Stanford University.
Pearson, Hazel. 2013. The Sense of Self: Topics in the Semantics of *De Se* Expressions. Ph.D. thesis, Harvard University.
Pitt, David. 2001. Alter Egos and Their Names. *The Journal of Philosophy* 98(10): pp. 531–552.
Pollard, Carl, and Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago, IL and Stanford, CA: The University of Chicago Press and CSLI Publications.
Potts, Christopher. 2005. *The Logic of Conventional Implicatures*. Oxford: Oxford University Press.
—. 2007. The Expressive Dimension. *Theoretical Linguistics* 33(2): 165–197.
Quine, Willard van Orman. 1953. *From a Logical Point of View*. Cambridge, MA: Harvard University Press. Revised edition 1980.
—. 1956. Quantifiers and Propositional Attitudes. *Journal of Philosophy* 53: 177–87.
—. 1960. *Word and Object*. Cambridge, MA: MIT Press.
Recanati, François. 1997. *Direct Reference: From Language to Thought*. Oxford: Blackwell.
Richard, Mark. 1990. *Propositional Attitudes: An Essay on Thoughts and How We Ascribe Them*. Cambridge: Cambridge University Press.
Russell, Bertrand. 1905. On Denoting. *Mind* 14: 479–493.
Saul, Jennifer. 2007. *Simple Sentences, Substitution, and Intuitions*. Oxford: Oxford University Press.
Saul, Jennifer M. 1997. Substitution and Simple Sentences. *Analysis* 57(2): 102–108.
Schiffer, Stephen. 1990. The Mode-of-Presentation Problem. In C. Anthony Anderson and Joseph Owens, eds., *Propositional Attitudes: The Role of Content in Logic, Language and Mind*, 249–268. Stanford, CA: CSLI Publications.
Schlenker, Philippe. 2003. A Plea for Monsters. *Linguistics and Philosophy* 26(1): 29–120.
Sells, Peter. 1987. Aspects of Logophoricity. *Linguistic Inquiry* 18(3): 445–479.
Shan, Chung-chieh. 2001. Monads for Natural Language Semantics. In Kristina Striegnitz, ed., *Proceedings of the ESSLLI-2001 Student Session*, 285–298. 13th European Summer School in Logic, Language and Information.
Stephenson, Tamina. 2007a. Judge Dependence, Epistemic Modals, and Predicates of Personal Taste. *Linguistics and Philosophy* 30(4): 487–525.
—. 2007b. Towards a Theory of Subjective Meaning. Ph.D. thesis, MIT.
Sundaresan, Sandhya. 2012. Context and (Co)rereference in the Syntax and its Interfaces. Ph.D. thesis, University of Tromsø and University of Stuttgart.
Wadler, Philip. 1992. Comprehending Monads. *Mathematical Structures in Computer Science* 2: 61–78.
—. 1994. Monads and Composable Continuations. *Lisp and Symbolic Computation* 7(1): 39–56.
—. 1995. Monads for Functional Programming. In Johan Jeuring and Erik Meijer, eds., *Advanced Functional Programming*, 24–52. Berlin: Springer.
Zimmermann, Thomas Ede. 1993. Scopeless Quantifiers and Operators. *Journal of Philosophical Logic* 22(5): 545–561.
—. 2006. Monotonicity in Opaque Verbs. *Linguistics and Philosophy* 29(6): 715–761.

A Further Formal Details

A.1 Monads

A.1.1 Interpretation Functions

To avoid introducing the complexities of the categorical formalism, we introduce monads as they are usually encountered in the computer science literature, as in our previous work (Giorgolo and Asudeh 2014a). A monad is defined as a triple $\langle \Diamond, \eta, \star \rangle$. $\Diamond$ is a *functor*, in our case a mapping between types and functions. We call the component of $\Diamond$ that maps between types $\Diamond_1$ and the one that maps between functions $\Diamond_2$. In our case, $\Diamond_1$ will map each type to a new type that corresponds to the original type with an added interpretation index parameter. Formally, if i is the type of interpretation indices, then $\Diamond_1$ maps any type τ to $i \to \tau$. The functor $\Diamond_2$ maps any function $f : \tau \to \delta$ to a function $f' : (i \to \tau) \to i \to \delta$. $\Diamond_2$ corresponds to function composition:

$$\Diamond_2(f) = \lambda g.\lambda i.f(g(i)) \tag{53}$$

The component $\Diamond_2$ will not be used below, so we will use $\Diamond$ as an abbreviation for $\Diamond_1$. This means that we will write $\Diamond\tau$ for the type $i \to \tau$.

η (pronounced 'unit') is a polymorphic function that maps inhabitants of a type τ to inhabitants of its image under $\Diamond$, formally $\eta : \forall\tau.\tau \to \Diamond\tau$. Using the computational metaphor, η should embed a value in a computation that returns that value without any side-effect. In our case η should simply add a vacuous parameter to the value:

$$\eta(x) = \lambda i.x \tag{54}$$

$\star$ (pronounced 'bind') is a polymorphic function of type $\forall\tau.\forall\delta.\Diamond\tau \to (\tau \to \Diamond\delta) \to \Diamond\delta$, and acts as a sort of enhanced functional application.[18] Again using the computational metaphor, $\star$ takes care of combining the side effects of the argument and the function and returns the resulting computation. In the case of the monad that we are interested in, $\star$ is defined as in (55).

$$a \star f = \lambda i.f(a(i))(i) \tag{55}$$

Another fundamental property of $\star$ is that, by imposing an order of evaluation, it will provide us with an additional scoping mechanism distinct from standard functional application. This will allow us to correctly capture the multiple readings associated with the expressions under consideration.

In sum, we add two operators, η and $\star$, to the lambda calculus and the reductions work as expected for (54) and (55). These reductions are implicit in our analyses in section 5.

A.1.2 Composition

For semantic composition we use a logical calculus adapted for the linear case[19] from the one introduced by Benton et al. (1998). The calculus is based on a language with two connectives corresponding to our type constructors: $\multimap$, a binary connective, that corresponds to (linear) functional types, and $\Diamond$, a unary connective, that represents monadic types.

The logical calculus is described by the proof rules in Figure 1.[20] The terms in the rules are reminiscent of terms in Glue Semantics (Dalrymple 1999, 2001, Asudeh 2012), which consist of a pairing of a term from a meaning language, formalized in the lambda calculus, and a term from a logic of composition, formalized in linear logic (Girard 1987). The correspondence between the linear logic terms and meaning terms is characterized by the Curry-Howard correspondence between proofs and terms (Curry and Feys 1958, Howard 1980, de Groote

[18] We use the argument order for $\star$ that is normally used in functional programming. We could swap the arguments and make it look more like standard functional application. We write $\star$ in infix notation.

[19] Various researchers have argued for linearity as a property of composition in natural language semantics (Moortgat 1999, Moortgat 2011, Asudeh 2012, among others). Asudeh (2012) discusses it under the rubric of 'resource sensitivity'.

[20] We can prove that the *Cut* rule is admissible, therefore the calculus becomes an effective (although inefficient) way of computing the meaning of a linguistic expression.

$$\frac{}{x : A \vdash x : A}\ id \qquad \frac{\Gamma \vdash B \quad B, \Delta \vdash C}{\Gamma, \Delta \vdash C}\ Cut$$

$$\frac{\Gamma, x : A \vdash t : B}{\Gamma \vdash \lambda x.t : A \multimap B} \multimap R \qquad \frac{\Delta \vdash t : A \quad \Gamma, x : B \vdash u : C}{\Gamma, \Delta, y : A \multimap B \vdash u[y(t)/x] : C} \multimap L$$

$$\frac{\Gamma \vdash x : A}{\Gamma \vdash \eta(x) : \Diamond A}\ \Diamond R \qquad \frac{\Gamma, x : A \vdash t : \Diamond B}{\Gamma, y : \Diamond A \vdash y \star \lambda x.t : \Diamond B}\ \Diamond L$$

Figure 1: Sequent calculus for a fragment of multiplicative linear logic enriched with a monadic modality, together with a Curry-Howard correspondence between formulae and meaning terms.

1995). For example, the linear implication, $\multimap$, functions like an undirected slash, $|$, in Categorial Grammar (Ades and Steedman 1982, Carpenter 1998, Morrill 2011): *modus ponens*/elimination for the implication corresponds to functional application. It is assumed that the logical calculus is associated with some syntactic representation, but the exact nature of the underlying syntactic formalism is not strictly relevant, so long as it can instantiate the terms for semantic composition.[21]

A key advantage of the monadic approach is that we are not forced to generalize all lexical entries to the "worst case", or richest type (as in, e.g., standard Montague Semantics or the semantics sketched in section **??**). With the logical setup we have just described we can freely mix monadic and non-monadic terms. For example, we can combine a pure version of a binary function with arguments that are either pure or monadic, as the following are all provable theorems in our logic.

$$A \multimap B \multimap C, A, B \vdash \Diamond C \tag{56}$$
$$A \multimap B \multimap C, \Diamond A, B \vdash \Diamond C \tag{57}$$
$$A \multimap B \multimap C, A, \Diamond B \vdash \Diamond C \tag{58}$$
$$A \multimap B \multimap C, \Diamond A, \Diamond B \vdash \Diamond C \tag{59}$$

In contrast, the following is not a theorem in the logic:

$$A \multimap B \multimap C, I \multimap A, I \multimap B \not\vdash I \multimap C \tag{60}$$

In short, if we were to instead simply lift the type of the lexical terms whose interpretation may be dependent on a specific perspective, we would be forced to lift all linguistic expressions that may combine with them, thus generalizing to the worst case. We do not have to do this, given our logic.

The monadic machinery also achieves a higher level of compositionality. In principle we could directly encode our monad using the $\rightarrow$ type constructor. However this alternative encoding wouldn't have the same deductive properties. Compare the pattern of inferences we have for the monadic type, in (56)–(59), with the corresponding pattern for the mooted simple type:[22]

$$A \multimap B \multimap C, A, B \vdash C \tag{61}$$
$$A \multimap B \multimap C, I \multimap A, B \vdash I \multimap C \tag{62}$$
$$A \multimap B \multimap C, A, I \multimap B \vdash I \multimap C \tag{63}$$
$$A \multimap B \multimap C, I \multimap A, I \multimap B \vdash I \multimap I \multimap C \tag{64}$$

[21]Glue Semantics is most closely associated with Lexical-Functional Grammar (Kaplan and Bresnan 1982, Dalrymple 2001, Bresnan et al. 2016), but Asudeh and Crouch (2002) have paired it with Head-Driven Phrase Structure Grammar (Pollard and Sag 1994) and Burke (2015) has recently paired it with the sorts of syntactic representations assumed in Tree-Adjoining Grammar (Joshi et al. 1975) and the Minimalist Program (Chomsky 1995).

[22]Rather than writing the types with $\rightarrow$, we write them with linear implication, $\multimap$, for better parity with the types above and to ensure that all other aspects of the logic are kept constant.

For the simple types (61)–(64), the final formula we derive depends in some non-trivial way on the entire collection of terms on the left-hand side of the sequent. In contrast, for the monadic types (56)–(59), the same result type can be derived for all configurations. What is important is that we can predict the final formula without having to consider the entire set of terms available. This shows that the compositionality of our monadic approach cannot be equivalently recapitulated in a simple type theory.

A.2 Fragment

$$\eta(\llbracket\text{not}\rrbracket\,(\llbracket\text{believe}\rrbracket\,(\llbracket\text{Hesperus}\rrbracket \star \lambda x.\, \llbracket\text{Phosphorus}\rrbracket \star \lambda y.\eta(\llbracket\text{is}\rrbracket\,(x)(y)))(\llbracket\text{Kim}\rrbracket))) \tag{65}$$

$$\llbracket\text{Hesperus}\rrbracket \star \lambda x.\eta(\llbracket\text{not}\rrbracket\,(\llbracket\text{believe}\rrbracket\,(\llbracket\text{Phosphorus}\rrbracket \star \lambda y.\eta(\llbracket\text{is}\rrbracket\,(x)(y)))(\llbracket\text{Kim}\rrbracket))) \tag{66}$$

$$\llbracket\text{Phosphorus}\rrbracket \star \lambda y.\eta(\llbracket\text{not}\rrbracket\,(\llbracket\text{believe}\rrbracket\,(\llbracket\text{Hesperus}\rrbracket \star \lambda x.\eta(\llbracket\text{is}\rrbracket\,(x)(y)))(\llbracket\text{Kim}\rrbracket))) \tag{67}$$

$$\llbracket\text{Hesperus}\rrbracket \star \lambda x.\, \llbracket\text{Phosphorus}\rrbracket \star \lambda y.\eta(\llbracket\text{not}\rrbracket\,(\llbracket\text{believe}\rrbracket\,(\eta(\llbracket\text{is}\rrbracket\,(x)(y)))(\llbracket\text{Kim}\rrbracket))) \tag{68}$$

Figure 2: Non-equivalent readings for *Kim doesn't believe Hesperus is Phosphorus*.

$$\begin{aligned}&\eta(\llbracket\text{but}\rrbracket\,(\llbracket\text{love}\rrbracket\,(\eta(\llbracket\text{Peter Parker}\rrbracket))(\llbracket\text{Mary Jane}\rrbracket))\\ &\qquad(\llbracket\text{not}\rrbracket\,(\llbracket\text{love}\rrbracket\,(\llbracket\text{Spider-Man}\rrbracket)(\llbracket\text{Mary Jane}\rrbracket))))\end{aligned} \tag{69}$$

$$\begin{aligned}&\llbracket\text{Spider-Man}\rrbracket \star \lambda x.\eta(\llbracket\text{but}\rrbracket\,(\llbracket\text{love}\rrbracket\,(\eta(\llbracket\text{Peter Parker}\rrbracket))(\llbracket\text{Mary Jane}\rrbracket))\\ &\qquad(\llbracket\text{not}\rrbracket\,(\llbracket\text{love}\rrbracket\,(\eta(x))(\llbracket\text{Mary Jane}\rrbracket))))\end{aligned} \tag{70}$$

Figure 3: Non-equivalent readings for *Mary Jane loves Peter Parker but she doesn't love Spider-Man*.

$$\eta(\llbracket\text{not}\rrbracket\,(\llbracket\text{believe}\rrbracket\,(\llbracket\text{Sandy}\rrbracket \star \lambda x.\, \llbracket\text{Sandy}\rrbracket \star \lambda y.\eta(\llbracket\text{is}\rrbracket\,(x)(y)))(\llbracket\text{Kim}\rrbracket))) \tag{71}$$

$$\llbracket\text{Sandy}\rrbracket \star \lambda x.\, \llbracket\text{Sandy}\rrbracket \star \lambda y.\eta(\llbracket\text{not}\rrbracket\,(\llbracket\text{believe}\rrbracket\,(\eta(\llbracket\text{is}\rrbracket\,(x)(y)))(\llbracket\text{Kim}\rrbracket))) \tag{72}$$

$$\llbracket\text{Sandy}\rrbracket \star \lambda x.\eta(\llbracket\text{not}\rrbracket\,(\llbracket\text{believe}\rrbracket\,(\llbracket\text{Sandy}\rrbracket \star \lambda y.\eta(\llbracket\text{is}\rrbracket\,(x)(y)))(\llbracket\text{Kim}\rrbracket))) \tag{73}$$

Figure 4: Non-equivalent readings for *Kim doesn't believe Sandy is Sandy*.

- The different contexts for the interpretation of referring expressions are completely determined by the order in which we evaluate monadic resources.
- This means that, just by looking at the linear order of the lambda term, we can check whether a referring expression is evaluated inside the scope of a perspective changing operator such as *believe*, or if it is interpreted using the standard interpretation.