On Mates' Puzzle*

Andrés Soria-Ruiz†

Abstract

I defend a metalinguistic account of Mates' puzzle: sentences where synonymous expressions cannot be substituted *salva veritate*. If Andrea thinks that attorneys are different from lawyers, and she thinks that Fiona is the former but not the latter, we may hesitate to substitute 'lawyer' for 'attorney' in 'Andrea believes that Fiona is an attorney', even though 'lawyer' and 'attorney' are synonymous. I argue that these sentences report *de re* beliefs about linguistic expressions, thereby blocking such substitutions, and I offer a compositional derivation of their meaning, relying on the concept generator technology from Percus and Sauerland 2003.

Keywords: Mates' puzzle, attitude reports, *de re*, metalinguistic, concept generators

Nobody doubts that any rational agent who believes that Fiona is an attorney believes that she is an attorney. But surely we can think of *some* rational agent who believes that Fiona is an attorney but does not believe that she is a lawyer:[1] suppose that Andrea doesn't know that attorneys are just lawyers (she thinks that 'attorney' means judge). When asked about her friend Fiona, who is a judge and not a lawyer, she may respond: 'Fiona is an attorney, not a lawyer'.[2] We may report Andrea's confusion by saying the following:

(1) a. Andrea believes that Fiona is an attorney.
b. Andrea believes that Fiona is not a lawyer.

Andrea is not irrational; she is simply wrong about what 'attorney' means. And yet, this seems to get us in trouble: do not 'lawyer' and 'attorney' mean the same thing? But then, should not the words 'lawyer' and 'attorney' be interchangeable in (1a) and (1b), thereby representing Andrea as believing a contradiction? This is *Mates' puzzle* (Mates 1952, p. 215).

Similarly, nobody doubts that any rational agent who believes that Ortcutt is a spy fails to believe that Ortcutt is not a spy. But surely we can think of *some* rational agent who believes of Ortcutt both that he is and that he is not a spy: Ralph sees a man by the docks and concludes that he is a spy. Ralph watches his neighbour Ortcutt at the beach one day and thinks that he is a pillar of the community—definitely *not* a spy. Unbeknownst to him, the man he saw by the docks was Ortcutt. We can say the following about Ralph:

(2) a. Ralph believes that Ortcutt is a spy.
b. Ralph believes that Ortcutt is not a spy.

*Forthcoming in *Mind & Language*.

†IFILNOVA, Lisbon. aruiz@fcsh.unl.pt.

[1] I assume that 'attorney' and 'lawyer' are synonymous.

[2] On notation: simple quotation marks (' ') enclose expressions of English and double quotation marks (" ") are for mixed quotation; *italics* are for non-English words and emphasis. When giving the semantic interpretation of natural language expressions (usually to the right of '$[\![...]\!] =$'), **boldface** represents denotations (model-theoretic objects like individuals and sets) and *italics* represent expressions.

It does not follow from (2a) and (2b) that Ralph is irrational; he simply failed to recognise the man he saw on each occasion as one and the same. And yet, that gets us in trouble: does not 'Ortcutt' denote the same individual in (2a) and (2b)? And if so, how can these sentences not represent Ralph as believing a contradiction? This is Quine's *double vision puzzle* (1956).

I argue that a standard solution to Quine's puzzle can be adapted to Mates'. The standard solution to Quine's puzzle is that the pair in (2) can report Ralph as standing in different relations of acquaintance to an individual without requiring him to recognise that individual as identical. The proposed solution to Mates' puzzle is that the pair in (1) can report Andrea as standing in different relations to the words 'attorney' and 'lawyer' without requiring her to recognise those words as synonymous. Put differently, (1a) and (1b) can report *de re* belief about particular words.

A compositional implementation of the proposed solution is offered, relying on the *concept generator* (CG) theory of Percus and Sauerland 2003 (see also Charlow and Sharvit 2014). The CG machinery was devised with the aim of providing specialised LFs for *de se/re* belief reports in possible worlds semantics and has been applied to double vision cases (Anand 2006, ch. 1; Rieppel 2017), as well as to objects of *de re* thought beyond individuals (Baron 2016). The basic idea behind CGs seems intuitively applicable to Mates' puzzle: CGs map the *res* in a belief report to a way in which the attitude holder thinks of that *res*. This way of thinking about an individual is an *individual concept* that picks out the *res* in the actual world but may pick out other individuals at the attitude holder's doxastic alternatives. Mates' puzzle is generated by sentences that report the beliefs of a subject who is confused about the meaning of a word. If we want a semantics for (1a) and (1b) that does not report Andrea as believing a contradiction, a theory that swaps the usual denotation of the word 'attorney' for Andrea's idiosyncratic concept of 'attorney' in (1a) would be ideal.

The paper is organised as follows. §1 presents the English (§1.1) and cross-linguistic (§1.2) data to be discussed. §2 argues that Mates' puzzle poses a challenge for the classical semantics for propositional attitudes. Subsequently, I look at ways of solving this challenge and find them wanting: in §2.1 I consider views that generalise *de re* belief about individuals to properties, qualities and kinds; in §2.2 I argue that Mates' puzzle is not quotation; in §2.3 I compare Mates' puzzle to Frege's, and I argue in favor of keeping the two phenomena separate. §3 presents and criticises Tancredi & Sharvit's 2019 account of Mates' puzzle. §4 introduces the standard account of *de re* belief reports about individuals relying on relations of acquaintance, as well as Percus & Sauerland's concept generator theory. §5 introduces the possibility of *de re* thought about linguistic expressions and offers informal truth conditions for Mates' sentences. §6 introduces the notion of a *metalinguistic concept generator* and sketches a compositional semantics for Mates' puzzle. §7 explores the semantics of 'know' (§7.1) and cross-linguistic data (§7.2). §8 concludes. A formal Appendix for this paper is accessible online.

1 Mates' puzzle

Mates' puzzle arises in situations in which an attitude holder is confused about the meaning of a word. In this section, I review the relevant reports in English (§1.1) and other languages (§1.2).

1.1 Same-language reports

Consider the case that we started with:

Context A: Andrea, a monolingual English speaker, is confused about the meaning of 'attorney': she thinks that it is a technical synonym of 'judge'. She knows that her friend Fiona is a judge and not a lawyer, so she says 'Fiona is an attorney, not a lawyer'.

(1) a. Andrea believes that Fiona is an attorney.
b. Andrea believes that Fiona is not a lawyer.

In (1) the words 'attorney' and 'lawyer' are not substitutable *salva veritate*. Call this the OPAQUE interpretation of (1).

Additionally, there is an interpretation of (1) according to which the words 'attorney' and 'lawyer' are substitutable *salva veritate*: in virtue of the fact that Andrea knows that Fiona is not a lawyer, it is natural to say that Andrea believes that Fiona is not an attorney. Under this interpretation, (1a) is false, while (1b) is true. This interpretation can be made salient as follows:

(3) Andrea believes that Fiona is not an attorney, although she wouldn't use those words.

Call this the TRANSPARENT interpretation of (1).

The sentences in (1) are not the only way that we can report Andrea's confusion. We can also say the following (Partee 1973, p. 414):

(4) Andrea believes that attorneys are not lawyers.

Finally, even though Andrea knows that Fiona is a judge, it is infelicitous (marked with #) to report Andrea's confusion using 'know' instead of 'believe' (Tancredi and Sharvit 2019, p. 3):

(5) # Andrea knows that Fiona is an attorney.

1.2 Cross-linguistic reports

One may wonder whether it is possible to report Andrea's confusion in a different language. 'Attorney' and 'lawyer' are both translatable as '*abogada*' in Spanish, and thus the following sentences translate (1a) and (1b):

(6) a. Andrea cree que Fiona es abogada.
Andrea believes that Fiona is an attorney.
b. Andrea cree que Fiona no es abogada.
Andrea believes that Fiona not is a lawyer.

However, only (6b) is true in context A. Intuitively, this is because Andrea knows that Fiona is not a lawyer. This suggests that the sentences in (6) only have a transparent interpretation.

One could think that this is due to the fact that 'attorney' and 'lawyer' are translated into the same Spanish word. To see that this does not play a role, consider a different case, this time involving a word with no synonyms in English ('mother-in-law') and a single translation in Spanish ('*suegra*'):

Context B: Mary, a monolingual English speaker, is confused about the meaning of 'mother-in-law': she thinks that it means step-mother. She knows that Ann is married to her father, and she says 'Ann is my mother-in-law'. Mary is single.

(7) a. Mary believes that Ann is her mother-in-law.

b. Mary cree que Ann es su suegra.
Mary believes that Ann is her mother-in-law.

While (7a) is true, (7b) is false for most of my informants,[3] myself included. Thus, (7b) hardly allows an opaque interpretation.

These observations suggest that it is difficult to report someone's linguistic confusion in a language other than the language *of* their confusion. But Tancredi & Sharvit (2019, p. 2) offer an example that can be reported in English and Japanese. Its Spanish translation is also true for a majority of my informants (19/29):

> Context C: John, a monolingual English speaker, is confused about the meaning of 'prime'. He takes it to be true of a number iff that number equals $x^3 - 1$ for some natural number x. He asserts '26 is prime', correctly calculating that $26 = 3^3 - 1$. He knows that 26 has exactly 4 factors.

(8) a. John believes that 26 is prime.

b. John-wa 26-ga sosuu de-aru to omotteiru.
John-TOP 26-NOM prime COP COMP believes

c. John cree que el 26 es primo.
John believes that ART 26 is prime.

Tancredi & Sharvit do not observe this, but if we manipulate the context so that John is not monolingual, but rather knows Spanish and knows the true meaning of 'primo', far fewer people find (8c) true (7/29):

> Context C*: John, a bilingual English/Spanish speaker, is confused about the meaning of the English word 'prime'. He takes it to be true of a number iff that number equals $x^3 - 1$ for some natural number x. He asserts '26 is prime', correctly calculating that $26 = 3^3 - 1$. However, John knows the true meaning of the Spanish word 'primo', and he knows that 26 has exactly 4 factors.

(8c) John cree que el 26 es primo.
John believes that ART 26 is prime.

2 A challenge for the semantics of propositional attitudes

Mates' puzzle poses a challenge for the classical possible worlds semantics for propositional attitudes.[4] This can be seen by noting three aspects of this analysis. First, atomic propositions result from semantically applying an individual (type e) as argument to a property (type $\langle e, st\rangle$), resulting in the set of worlds where the individual has the property.[5] Secondly, all expressions are interpreted as defined in the relevant model of interpretation. Thirdly, propositional attitude reports are cashed out in terms of relations between sets of possible worlds: for any individual x and proposition p, a sentence of the form 'x believes that p' is true iff every possible world compatible with the beliefs of x belongs to the set of p-worlds.

[3] I ran a survey with 29 Spanish speakers, in which I asked them to provide truth-value judgments for some of the sentences *cum* contexts described here. I offered participants a choice between saying that the relevant sentence was 'true', saying it was 'false' or saying 'I don't know'. Where relevant, I report results by giving the proportion of 'true' answers. In this case, the result was 9/29.

[4] See Hintikka 1962; Heim and Kratzer 1998, Ch. 12. For a different direction, see Kratzer 2006.

[5] Perhaps more commonly, properties are assigned type $\langle s, et\rangle$, but I follow the convention of conceiving of them as functions from individuals to propositions, which will make our compositional semantics more explicit down the line.

Since 'attorney' and 'lawyer' have the same meaning, this analysis predicts that the proposition *Fiona is an attorney* is identical to the proposition *Fiona is a lawyer*. Thus, the sole interpretation of (1) that this analysis can generate is a transparent interpretation.

To generate an opaque interpretation, we need to assign a denotation for 'attorney' in (1a) that matches Andrea's understanding of that word, namely, *being a judge*. But we need to do this in a principled way: we cannot simply assign the property of *being a judge* as the denotation of 'attorney', as this would massively over-generate incorrect interpretations. In this section, I review and criticise three principled approaches to this challenge.[6]

2.1 Generalised *de re* accounts

Consider again (1a) in its opaque interpretation. Intuitively, we want to interpret it as equivalent to 'Andrea believes that Fiona is a judge'. Thus, what we need is for the property of *being an attorney* to be substituted for the property of *being a judge*. In the literature on propositional attitudes, there are proposals designed to provide this result. Call these 'generalised *de re* accounts'.

These accounts extend available machinery for *de re* reports about individuals (such as (2)) to solve persistent problems in the semantics of propositional attitudes. The basic idea is that we can report *de re* attitudes about things other than individuals, e.g., numbers, properties, qualities and kinds (Cresswell and Stechow 1982; Fodor 1970; Schwager 2011; Sharvit 1998; Sudo 2014).

However, none of these approaches are useful *vis-à-vis* Mates' puzzle, for the following reason: the property (or quality or kind) denoted by 'attorney' is the same as that denoted by 'lawyer'. Hence, interpreting (1a) or (1b) as reporting *de re* belief about properties will not result in different truth conditions for each sentence. Nonetheless, it is worth quickly working through the most relevant generalised *de re* accounts to see why they cannot accommodate Mates' puzzle.

Cresswell & von Stechow (1982) tackle the problem of ascribing mathematical belief in possible worlds semantics. Since mathematical formulae are true in either *all* or *no* possible worlds, any ascription of a mathematical belief in possible worlds semantics assigns as its content either the set of all possible worlds or the empty set. This is counterintuitive, as one person may believe one false mathematical formula but not another. In possible worlds semantics, the same belief is ascribed in both cases. To avoid this, Cresswell & von Stechow propose that ascriptions of mathematical beliefs are sensitive to the properties that the beliefs are about. They propose that the reason why one can ascribe to someone the (false) belief that 8 is prime without thereby ascribing to them the (false) belief that 8 is a multiple of 3 is that in each case one is ascribing *de re* beliefs about different properties (*being prime* and *being a multiple of 3*).

Suppose we adopt this approach and claim that (1a) reports a *de re* belief about the property of *being an attorney*, while (1b) reports a *de re* belief about *being a lawyer*. Since these properties are identical, we would be ascribing the same belief in (1a) and (1b), and these sentences could not have different truth conditions.

Similar considerations apply to later incarnations of this proposal, designed to handle cases like the following (Schwager 2011, based on Fodor 1970; Sudo 2014):

> Context D: Malte and Adrian do not know each other. Adrian has seen a green Bench jacket in a catalogue and wants to buy one. Malte happens to own precisely such a green Bench jacket.

[6] It bears pointing out that none of the views to be discussed in this section are *actual* accounts of Mates' puzzle; they are accounts of closely related phenomena that could be extended to Mates' puzzle.

(9) Adrian wants to buy a jacket like Malte's.

(9) is true in context D, even though we would not expect Adrian to say 'I want a jacket like Malte's'. But since he would say 'I want this green Bench jacket' and *we know* that Malte has such a jacket, (9) is licensed. Simplifying greatly, Schwager proposes that we substitute the property that is explicitly mentioned in (9), *being a jacket like Malte's*, for the property that Adrian is thinking of, *being a green Bench jacket*.

Similarly, to generate an opaque interpretation of (1a), we need to substitute the property that is explicitly mentioned in (1a), *being an attorney*, for the property that Andrea is thinking of, *being a judge*. The problem, again, is that attorneys are lawyers, and thus whatever mechanism substitutes *being an attorney* for *being a judge* in (1a) will have the same effect on (1b).[7]

To summarise, generalised *de re* accounts face a simple problem: whatever mechanism generates the desired interpretation of (1a) will produce the same result for (1b). Schematically: for any mechanism m such that $m(\lambda x\lambda w.\textbf{attorney}(x)(w)) = \lambda x\lambda w.\textbf{judge}(x)(w)$: $\lambda x\lambda w.\textbf{attorney}(x)(w) = \lambda x\lambda w.\textbf{lawyer}(x)(w)$, therefore $m(\lambda x\lambda w.\textbf{lawyer}(x)(w)) = \lambda x\lambda w.\textbf{judge}(x)(w)$.

2.2 Quotational accounts

We want a mechanism that generates the property of *being a judge* when 'attorney', but not 'lawyer', is fed into it. This cannot be achieved by letting this mechanism be a function of the denotation of these words, since they are identical. But it could be a function of the *words* themselves. There are various ways of spelling out this type of metalinguistic approach, and one of them will form the basis of our account. First, however, let's look at the most salient way of giving a metalinguistic theory of Mates' puzzle, which invokes quotation.[8]

There are two ways of spelling out a quotational view, depending on whether one thinks that (1a) (under its opaque interpretation) involves quotation of the whole complement clause (*full-clause* quotation) or only part of it (*mixed* quotation).

Consider the full-clause analysis. The idea would be that when a speaker utters (1a), they are quoting Andrea as believing the sentence 'Fiona is an attorney'. Since we know that Andrea is confused about 'attorney' but not 'lawyer', this would explain why (1a) is logically independent of (1b).

However, it is hardly arguable that the complement clause of (1a) is quoted *in full*. To see this, we may recruit arguments against the hypothesis that indexical shifting amounts to quotation (Anand and Nevins 2004; Deal 2017, 2020): for example, we can *wh*-extract components from the complement clause in (1a). This should be impossible if the clause is quoted (* marks ungrammaticality):

(10) a. Who does Andrea believe is an attorney?
b. * Who does Andrea believe 'is an attorney'?

Here, the defender of the quotational view could reply that (1a) involves *mixed* quotation (Cappelen and Lepore 1997; Maier 2007, 2014a, 2020; Shan 2010). Mixed quotations are sentences in which we report someone's utterance by quoting only some of the words they used:

[7] The fact that this account predicts same truth conditions for (1a) and (1b) does not mean that it predicts a transparent interpretation. To the contrary, we would ascribe to Andrea the belief that Fiona is a judge in both cases.

[8] A metalinguistic account of Mates' puzzle was originally proposed by Church 1954 and Sellars 1955.

(11) a. Bush said that the enemy "misunderestimated" him.
b. Quine said that quotation "has a certain anomalous feature".

Following this model, it could be argued that (1a) should be understood as:

(12) Andrea believes that Fiona is an "attorney".

However, this reconstruction of (1a) is not immediately transparent as a case of mixed quotation. First, (1a) does not have overt quotation marks (which in mixed quotes delimit the quoted material). Secondly, mixed quotations may feature expressions from other languages, and even malapropisms ('misunderestimate'); Mates' puzzle disallows this. Thirdly, mixed quotations carry an inference that the quoted words were uttered: (11a) invites the inference that Bush uttered 'misunderestimated'. By contrast, (1a) does not trigger the inference that Andrea uttered 'attorney'. Finally, this view predicts that (5) (repeated here) should be felicitous in context A, contrary to what we have observed (Tancredi and Sharvit 2019, pp. 16–17). The reason for this is that if 'attorney' were (covertly) mixed-quoted in (1a), then it should also be (covertly) mixed-quoted in (5). But if we add overt quotation marks to (5), the resulting sentence is felicitous (and easily interpretable):

(5) # Andrea knows that Fiona is an attorney.

(13) Andrea knows that Fiona is an "attorney".

Notwithstanding the contrast with mixed quotation, we should not conclude that Mates' sentences are a *wholly* different phenomenon. In my analysis, (1a) will share some features with mixed quotation: (1a) will involve mentioning the word 'attorney', and although (1a) does not trigger the inference that Andrea uttered 'attorney', (1a) suggests that Andrea *would* use the word 'attorney' to talk about judges. Based on these facts, it could be argued that Mates' sentences showcase a special type of mixed quotation (perhaps a variety of *modalised* mixed quotation; cf. Wiślicki 2021).

2.3 Is Mates' puzzle just Frege's puzzle?

Finally, let us consider the possibility that Mates' puzzle is just (a version of) Frege's puzzle. Frege's puzzle involves failures of substitution of individual-referring expressions. A classic example: although Phosphorus is Hesperus, Plato believed that each was a different celestial body. Thus, 'Phosphorus' and 'Hesperus' are not substitutable *salva veritate* across (14):

(14) a. Plato believed that Phosphorus was visible in the morning.
b. Plato believed that Hesperus was not visible in the morning.

Mates' and Frege's puzzles share central properties: linguistically, they both involve failures of substitution of putatively synonymous words; psychologically, they both concern people with different ways of conceiving of one and the same entity (in the case of Frege's puzzle, an individual; in the case of Mates' puzzle, a property). This may suggest that both phenomena deserve a uniform treatment.[9]

However, there are important differences as well. First, Frege's puzzle does not seem to be sensitive to considerations about the language spoken by the attitude holder: we can report Plato's confusion regardless

[9] Additionally, Moffett (2002) argues that every instance of Mates' puzzle implies an instance of Frege's.

of what language(s) he spoke (although see Kripke 1979). By contrast, we saw in §1.2 that multiple instances of Mates' puzzle are sensitive to the language spoken by the attitude holder. Secondly, animals can be subject to Frege's, but not Mates', puzzle. Consider the following case:

> Context E: David's dog, Repu, is terrified of the vet, Néstor. However, Néstor also has a second job as a clown—called Ernesto—and regularly performs at birthday parties held by Repu's owners. Whenever Néstor shows up disguised as Ernesto, Repu is delighted.

(15) a. Repu is terrified of Néstor.

b. Repu is not terrified of Ernesto.

This suggests that the confusion involved in Mates' puzzle is distinctively metalinguistic (Kripke 1979, n.46, p. 459): without stating that Andrea is wrong about the meaning of 'attorney', we cannot generate the opaque reading of (1). Additionally, Mates' puzzle does not require a failure in the representations of the relevant properties: Andrea may know perfectly well what judges and lawyers are, and the puzzle can still arise.

By contrast, the confusion involved in Frege's puzzle concerns representations which may not be linguistic. Plato has two different representations of the same individual, Venus, which we report with two different names. Similarly, Repu has two different representations of the same individual, Néstor, which we associate with the names 'Néstor' and 'Ernesto'. But neither Plato nor Repu have wrong beliefs about words in their languages; it seems more appropriate to say that they have wrong beliefs about the world.[10]

For these reasons, I keep these phenomena apart and offer an account of Mates' puzzle that is not a descendant of, or clearly applicable to, Frege's. Regardless, there are solutions to Frege's puzzle that could potentially be extended to Mates' puzzle. This includes contextualist views, such as Crimmins 1992; Crimmins and Perry 1989 or more recently Goodman and Lederman 2019, as well as Richard 1990's *quasi*-quotational view. These views are Millian, as they assume that names are rigid designators that retain their normal reference wherever they occur, and they account for Frege's puzzle by co-opting additional semantic machinery in addition to Millianism.[11]

Contextualists about Frege's puzzle claim that some constituent of the sentences in (14) is context-dependent and can therefore take different values in different contexts of use.[12] They hold that each of the sentences in (14) is uttered in a different context, which is what generates coherent readings of otherwise incompatible sentences: (14a) is uttered in a context in which the representation associated with the name 'Phosphorus' (e.g., being the star visible in the morning) is made salient, whereas (14b) is uttered in a different context, where the salient representation is the one associated with the name 'Hesperus'.

Contextualists, however, face a challenge when it comes to sentences in which non-substitutable names occur within a single clause (Goodman and Lederman 2019; Lederman 2021):

(16) Plato did not believe that Hesperus is Phosphorus.

[10] Admittedly, one could say that Plato has a wrong belief about the word 'Hesperus', insofar as he thinks that it picks out a different individual from the word 'Phosphorus'. But to say that he is *confused about its meaning* (in the same way that we say that Andrea is confused about the meaning of the word 'attorney') seems odd. In addition, we can restate (14) in other languages without the words 'Hesperus' and 'Phosphorus', which suggests that Plato's confusion is not about these particular words. Moreover, to say that Repu is confused about the meaning of 'Néstor'/'Ernesto' is even odder.

[11] There are, of course, properly *Fregean* accounts of Frege's puzzle, which rely on Fregean senses or modes of presentation in order to account for failures of substitutivity (Yalcin 2015). I do not explore these views here.

[12] There are different theoretical choices depending on what element of the sentence is thought to carry context-dependence, such as *hidden-indexical* views (Schiffer 1992), Crimmins & Perry 1989's view that belief reports include *unarticulated constituents* and Dorr 2014's view that attitude verbs are context-sensitive.

The trouble with these sentences is that, in order to avoid attributing a contradiction to Plato, one would need to posit that a context shift occurs halfway through a clause, specifically between the names 'Hesperus' and 'Phosphorus'. This is hard to implement, at least on the standard Kaplanian view of contexts (but see Dorr 2014). Mates' puzzle also gives rise to single-clause substitution failures (4), and thus whatever challenge contextualists about Frege's puzzle face would be inherited by any extension of these views to Mates' puzzle.[13]

3 Tancredi & Sharvit

Tancredi & Sharvit (2019; henceforth T&S) offer an account of Mates' puzzle (which they call *de translato* attitude reports) that accommodates many of the data discussed in §1 while avoiding most of the pitfalls faced by the proposals reviewed in §2. T&S propose that expressions in the complement clause of a belief report can be interpreted relative to the believer's presumed I-language, rather than the speaker's. On T&S's view (2019, p. 21),

1. all interpretation takes a language and a perspective as parameters (in addition to worlds);
2. expressions can be indexed with a dedicated language transformation index T; and
3. for any expression σ, language L and perspective p, $[\![\sigma_T]\!]^{L,p} = [\![\sigma]\!]^{T_p(L),p}$, where $T_p(L)$ is a transformation that differs from L at most in assigning to some expressions a value consistent with p's presumed I-language.

In a matrix context, the perspective parameter p takes the value of the speaker's perspective, and indexed expressions are interpreted relative to the speaker's language. When embedded under a propositional attitude, however, the perspective parameter takes the value of the attitude holder's perspective, and indexed expressions are interpreted relative to their presumed I-language.

To account for Mates' puzzle, T&S subscript 'attorney' in (1a) with a dedicated language transformation index T. The language transformation index indicates that the word 'attorney' is to be interpreted not at the speaker's language L (that is, English) but at the minimal modification of English needed to make L consistent with Andrea's presumed I-language, $T_a(L)$. $T_a(L)$ is like English except for the fact that 'attorney' means $\lambda x \lambda w.\textbf{judge}(x)(w)$ instead of $\lambda x \lambda w.\textbf{attorney}(x)(w)$. In other words, the language transformation index swaps the actual meaning of 'attorney' for Andrea's interpretation of 'attorney'. Thus, where a stands for Andrea's perspective,

[13] Richard 1990 is a better candidate for a proposal about Frege's puzzle that could be extended to Mates', as it avoids the problem just described. In Richard's view, declarative sentences express RUSSELLIAN ANNOTATED MATRICES (RAMs). A RAM is a tuple formed by pairs consisting of each linguistic expression in a sentence together with its Russellian meaning (a name paired with an individual, a predicate paired with a property, etc.). Each of these pairs is called an ANNOTATION. For example, the annotation corresponding to the name 'Phosphorus' is the pair ⟨'Phosphorus', Venus⟩; the annotation corresponding to 'Hesperus' is ⟨'Hesperus', Venus⟩, etc. Richard proposes that we represent the beliefs of individuals with RAMs and calls the set of a subject's RAMs their REPRESENTATIONAL SYSTEM (RS). Very roughly, the truth of a belief report depends on whether the RAM determined by its complement clause matches a RAM in the believer's RS at least with respect to its Russellian component (additionally, context tells us the extent to which these RAMs must match with respect to their *quasi*-linguistic component as well). Given the *quasi*-linguistic component, this view can generate truth conditions for (1) that do not jointly attribute a contradiction to Andrea. Each report in (1) would ascribe a different RAM to Andrea in virtue of the fact that each contains a different word. And since the words 'attorney' and 'lawyer' contribute different annotations anywhere they occur, worries about single-clause failures of substitution vanish. However, this isn't yet an account of Mates' puzzle. Remember that in (1a) we need 'attorney' to swap its standard meaning for Andrea's idiosyncratic interpretation. As far as I can see, Richard's view does not offer a procedure for doing this. Regardless, I do not foreclose the possibility that an account of Mates' puzzle based on Richard 1990 can be developed, but I leave the matter for future consideration. Thanks to an anonymous reviewer for suggesting this.

- $[\![\text{attorney}_T]\!]^{L,a} = [\![\text{attorney}]\!]^{T_a(L),a} = \lambda x \lambda w.\textbf{judge}(x)(w)$; whereas
- $[\![\text{attorney}]\!]^{L,a} = \lambda x \lambda w.\textbf{attorney}(x)(w)$

T&S predict truth-conditions for (1) that do not jointly ascribe a contradiction to Andrea. This result is obtained by subscripting 'attorney' and 'lawyer' with language transformation indices, so that these words are interpreted relative to Andrea's presumed I-language. These are the truth-conditions for (1a), relative to a language-perspective-world index $\langle L, p, w\rangle$:[14]

(17) $[\![\text{Andrea believes that Fiona is an attorney}_T]\!]^{\langle L,p,w\rangle} = 1$ iff Andrea ascribes $[\![\text{attorney}_T]\!]^{L,a}$ to $[\![\text{Fiona}]\!]^{L,a}$ in w

That is, (1a) is true just in case Andrea ascribes the property of *being a judge* to Fiona.

By allowing synonymous words to denote different properties, T&S's view avoids the main trouble faced by generalised *de re* accounts. Moreover, since their proposal does not rely on quotation, it steers away from the problems discussed in §2.2 as well.

However, T&S's view faces three main challenges. First, by populating the complement clauses of belief reports with language transformation indices, T&S predict the opaque but not the transparent interpretation of (1). T&S could claim that there exists an LF for (1) where 'attorney'/'lawyer' are not subscripted with language transformation indices, allowing for a transparent reading. But this begs a pressing question about the distribution of such indices, which T&S do not tackle. We return to this issue below.

Secondly, T&S propose (tentatively, see 2019, p.21 n.19) that the perspective parameter, which determines the language of interpretation of subscripted expressions, is the same parameter that guides the interpretation of *predicates of personal taste* (PPTs, e.g., 'fascinating'). But this implies predictions which are not borne out.

PPTs are expressions whose semantics depend on subjective perspectives or opinions, and this is traditionally accounted for by proposing that their semantic interpretation depends on a judge parameter (see e.g. Bylinina 2017; Coppock 2018; Lasersohn 2005, *a.m.o.*). T&S propose that this judge parameter can be used, in addition, to determine the *language* of interpretation of expressions subscripted with language transformation indices. This predicts that subscripted predicates should behave like PPTs in environments that determine the same judge parameter.[15] Specifically, at any index of evaluation i, the judge of i determines both the perspective relative to which a PPT should be interpreted and the language relative to which subscripted predicates are interpreted. This is initially attractive, as it can account for the contrast between 'believe' and 'know' observed in (5). Note that, when embedded under 'believe', 'fascinating' is interpreted as fascinating *from the believer's perspective*, while when embedded under 'know', it is interpreted as fascinating *from the speaker's perspective*:

Context F: Andrea finds Fiona fascinating; the speaker does not.

(18) a. Andrea believes that Fiona is fascinating. (fascinating from Andrea's perspective)
b. # Andrea knows that Fiona is fascinating. (fascinating from the speaker's perspective)

[14] T&S frame their proposal in Cresswell and Stechow 1982's structured proposition view, according to which belief is analysed in terms of a relation of *ascription* towards the constituents of a Russellian structured proposition, all of which are analyzed *de re*. I set aside this feature of their analysis.

[15] T&S draw a close connection between two *a priori* very different semantic phenomena, PPTs and Mates' puzzle: PPTs belong to a broader class of lexically subjective expressions; by contrast, the type of shift in interpretation illustrated by Mates' puzzle is not restricted to a specific class of expressions. However, I do not mean to suggest that T&S treat subscripted predicates as PPTs.

The fact that the speaker does not find Fiona fascinating explains why (18b) is infelicitous. As we saw, a parallel contrast arises in Context A: when embedded under 'believe', 'attorney' can be interpreted *according to the believer's language*; when embedded under 'know', it is interpreted *according to the speaker's language*.

Context A

(1a) Andrea believes that Fiona is an attorney. ('attorney' according to Andrea's language)

(5) # Andrea knows that Fiona is an attorney. ('attorney' according to the speaker's language)

The explanation for this contrast could be parallel as well: under 'know', 'attorney' is obligatorily interpreted as the speaker understands it, and since in context A the speaker knows that Fiona is not an attorney, (5) is infelicitous.

However, two points of contrast between 'attorney' (in contexts like A) and PPTs arise here. First, while the most natural reading of (18b) is one according to which it (i) presupposes that Fiona is fascinating *to the speaker* but (ii) asserts that Andrea believes that Fiona is fascinating *to Andrea*, it is difficult to access a parallel reading of (5) according to which it (i) presupposes that Fiona is an "attorney" *according to the speaker* (i.e., an attorney) but (ii) asserts that Andrea believes that Fiona is an "attorney" *according to Andrea* (i.e., a judge).[16]

Secondly, while it is unproblematic to understand unembedded PPTs as indexed to the speaker's perspective, it is impossible to understand unembedded 'attorney' as indexed to the speaker's I-language (e.g. as *judge*).

(19) a. Andrea: 'Fiona is fascinating!'

b. Andrea: 'Fiona is an attorney!'

Again, T&S could reply that, when unembedded, 'attorney' is not subscripted with a language transformation index. This brings us back to the question of the distribution of these indices. There are 3 options:

1. Language transformation indices may subscript expressions below and above attitude verbs.
2. Language transformation indices obligatorily subscript all expressions below attitude verbs and no expressions above them.
3. Language transformation indices may subscript expressions below, but never above, attitude verbs.

The first, maximally liberal option predicts unattested interpretations of (19b) where 'attorney' means judge. The second option avoids this but does not predict transparent interpretations of (1), as we have seen. The third option solves both problems: unembedded expressions cannot shift their interpretation; embedded expressions may or may not. Thus, T&S's view should be complemented with a claim about the distribution of indices along option 3. Until then, something remains unexplained: why is it that language transformation indices are only found below attitude verbs—after all, in T&S's view such indices interact with attitude verbs but are semantically independent of them.

The third difficulty with T&S's proposal is that, based on their observation about successful cross-linguistic *de translato* reports (repeated here), they predict the wholesale translatability of such reports. As we saw in §1.2 however, this is incorrect.

Context C

[16] T&S's proposal for 'know' (2019, p.31 and ff) predicts just this reading of (5). T&S discuss this problem in fn.28, p.33. We return to this issue in §7.1.

(8) a. John believes that 26 is prime.

b. John-wa 26-ga sosuu de-aru to omotteiru.
John-TOP 26-NOM prime COP COMP believes

c. John cree que el 26 es primo.
John believes that ART 26 is prime.

Consider (8a) first. To generate an opaque interpretation of (8a), 'prime' has to be subscripted with a language transformation index. Where j is John's perspective,

- $[\![\text{prime}_T]\!]^{L,j} = [\![\text{prime}]\!]^{T_j(L),j} = \lambda n.\exists x(n = x^3 - 1)$

To produce a *de translato* interpretation of (8b), '*sosuu*' should denote $\lambda n.\exists x(n = x^3 - 1)$ as well. The step from '*sosuu*' to $\lambda n.\exists x(n = x^3 - 1)$ is presumably identical to the step from 'prime' to $\lambda n.\exists x(n = x^3 - 1)$, except for the fact that the language of interpretation L is Japanese. Substituting 'prime' for '*sosuu*', T&S's proposal is to take the word '*sosuu*' and evaluate it, not at the language of evaluation (= Japanese) but at the *minimal modification of Japanese needed to make Japanese consistent with John's presumed I-language*:

- $[\![\textit{sosuu}_T]\!]^{L,j} = [\![\textit{sosuu}]\!]^{T_j(L),j} = \lambda n.\exists x(n = x^3 - 1)$

The problem is that the minimal modification of Japanese needed to make Japanese consistent with John's presumed I-language will not necessarily guarantee that '*sosuu*' means $\lambda n.\exists x(n = x^3 - 1)$, because by hypothesis '*sosuu*' does not match anything in John's presumed I-language (by contrast, 'prime' does). Moreover, given that John is a monolingual English speaker, the minimal modification of Japanese needed to make Japanese consistent with John's presumed I-language would be to turn it into English (perhaps including John's idiosyncratic understanding of 'prime'). But in that modification, there would still be no entry for '*sosuu*'.

To generate the right cross-linguistic results, T&S propose the following principle:[17]

Definition 1 (*Idiolect Preservation across Translation* (IPT)) *Impose on a proper translation σ' of σ not only that the language L' interpreting σ' give the same value that L gives to σ but also that for all transformations T, $T(L')(\sigma') = T(L)(\sigma)$.*

IPT guarantees that, if T_j(English)('prime') $= \lambda n.\exists x(n = x^3 - 1)$, then T_j(Japanese)('*sosuu*'), T_j(Spanish)('*primo*'),... $= \lambda n.\exists x(n = x^3 - 1)$. Thus, IPT would correctly predict cross-linguistic reports like those in (8).

IPT over-generates however, as it also predicts that (6) and (7b) should have opaque interpretations. Additionally, IPT predicts that manipulating context C so that John knows Spanish and knows the right meaning of 'prime' (Context C*) should have no bearing on (8c), contrary to what we observed in §1.2.

4 *De re* belief about individuals and concept generators

In the following sections, I propose a view that overcomes the aforementioned difficulties while preserving the right predictions of T&S's view. My view takes inspiration from accounts of *de re* attitude reports. Following Kaplan 1968, *de re* reports introduce tacit existential quantification over modes of presentation of the relevant *res*. Lewis (1979, 1983) proposes that we think of such modes of presentation in terms of the self-ascription of properties; *de re* ascriptions in particular involve the self-ascription of the property of

[17] T&S formulate this principle in passing, on p.20. The label is mine.

inhabiting a world where a certain relation of acquaintance holds uniquely between the believer and the *res*. Here is a standard formulation:[18]

(20) $[\![a \text{ believes}^{\text{DE RE}} \text{ that } b \text{ is } P]\!]^w = 1$ iff
$\exists R \forall x((R(a,x,w) \leftrightarrow x = b)\ \&\ \forall w' \in DOX_{a,w} : \exists y \forall z((R(a,z,w') \leftrightarrow z = y)\ \&\ P(y)(w')))$

To see this proposal in action, recall Ralph's confusion:

(2) a. Ralph believes that Ortcutt is a spy.
b. Ralph believes that Ortcutt is not a spy.

By giving truth-conditions to (2) along (20), we block the conclusion that Ralph believes a contradiction:

(21) $[\![(2a)]\!]^w = 1$ iff $\exists R \forall x((R(\text{Ralph},x,w) \leftrightarrow x = \text{Ortcutt})\ \&\ \forall w' \in DOX_{\text{Ralph},w} :$
$\exists y \forall z((R(\text{Ralph},z,w') \leftrightarrow z = y)\ \&\ \textbf{spy}(y)(w')))$

(22) $[\![(2b)]\!]^w = 1$ iff $\exists R \forall x((R(\text{Ralph},x,w) \leftrightarrow x = \text{Ortcutt})\ \&\ \forall w' \in DOX_{\text{Ralph},w} :$
$\exists y \forall z((R(\text{Ralph},z,w') \leftrightarrow z = y)\ \&\ \sim \textbf{spy}(y)(w')))$

In Ralph's predicament, there exists a relation of acquaintance between Ralph and Ortcutt—being seen by the docks—and, since Ralph thinks that the man he saw is a spy, that same relation also holds between him and a spy across Ralph's doxastic alternatives, making (2a) true. There also exists another relation of acquaintance between Ralph and Ortcutt—being seen at the beach—but, since Ralph thinks that the man he saw at the beach is no spy, that same relation holds between Ralph and a non-spy across his doxastic alternatives, making (2b) true as well. In Ralph's doxastic alternatives, the spy and the non-spy are different people, which is why Ralph does not believe a contradiction.

Percus & Sauerland (henceforth: P&S; 2003, p.10 ff.) provide a perspicuous way of generating truth conditions for *de re* reports compositionally, with a function that they dub a CONCEPT GENERATOR.[19] A concept generator is a function G from individuals to individual concepts that, for an attitude holder x at a world w, maps the individuals that x is acquainted with at w onto individual concepts (descriptions) of those individuals at w:

Definition 2 (*Concept generator*) *G is a concept generator for x at w iff*

(a) *G is an object of type $\langle e, se\rangle$, and*
(b) *$Dom(G) = \{y : x$ is acquainted with y at $w\}$*

An *acquaintance-based* concept generator $\mathcal{G}$ for x at w is a specific type of concept generator that preserves the acquaintance relations that, at w, hold between the individuals in its domain and x:

Definition 3 (*Acquaintance-based CG*) *$\mathcal{G}$ is an acquaintance-based concept generator for x at w (written $ACG_{x,w}$) iff*

[18] Strictly speaking, the individual referred to as 'a' within the proposition denoted by the complement clause of (20) may not be a but rather the individual that a self-identifies with. A standard way of capturing this is to state these truth-conditions in terms of centered worlds rather than possible worlds (Lewis 1979).

[19] P&S propose this function while arguing for the need to distinguish between *de re* and *de se* readings of attitude reports. Anand 2006; Rieppel 2017 subsequently applied this machinery to double vision cases. See also Charlow and Sharvit 2014; Lederman 2021 for further development and discussion of P&S's view.

(a) *$\mathcal{G}$ is a concept generator for x at w, and*

(b) *the individual concepts that $\mathcal{G}$ yields are acquaintance-based in the sense that, for all z in $Dom(\mathcal{G})$,*

- *there is some acquaintance relation R that x bears uniquely to z in w, and*
- *$\forall w' \in DOX_{x,w}$: x bears R uniquely to $\mathcal{G}(z)(w')$ in w'*

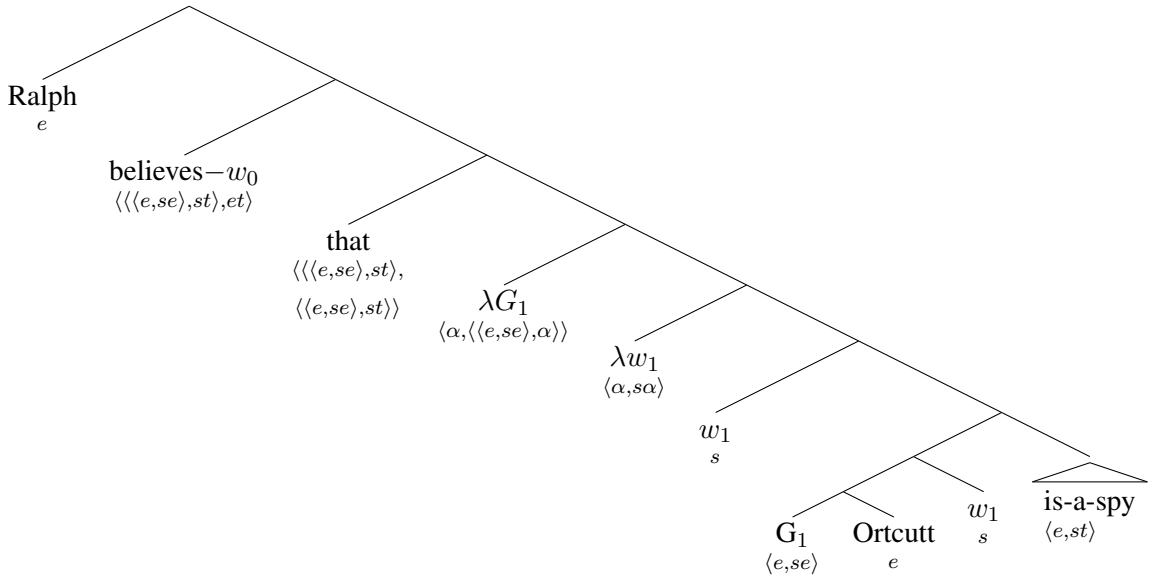
Ralph is acquainted with Ortcutt *via* relation R (where R is *being seen by the docks*) at w. An acquaintance-based concept generator for Ralph at w is a function $\mathcal{G}$ that, taking Ortcutt as argument, returns an individual concept of Ortcutt for Ralph—let us abbreviate it as m_{Ortcutt}—such that, at Ralph's doxastic alternatives w', m_{Ortcutt} picks out the unique individual that Ralph saw by the docks in w' (which may or may not be Ortcutt).

Concept generators allow for the following, simpler formulation of the target truth conditions for *de re* belief reports:

(23) $[\![a \text{ believes that } b \text{ is } P]\!]^w = 1$ iff $\exists G(G \text{ is an } ACG_{a,w} \;\&\; \forall w' \in DOX_{a,w} : P(G(b)(w'))(w'))$

To provide a full account of *de re* belief reports, these truth conditions have to be derived compositionally from the meaning of each constituent in the sentence. P&S enrich the LFs of the complement clauses of belief reports with concept generator *variables* and *abstractors* over these variables, in the same way that possible world variables and abstractors are introduced in order for those complement clauses to have an intensional denotation.[20] The complement clauses of *de re* reports contain, in addition to world-variables and abstractors over those variables, concept generator variables and abstractors over such variables. CG-variables are placed in LF as sisters of the *rei* that we need individual concepts of. To a *de re* report such as (2a), we assign the following LF (to facilitate following the computation along the tree, each node has its semantic type underneath it. For any semantic types α, β and any variable υ_β, $\lambda\upsilon$ has type $\langle\alpha, \beta\alpha\rangle$):

(24) Ralph [believes-w_0 [that [λG$_1$ [λw_1[[[[G$_1$ Ortcutt]w_1] is a spy]w_1]]]]]



[20] This is one possibility; an alternative is to complement the rule of Functional Application with a rule of *Intensional* Functional Application which is sensitive to the intensions of the expressions to which it applies (see Heim and Kratzer 1998, ch. 12).

According to this LF, *believe* takes as its internal argument a function from concept generators to propositions:[21]

(25) $[\![\text{believe}]\!]^w = \lambda\omega_{\langle\langle e,se\rangle,st\rangle}\lambda x.\exists G(G \text{ is an } ACG_{x,w} \,\&\, \forall w' \in DOX_{x,w} : \omega(G)(w'))$

In this account, the *de re* character of a belief report is provided by the presence of a CG-variable as sister to the appropriate *res* in the LF of its complement clause. *De dicto* reports are assigned a structure that contains no CG-variable in the complement clause. In the absence of any such variable, the existential quantifier introduced by 'believe' does not bind anything in the complement clause.

In the following sections, I take inspiration from this account of *de re* belief about individuals in order to derive the opaque and transparent interpretation of (1). The basic idea is that, just as one can have *de re* thoughts about individuals, one can have *de re* thoughts about *expressions of a language*.

5 *De re* belief about words

Recall Andrea's confusion. She doesn't know that 'attorney' is just another word for lawyer. She thinks (mistakenly) that 'attorneys' are judges, and she thinks (rightly) that her friend Fiona is a judge and not a lawyer. It seems reasonable to say that what makes (1a) true is partially the fact that Andrea has a certain belief about a particular word, 'attorney'. What also makes (1b) true is the fact that she does not have the same belief about 'lawyer'. The belief, in both cases, concerns whether the relevant word applies to Fiona. For Andrea, this is true of 'attorney' but not of 'lawyer'.

How can Andrea come to have different beliefs towards synonymous words? One possible starting point is to assume that (part of) the object of belief is the words themselves and not, or not just, their meaning.[22] Since 'attorney' and 'lawyer' are different *words*, speakers can stand in different relations to them. The relation that is of interest to us is the relation of *identifying a property as the denotation of a word*. This is the relation in which one stands to a word in virtue of having beliefs about its meaning. We can formulate this relation as follows:

Definition 4 (*Word Identification*) *For any word 'F', individual x, property P and world w, x identifies P as the denotation of 'F' at w iff 'F' as evaluated by x at w picks out property P.*

Following this thread, we can adapt the Kaplan-Lewis story about individual *de re* belief in order to obtain the following (opaque) truth-conditions for (1a) and (1b):

(26) $[\![(1a)]\!]^w = 1$ iff there exists a property P that Andrea identifies as the denotation of 'attorney' at w and for every world w' in Andrea's doxastic alternatives at w, Fiona has property P at w'.

(27) $[\![(1b)]\!]^w = 1$ iff there exists a property P that Andrea identifies as the denotation of 'lawyer' at w and for every world w' in Andrea's doxastic alternatives at w, Fiona lacks property P at w'.

[21] See Appendix for a full derivation of (2a).

[22] This is not the only way that Andrea could come to have different attitudes towards synonymous words. Andrea might come to have such different beliefs by intersecting (or not) the relevant extension (of attorneys and lawyers), or by providing counterparts of the relevant individuals. But presumably, both options presuppose the possibility of interpreting each word differently, something which could be achieved by a procedure of interpretation-shifting along the lines of Tancredi and Sharvit 2019. Another idea might be to implement a metalinguistic mechanism similar to that required to interpret *metalinguistic comparatives* (Giannakidou and Yoon 2011; Morzycki 2011). I thank two anonymous reviewers for pressing me in this regard.

In context A, the truth of (1a) is evidenced by the fact that there exists a property, *being a judge*, that Andrea (i) identifies as 'attorney' in w and (ii) holds of Fiona across Andrea's doxastic alternatives. And the truth of (1b) is evidenced by the fact that there exists a property, *being a lawyer*, that Andrea (i) identifies as 'lawyer' in w and (ii) does not hold of Fiona across Andrea's doxastic alternatives.

In the next section, I adapt the concept generator machinery to generate these truth-conditions compositionally. But before that, it bears pointing out an important contrast between *de re* thought about individuals and *de re* thought about words. When one speaks of the *de re/dicto* distinction (about individuals), it is standard to characterise the former as admitting the substitution of co-referential expressions and the latter as blocking them. When we interpret (2a) *de re*, we can substitute 'Ortcutt' for 'the man Ralph saw at the beach' *salva veritate*. By contrast, if we interpret this sentence *de dicto*, substitution is blocked.

When we consider *de re* belief about expressions, however, things work the other way around: in the opaque interpretation of (1a), where (as I am claiming) Andrea has a *de re* thought about the word 'attorney', 'attorney' is not substitutable for a co-intensional expression, e.g., 'lawyer'. By contrast, in its transparent interpretation, (1a) allows for the substitution of 'attorney' for 'lawyer'. So, in the view I am putting forward, a *de re* interpretation blocks the relevant substitutions rather than allowing for them. This is surprising, but the rationale behind it is simple: *de re* reports require that the *res* be preserved. In the case of *de re* reports about individuals, we can substitute co-referential expressions because co-referential expressions preserve the *res*, that is, the individual. In the case of *de re* reports about linguistic expressions, however, substituting the expression changes the *res*, therefore blocking substitution.

6 Metalinguistic concept generators

In this section, I introduce the formal apparatus required to generate the above-discussed truth conditions compositionally. My proposal is importantly similar to that of P&S: sentences like (1a) are ambiguous between their opaque and their transparent interpretation, and this is cashed out by assigning different LFs to these sentences, some of which contain concept-generator variables in their complement clauses and some of which do not. The type of concept generator variable that we need should be a function that shifts the interpretation of a word like 'attorney' so that it is interpreted in the way that Andrea understands it.

6.1 A function from properties or from expressions?

In principle, a concept generator variable should be a function that takes 'attorney'—in (1a)—and returns the property of *being a judge*. But we saw in §2.1 that this function is not so easy to characterise: if one defines a function m so that $m(\lambda x \lambda w.\textbf{attorney}(x)(w)) = \lambda x \lambda w.\textbf{judge}(x)(w)$, then, since $\lambda x \lambda w.\textbf{attorney}(x)(w) = \lambda x \lambda w.\textbf{lawyer}(x)(w)$, $m(\lambda x \lambda w.\textbf{lawyer}(x)(w)) = \lambda x \lambda w.\textbf{judge}(x)(w)$.[23]

Alternatively, we can think of the concept generator variable as behaving formally like a quotation device, that is, as a function from expressions, not properties. If we define a function m such that $m(\textit{attorney}) = \lambda x \lambda w.\textbf{judge}(x)(w)$, then, since 'attorney' and 'lawyer' are different *expressions*, we no longer predict that $m(\textit{lawyer}) = \lambda x \lambda w.\textbf{judge}(x)(w)$. But we saw the pitfalls of this type of account in §2.2.

However, quotation is not the only way to characterise the kind of metalinguistic mechanism we need. In particular, we can retain the idea that a function can be sensitive to the expressions it takes as arguments

[23] Roughly this type of application of the concept generator machinery can be found in Baron 2016, designed to account for the cases that Sudo 2014 and Schwager 2011 worry about.

without *quoting* anyone's words.

I propose that we think of this function as a variable version of a mixed quotation operator (Maier 2014a). A mixed quotation operator is a function from expressions to properties. Whereas this operator has a constant meaning, our concept generator function will be a variable whose exact value will be provided by the belief operator, similarly to what we saw in the case of standard *de re* reports in P&S 2003. When bound by 'believe', this function will map any expression in its argument to whatever the believer thinks the expression means.

6.2 Maier 2014a's mixed quotation operator

This will become clearer if we briefly introduce Maier's mixed quotation operator. Recall the examples from §2.2:

(11) a. Bush said that the enemy "misunderestimated" him.
b. Quine said that quotation "has a certain anomalous feature".

In (11b), the expression ' "has a certain anomalous feature" ' is both mentioned and used. It is mentioned because the meaning of the sentence relies on the fact that Quine used those very words, and it is used because Quine did attribute a property to quotation, namely the property of having a certain anomalous feature.

To account for this, Maier proposes that mixed quotation defers the interpretation of an expression to someone other than the speaker. Specifically, ' "has a certain anomalous feature" ' in (11b) will mean *whatever Quine meant by 'has a certain anomalous feature'*.

Maier achieves this by defining a mixed quotation operator " " with particular properties at the syntax-semantics interface. To understand how this operator works, it is best to say something about Maier's characterisation of a language. In Maier's view, a language is composed of a set L_{phon} of phonologically well-formed expressions; a set L_{syn} of grammatically well-formed expressions (a subset of $L_{\text{phon}} \times$ CAT, where CAT is the set of grammatical categories $N, NP, NP/S, N/N, ...$[24]); and L_{sem}, formed by the usual model-theoretic objects, but crucially including L_{syn} (members of which are assigned type ε). Including L_{syn} in L_{sem} affords the possibility of incorporating expressions into semantic interpretation. This is crucial to account for quotation.

We are ready to characterise Maier's mixed quotation operator: (i) Phonologically, it is generated by enclosing any string within double quotation marks " ". (ii) Syntactically, it is captured by the following L_{syn} composition rule: if $\langle \sigma, C \rangle \in L_{\text{syn}}$, then $\langle \text{“}\sigma\text{”}, C \rangle \in L_{\text{syn}}$ (where C is an arbitrary syntactic category). That is, this operator takes well-formed syntactic expressions as arguments (but also malapropisms such as 'misunderestimate', so long as their syntactic category is clear). And (iii) semantically, "σ" stands for whatever some salient speaker expressed when she used the form σ. More formally, where E is a three-way relation of *expression*,

(28) $[\![\text{“}\sigma\text{”}]\!] = \iota X[E(x, \ulcorner\sigma\urcorner, X)]$ = the X that the source x expressed with her use of the phonological string σ

Applied to (11b), and assuming that the relevant source is Quine, this analysis predicts the following:

[24] Maier uses a simple categorial grammar with basic categories N, NP & S, combinable with a left (\) and right slash (/). See Maier 2014a, p.19 ff. as well as Shan 2010.

(29) $[\![$"has a certain anomalous feature"$]\!] = \iota X[E(\text{Quine}, \ulcorner \text{has a certain anomalous feature} \urcorner, X)] =$ the X that Quine expressed with his use of the phonological string 'has a certain anomalous feature'.

Crucially, the argument of the double quotes is not an interpreted expression but a well-formed expression (from L_{syn}). In other words, the double quotes *see* an expression, not its meaning, and return a (possibly shifted) denotation.

I propose that we think of concept generator variables as a formally similar mechanism but with two important differences: first, concept generators have a variable, rather than a constant, meaning. Second, instead of letting concept generators range over L_{syn}, as Maier 2014a proposes for mixed quotation, concept generators will range over a specialised basic domain D_u of *well-formed linguistic entities*, or simply *expressions*, of type u. The members of D_u are pairs of a phonological representation and a syntactic category, but they carry no assigned semantic interpretation.[25]

Thus, let us assume that there exists a variable function $\mathcal{H}$ with the following properties:

1. It is phonologically silent.
2. Syntactically, it is captured by the following syntactic rule: if $\langle \sigma, NP\backslash S\rangle \in L_{\text{syn}}$, then $\langle \mathcal{H}(\sigma), NP\backslash S\rangle \in L_{\text{syn}}$ (where $NP\backslash S$ is the category of simple predicates).
3. Semantically, $\mathcal{H}$ is a function from entities of type u ($\in D_u$) to properties (type $\langle e, \langle s, t\rangle\rangle$).

'Believe' will bind this variable, similarly to what occurs in the CG-machinery of P&S. Let's see how.

6.3 'Believe' introduces metalinguistic concept generators

As we saw in §5, concept generators are functions from individuals to individual concepts. In our account, we require concept generators to be functions from well-formed linguistic expressions (type u) to properties (type $\langle e, st\rangle$). So, let us define a METALINGUISTIC CONCEPT GENERATOR (MCG for short) as follows:

Definition 5 (*Metalinguistic CG*) *$\mathcal{H}$ is a metalinguistic concept generator for x at w (written $MCG_{x,w}$) iff*

(a) *$\mathcal{H}$ is an object of type $\langle u, \langle e, st\rangle\rangle$,*

(b) $\textsc{Dom}(\mathcal{H}) = \{\sigma \in D_u : \sigma$ *is identified by x at w as denoting some property $P\}$ and*

(c) *for any property P and any expression σ in* $\textsc{Dom}(\mathcal{H})$, *$\mathcal{H}(\sigma) = P$ iff x identifies P as the denotation of σ at w.*

In P&S's original proposal, the way to obtain the desired *de re* interpretations was to assume that there were concept generator variables in the appropriate places in logical form. In order to introduce MCGs, we need to do the same. To this effect, we take the following steps: first, the complement clause of an attitude verb can have an LF where there are MCG-variables next to predicates (as well as the corresponding abstractor at the top of the clause), and furthermore, the predicate is *mentioned* rather than used. So what figures in the interpretation of these sentences is not the usual denotation of the predicate but the predicate itself, as sister to the appropriate MCG-variable.

[25] See Potts 2007, p. 410, for the characterisation of D_u and Maier 2014b for further discussion. Letting metalinguistic concept generators range over D_u ensures that they take uninterpreted linguistic entities as their arguments, so that one and the same expression can receive different interpretations, as desired. I thank a reviewer for drawing my attention to this.

Secondly, once MCG-variables and abstractors are introduced, the complement clause of an attitude verb no longer denotes a standard proposition, $\langle s,t\rangle$. In our account, complement clauses are functions from metalinguistic concept generators (type $\langle u,\langle e,st\rangle\rangle$) to propositions. When this type of denotation appears unembedded, it cannot simply be evaluated for truth relative to a possible world. To remedy this, we will assume that every index of evaluation contributes a possible world together with a (constant) DISQUOTATIONAL FUNCTION Δ, defined as follows:[26]

Definition 6 (*Disquotational function*) *Δ is a disquotational function at w (written Δ_w) iff*

(a) *Δ is an object of type $\langle u,\langle e,st\rangle\rangle$,*

(b) $\textsc{Dom}(\Delta) = D_u$ *and*

(c) *for any property P and any expression σ in $\textsc{Dom}(\Delta)$, $\Delta(\sigma) = P$ iff the linguistic community identifies P as the denotation of σ at w.*

When applied to any clause, Δ maps expressions that carry an MCG-variable to their interpretation according to the linguistic community and is otherwise idle.[27] Let's see how this works for 'Fiona is an attorney' with an LF containing an MCG-variable and one without it:[28]

(30) $\lambda\mathcal{H}_1$ [λw_1 [w_1 [Fiona [$\mathcal{H}_1$ is-an-attorney]]]] $= \lambda\mathcal{H}\lambda w.\mathcal{H}$(*attorney*)(Fiona)$(w)$
$[\![(30)]\!]^{\Delta,w} = 1$ iff Δ(*attorney*)(Fiona)(w) = **attorney**(Fiona)(w)

(31) $\lambda\mathcal{H}_1$ [λw_1 [w_1 [Fiona is-an-attorney]]] $= \lambda\mathcal{H}\lambda w.$**attorney**(Fiona)$(w)$
$[\![(31)]\!]^{\Delta,w} = 1$ iff **attorney**(Fiona)(w)

Thirdly, we modify the lexical entry of 'believe' so that it quantifies over metalinguistic concept generators:[29]

(32) $[\![\text{believe}]\!]^{\Delta,w} = \lambda\omega_{\langle\langle u,\langle e,st\rangle\rangle,st\rangle}\lambda x.\exists\mathcal{H}(\mathcal{H}$ is an $MCG_{x,w}$ & $\forall w' \in DOX_{x,w} : \omega(\mathcal{H})(w'))$

The LF for (1a) (in its opaque interpretation), which results from combining (32) and (30), would be as shown in (33a), generating the truth conditions in (33b) (full derivation in the Appendix):

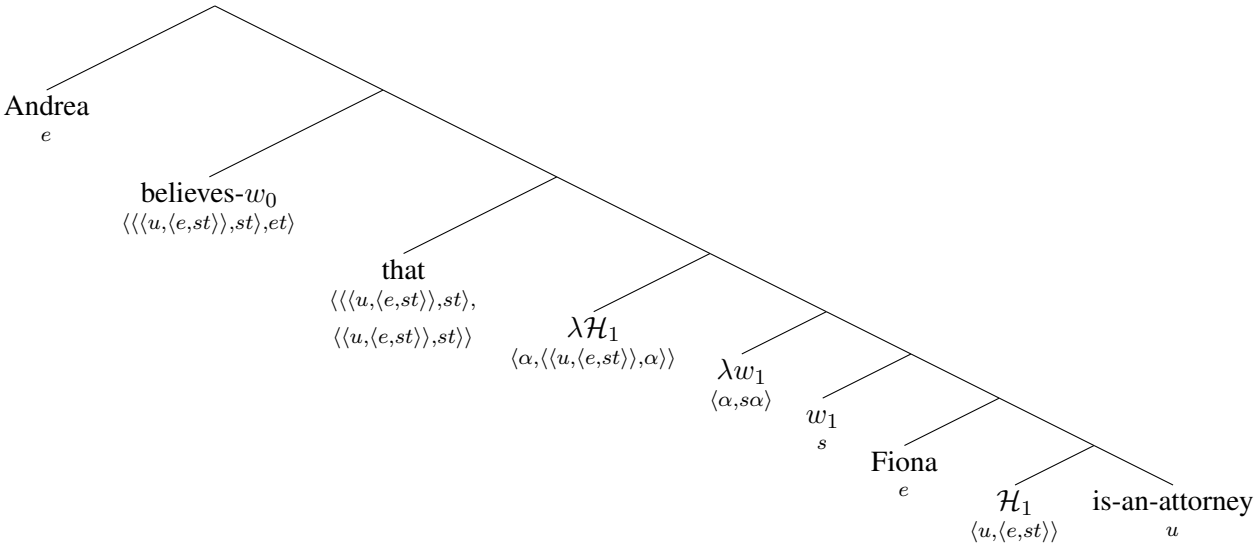
(33) a. Andrea [believes-w_0 [that [$\lambda\mathcal{H}_1$ [λw_1 [w_1 [Fiona [$\mathcal{H}_1$ is-an-attorney]]]]]]]

[26] Note that both metalinguistic concept generators and the disquotational function defined here rely on pure quotation, in the same way that Maier's mixed quotation operator does. Given that our model includes a domain of linguistic expressions, this is not problematic.

[27] This is somewhat imprecise; it would be best to formally define the interpretation assigned to any expression σ by the linguistic community at a world w and context c in virtue of the intersection of salient *individual* interpretations of σ at c, w. This would provide an interpretation of σ shared by the most salient linguistic community at w, c.

[28] To the right of $[\![.]\!]$, '$^{\Delta,w}$' abbreviates '$^{\Delta_w,w}$' to avoid clutter.

[29] In order to combine P&S's account of *de re* belief about individuals with our proposal, the entry for 'believe' should introduce existential quantification over both types of concept generators, individual and metalinguistic. I will nevertheless ignore the former here so as to avoid clutter.



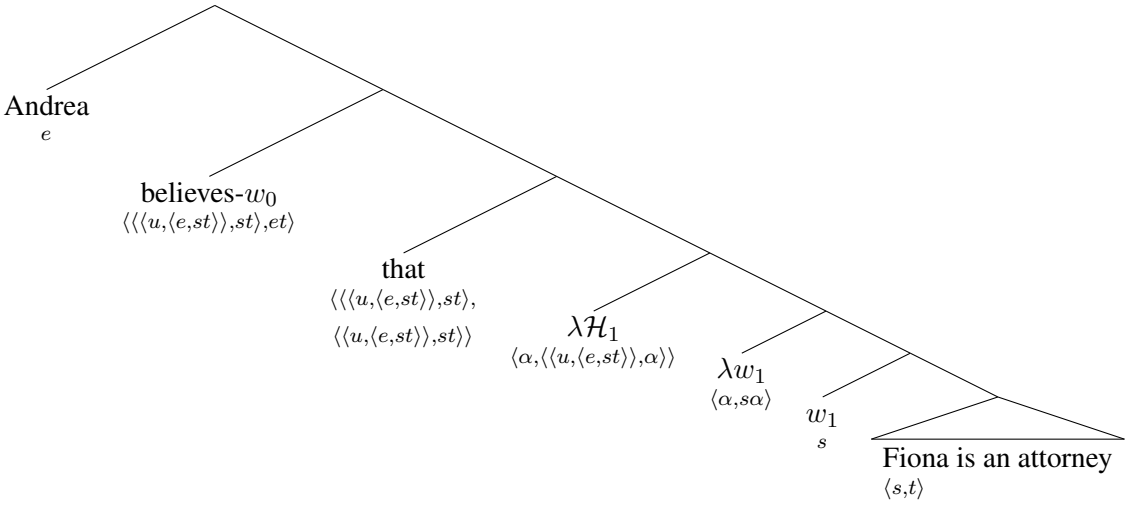
b. $[\![(1a)]\!]^{\Delta,w} = 1$ iff $\exists\mathcal{H}(\mathcal{H}$ is an $MCG_{\mathrm{A},w}$ & $\forall w' \in DOX_{\mathrm{A},w} : \mathcal{H}(\textit{attorney})(\text{Fiona})(w'))$

In words: (1a) is true relative to Δ and a world w just in case there exists a metalinguistic concept generator for Andrea at w such that, when applied to the word 'attorney', it returns a property that holds of Fiona throughout Andrea's doxastic alternatives at w. This is evindenced by the fact that, in Context A, there is such a property, namely *being a judge*.

It bears repeating that what appears next to $\mathcal{H}$ in the LF in (33a) is an expression, not its interpretation—in the same way that what appears inside quotation marks in ' "has an anomalous feature" ' is an expression and not the property of *having an anomalous feature*. Thus, the LF of (1b)—also in its opaque interpretation—contains the word 'lawyer' instead of 'attorney', and thus its truth requires that a (potentially different) relation hold between Andrea and the word 'lawyer'. A derivation of (1b) can be found in the Appendix. Both sentences turn out to be truth-conditionally independent of each other, as desired.

To account for Mates' sentences in their transparent interpretation, we need the LF in (31), where the predicate has its transparent meaning and there is no MCG-variable as sister to it. The LF for (1a)—in its transparent interpretation—would be as follows (see Appendix):

(34) a. Andrea [believes-w_0 [that [$\lambda\mathcal{H}_1$ [λw_1 [w_1 [Fiona is an attorney]]]]]]



b. $[\![(1a)]\!]^{\Delta,w} = 1$ iff $\exists\mathcal{H}(\mathcal{H}$ is an $MCG_{\mathrm{A},w}$ & $\forall w' \in DOX_{\mathrm{A},w} : \textbf{attorney}(\text{Fiona})(w'))$

In words: (1a) is true relative to Δ and a world w just in case there exists a metalinguistic concept generator for Andrea at w and Fiona is an attorney throughout Andrea's doxastic alternatives at w.

As there are no MCG-variables in the LF of the complement clause, existential quantification over metalinguistic concept generators introduced by the attitude verb is idle, and every constituent in the complement clause receives its usual semantic value.

This extension of P&S's proposal makes the necessary predictions relative to (1): quantification over the believer's metalinguistic concept generators ensures that the report is appropriately sensitive to the believer's idiolect, and we compositionally derive independent truth conditions for (1a) and (1b) in their opaque interpretation that do not jointly attribute a contradiction to Andrea. The contrast between the opaque and the transparent interpretation of the sentences in (1) lies in the fact that there are two different LFs that the complement of these belief reports can have. One of them contains an MCG-variable next to a predicate that is not interpreted; the other contains nothing of the sort.

Let us take stock. I have offered an account of the sentences in (1) according to which they are ambiguous between an opaque and a transparent interpretation. This is already an accomplishment *vis-à-vis* generalised *de re* accounts and T&S's proposal: the former could not predict independent truth conditions for (1), while the latter predicted only an opaque interpretation (at least if their proposal is not adorned with a claim about the distribution of language transformation indices). We achieved this result without quotation. Thus, in relation to the problems pointed out in §2.2, our view predicts the absence of quotation marks, as well as the lack of an inference that the purportedly quoted material was uttered. I will set aside the question of whether this proposal can be applied to Frege's puzzle (§2.3), although we may point out that, in contrast to contextualism about Frege's puzzle, our view can account for sentences like (4). In our view, multiple *rei* in the complement clause of a belief report can have MCG-variables, which can all be bound by 'believe' and return different properties. The truth conditions for (4) would be:

(35) $[\![(4)]\!]^{\Delta,w} = 1$ iff $\exists\mathcal{H}(\mathcal{H}$ is a $MCG_{\mathrm{A},w}$ & $\forall w' \in DOX_{\mathrm{A},w} : \mathcal{H}(\textit{attorney}) \neq \mathcal{H}(\textit{lawyer})(w'))$

Since the MCG-variables see the expressions and not their meaning, they are not predicted to return the same property for 'attorney' and 'lawyer'.[30,31]

[30] Lederman 2021 observes that a straightforward application of P&S's theory to Frege's puzzle, in combination with standard Millianism about proper names, cannot account for single-clause failures of substitutivity, such as (16):

(16) Plato did not believe that Hesperus was Phosphorus.

The reason for this is that, *per* Millianism, [Hesperus] = [Phosphorus], and thus placing the same concept generator variable next to both names will return the same value, i.e., $[\![G(\text{Hesperus})]\!] = [\![G(\text{Phosphorus})]\!]$. Even though I adapt P&S's proposal in crucial ways, my view about Mates' puzzle escapes this concern. The reason for this is that the argument of an MCG-variable is an expression. Since 'attorney' and 'lawyer' are different expressions, they can be mapped onto different properties by the same MCG-variable. I thank an anonymous reviewer for drawing my attention to this. See Lederman 2021 for a sophisticated account of Frege's puzzle that tackles this issue.

[31] One may wonder whether my proposal could handle Paderewski-style versions of Mates' puzzle, that is, cases based on a single predicate with different meanings. Suppose that Andrea thinks both that there are two different properties associated with the word 'lawyer' and that Fiona satisfies one of them but not the other. In this context, the following seem true:

(i) Andrea thinks that Fiona is a lawyer.

(ii) Andrea thinks that Fiona is not a lawyer.

The view offered here can handle these examples. Mates' sentences (in their opaque interpretation) introduce existential quantification over metalinguistic concept generators. In the situation described here, we can assume that the truth of (i) and (ii) is witnessed by two different metalinguistic concept generators, each mapping 'lawyer' to a different property. In addition, our view predicts

Moreover, this view does not tie Mates' puzzle to the interpretation of PPTs and thus avoids the problems that this feature of T&S's proposal causes. In particular, it does not predict that unembedded predicates can be interpreted relative to the speaker's understanding. This can only occur below attitude verbs, as only attitude verbs can bind MCG-variables. However, this does not mean that unembedded sentences like 'Fiona is an attorney' do not contain MCG-variables. We impose no such restriction. Instead, by assuming that indices of evaluation contribute a function Δ that maps any MCG-variable scoping over an expression to the interpretation of that expression according to the relevant linguistic community, we guarantee that unembedded sentences will always be interpreted transparently, regardless of whether they contain MCG-variables. Finally, the view proposed here makes use of the same basic mechanism devised by Kaplan & Lewis—*via* P&S's compositional implementation—to account for Quine's *double vision* puzzle. I take this to be a methodological advantage.

7 Pending issues

There still remain some loose ends, however. First, we owe an account of 'know' that accommodates the *datum* in (5). Secondly, the proposal just presented predicts that cross-linguistic reports should be unsuccessful—a prediction that fails for 'prime' (8). I consider these issues in turn.

7.1 'Know'

Recall the contrast between 'believe' and 'know' that T&S observed (I will switch back to their examples):

Context C

(8) John believes that 26 is prime

(36) # John knows that 26 is prime

Intuitively, (36) is infelicitous because 26 is not a prime number. Thus, 'know' blocks an opaque interpretation of 'prime'.

As we saw, T&S account for this contrast in a way that parallels the behavior of PPTs in similar environments: when embedded under 'believe', a PPT is evaluated relative to the believer's perspective; when embedded under 'know', it is evaluated relative to the speaker's. T&S offer the following analysis of 'know' (where $[\![p]\!]^s$ represents that the proposition p is evaluated relative to the speaker's perspective):

(37) S: b knows that p
  a. Presupposes: $[\![p]\!]^s$
  b. At issue:
    (i) b believes that $[\![p]\!]^b$
    (ii) b's belief is justified

In the *at-issue* component, the complement clause is evaluated relative to the believer's perspective, whereas in the presuppositional component, it is evaluated relative to the speaker's.

that there should be alternative, transparent interpretations of (i) and (ii) according to which Andrea believes a contradiction. I think this reading is possible too.

Assuming that 'prime' in (36) is subscripted with a language transformation index, the prediction of this analysis is that (36) presupposes a transparent interpretation of 'prime':

Context C

(38) S: John knows that 26 is prime$_T$

a. Presupposes: $[\![26 \text{ is prime}_T]\!]^s$ = 26 is **prime**

b. At issue:

(i) John believes that $[\![26 \text{ is prime}_T]\!]^j$ = 26 is $\lambda n.\exists x(n = x^3 - 1)$

(ii) John's belief is justified

As we saw, this gives an account of the infelicity of (36) in terms of standard presuppositional failure. Indeed, it seems impossible to interpret (36) as 'John knows that 26 is one less than a perfect cube'.

However, we resisted saying that this account of (36) offers the right analysis of its *at-issue* component. Intuitively, (36) seems not only to presuppose that 26 has exactly 2 factors, but also to assert that John thinks so.

This can be seen by considering contexts where the presupposition of (36) is satisfied and the belief of the ascribee matches either the ascribee's interpretation or the speaker's. T&S consider just this type of case (2019, n.28, p.33). They acknowledge that according to their semantics for 'know', (39) should be true in context K, even though many find it false:

> Context K: John, a monolingual English speaker, is confused about the meaning of 'prime'. He takes it to be true of a number iff that number equals $x^3 - 1$ for some natural number x. He asserts '7 is prime', correctly calculating that $7 = 2^3 - 1$. He mistakenly believes that 7 has exactly 4 factors.

(39) John knows that 7 is prime.

Since 7 is prime, the presupposition of (39) is satisfied, and since John thinks that 7 is $\lambda n.\exists x(n = x^3 - 1)$, the asserted content of (39) is true. However, I concur with T&S's reported judgments: few of my informants, when consulted about this context and sentence in Spanish, found it true (7/29).[32]

At this point, given that (36) is infelicitous in context C and that (39) seems felicitous but false in context K, one might conclude that 'know' simply has no opaque interpretation. But that is not the case, as the following variation on K shows:

> Context K*: John, a monolingual English speaker, is confused about the meaning of 'prime'. He takes it to be true of a number iff that number equals $x^3 - 1$ for some natural number x. He would not assert '13 is prime', correctly calculating that there is no number n s.t. $13 = n^3 - 1$. John knows that 13 has exactly 2 factors.

(40) John knows that 13 is prime.

[32] T&S (2019, n.28, p.33) claim that if we add to the at-issue component of their semantics for 'know' the assumption that John's justification for believing that 7 is $\lambda n.\exists x(n = x^3 - 1)$ is identical to the speaker's justification for believing that 7 is prime, the issue is solved, as such justifications are not identical in Context K. This solution depends on what T&S mean by 'having identical justification'. At least on a natural reading of that phrase, however, this is not enough, as people can have identical justification for believing different things. Suppose, for example, that John and Mary have the same math teacher. One day, she tells them: '7 is prime'. John's and Mary's justification for believing (what each would express by) '7 is prime' would be, in a sense, identical: they both received this information from the same teacher at the same time. Still, Mary's utterance of (39) would be false in this context.

According to transparent 'know', (40) should be true: 13 has exactly 2 factors, and John believes this. However, very few of my informants find this sentence true (5/29). In this case, it seems that speakers access an opaque interpretation of the *at-issue* component of 'know', and since John does not think that 13 is $\lambda n.\exists x(n = x^3 - 1)$, they judge (40) false.

T&S's view predicts a transparent interpretation of the presuppositional component of 'know' but only an opaque interpretation of its *at-issue* component. But they could solve this problem by letting 'know' take a complement clause without language transformation indices. This would offer a way of generating a transparent 'know' that is similar to how they could generate a transparent 'believe'. However, we saw in §3 that simply saying that the complement clauses of attitude verbs may sometimes contain language transformation indices is not sufficiently explanatory, as such indices are not compositionally tied to attitude verbs, and yet they can only appear below them.

I propose that we assume an ambiguity for 'know' in its *at-issue* component that parallels the ambiguity I proposed with regard to 'believe'. That is, 'know' presupposes (unambiguously) a transparent interpretation of its complement, while allowing for opaque and transparent interpretations of its *at-issue* component, in virtue of the presence or absence of an MCG-variable (as sister to an uninterpreted expression). Schematically:

(41) b knows that p.
   a. Presupposes: p
   b. At issue: $\exists\mathcal{H}(\mathcal{H}$ is an MCG for b & b believes that $p(\mathcal{H}))$

Recall that the complement clauses of attitude verbs are functions of type $\langle\langle u, \langle e, st\rangle\rangle, \langle s, t\rangle\rangle$. To generate a transparent interpretation of the complement clause in the presuppositional component of 'know', all we need to do is to plug in the index of evaluation, which includes the disquotational function Δ. The lexical entry for 'know' looks like this (presupposed material appears after the first semicolon):

(42) $[\![\text{know}]\!]^{\Delta,w} = \lambda\omega\lambda x : \omega(\Delta)(w).\exists\mathcal{H}(\mathcal{H}$ is an $MCG_{x,w}$ & $\forall w' \in DOX_{x,w} : \omega(\mathcal{H})(w'))$

Returning to our example, (5) will invariably have a transparent *presuppositional* component, but it will have either an opaque or a transparent *at-issue* component depending on whether 'know' takes (30) or (31) as its complement:

(43) Andrea [knows-w_0 [$\lambda\mathcal{H}_1$ [λw_1 [w_1 [Fiona [$\mathcal{H}_1$ is-an-attorney]]]]]]
   a. $[\![(5)]\!]^{\Delta,w}$ is defined only if $[\![(30)]\!]^{\Delta,w} = 1$. If defined,
   b. $[\![(5)]\!]^{\Delta,w} = 1$ iff $\exists\mathcal{H}(\mathcal{H}$ is an $MCG_{\text{A},w}$ & $\forall w' \in DOX_{\text{A},w} : \mathcal{H}(\textit{attorney})(\text{Fiona})(w'))$

(44) Andrea [knows-w_0 [$\lambda\mathcal{H}_1$ [λw_1 [w_1 [Fiona is an attorney]]]]]
   a. $[\![(5)]\!]^{\Delta,w}$ is defined only if $[\![(31)]\!]^{\Delta,w} = 1$. If defined,
   b. $[\![(5)]\!]^{\Delta,w} = 1$ iff $\exists\mathcal{H}(\mathcal{H}$ is an $MCG_{\text{A},w}$ & $\forall w' \in DOX_{\text{A},w} :$ **attorney**$(\text{Fiona})(w'))$

The correct interpretation of 'know' will remain an open issue. Suffice it to note that both interpretations seem preliminarily attested and that our account can make room for them.

7.2 Cross-linguistic reports

In our semantics for (1), these sentences are not guaranteed to retain their truth across translations. The reason for this is that the proposed truth conditions make crucial reference to expressions. Consider (6a):

(6a) Andrea cree que Fiona es abogada
Andrea believes that Fiona is an attorney

According to our semantics, this sentence is true just in case there is a metalinguistic concept generator for Andrea that maps '*abogada*' to a property that she thinks Fiona has. But in context A, Andrea is a monolingual English speaker, so there is no such metalinguistic concept generator. (6a) comes out false, as expected.

Something similar applies to (7): (7a) is true in virtue of the fact that Mary associates 'mother-in-law' with a property that she ascribes to Ann, while (7b) comes out false because there is no property that Mary associates with '*suegra*' and that she ascribes to Ann.

So far, so good. But recall that T&S report that their example with 'prime' can be translated into Japanese (and Spanish):

Context C

(8) a. John believes that 26 is prime.

b. John-wa 26-ga sosuu de-aru to omotteiru.
John-TOP 26-NOM prime COP COMP believes

c. John cree que el 26 es primo.
John believes that ART 26 is prime.

(8) motivated T&S to adopt the principle that I dubbed IPT, which wrongly predicts the wholesale translatability of Mates' sentences. To avoid this problem, rather than relying on IPT, I want to (tentatively) propose that the success of the translations in (8) is due to an extra-semantic assumption about the (potential) linguistic beliefs of monolingual speakers (for similar assumptions, see Heim 1992; Sudo 2014):

Observation 1 (*Default Cross-linguistic Assumption* (DCA)) *For any synonymous expressions σ,σ' s.t. $\sigma \in$ language L and $\sigma' \in$ language L': unless mentioned otherwise, assume that any monolingual speaker of L would stand in the same relations to σ' that they actually stand in to σ if they learned L'.*

DCA helps with (8). First, DCA predicts the truth of both (8b) and (8c) in C: according to DCA, John would stand in the same relation to '*sosuu*' or '*primo*' that he actually stands in to 'prime' if he learned Japanese or Spanish (that is, he would think that '*sosuu*' = '*primo*' = $\lambda n.\exists x(n = x^3 - 1)$).

Secondly, the failure of (8c) in context C* is also predicted, because we are told explicitly that John does not have the same beliefs about 'primo' that he has about 'prime'.

Thirdly, however, DCA predicts that (6a) and (7b) should be acceptable, contrary to what we observed. But this is not devastating, as we can introduce restrictions on DCA that would leave the relevant predicates beyond its scope. For example, successful translations could be restricted to *mathematical* expressions:

Observation 2 (*Default Cross-linguistic Assumption for mathematical terms* (DCA+)) *For any synonymous mathematical expressions σ,σ' s.t $\sigma \in$ language L and $\sigma' \in$ language L': unless mentioned otherwise,*

assume that any monolingual speaker of L would stand in the same relations to σ' that they actually stand in to σ if they learned L'.

DCA+ would predict successful translations of (8) but not of (6a) or (7b), as 'attorney' and 'mother-in-law' are not mathematical expressions.

In sum: in our semantics, Mates' sentences are not guaranteed to retain their semantic properties across translations. This is not problematic, however, for two reasons: (i) there appears to be more data *against* successful translations than in favor of them; (ii) successful translations could be explained as a result of an extra-semantic assumption along the lines of DCA or DCA+. A pressing question concerns the scope of successful translations. I have proposed that we restrict them to mathematical terms, but perhaps other scientific or technical expressions behave in the same way. I will leave this issue for future work.

8 Conclusion

The account presented here vindicates the claim that Mates' sentences are ambiguous between two interpretations, only one of which actually generates Mates' puzzle. To account for this interpretation, I have adapted a well-known account of *de re* belief about individuals. The basic thought is that, in their opaque interpretation, Mates' sentences are *de re* reports about expressions. Failures of substitutivity across these sentences are due to the fact that their truth crucially relies on the attitude holder's standing in certain relations to particular word(s) in the sentences that report their attitude; change these words, and the relations will be altered as well. Furthermore, I have put forward an extension of Percus & Sauerland's (2003) proposal, placing metalinguistic concept generator variables and abstractors in LF. In doing so, this theory co-opts the semantic machinery devised by Kaplan and Lewis to account for Quine's double vision cases, thus relying largely on existing technology to account for an additional range of puzzles in the semantics of propositional attitudes.

Mates' puzzle calls for at least two lines of further research, hinted at in §7. First, the scope of Mates' puzzle *vis-à-vis* propositional attitude verbs requires more attention. Factivity plays a role, and it would be interesting to explore how this phenomenon behaves in relation to other propositional attitudes, such as *semi-factive* or *veridical* verbs. Secondly, the cross-linguistic aspect of Mates' puzzle remains reasonably obscure. Sentences reporting monolingual speakers' confusion about mathematical terms seem translatable, but this does not carry over to reports about, e.g., kinship terms. In §7.2 I offer a way of generating different predictions for these classes of reports, but I do not attempt to explain the contrast.[33]

References

Anand, P. (2006). De de se. PhD thesis. MIT.

Anand, P. & A. Nevins (2004). Shifty operators in changing contexts. In: *Semantics and Linguistic Theory*. Vol. 14, 20–37.

Baron, C. (2016). Generalized Concept Generators. In: *The Proceedings of NELS*. Ed. by C. Hammerly & B. Prickett. Vol. 46. Amherst, MA: Graduate Linguistics Student Association.

[33]I'm specially indebted to Harvey Lederman for encouraging me to pursue this project. Thanks to Amir Anvari, Mora Maldonado, Uli Sauerland, Matheus Valente, Jose Manuel Viejo, and Neftalí Villanueva. This work has been funded by Portuguese national funds through FCT - Fundação para a Ciência e a Tecnologia, I.P., within project UIDB/00183/2020 and by COST Action CA17132, funded by the Horizon 2020 Framework Programme of the European Union.

Bylinina, L. (2017). Judge-dependence in degree constructions. *Journal of Semantics* 34(2), 291–331. DOI: 10.1093/jos/ffw011.
Cappelen, H. & E. Lepore (1997). On an Alleged Connection Between Indirect Speech and the Theory of Meaning. *Mind & Language* 12(3-4), 278–296. DOI: 10.1111/1468-0017.00049.
Charlow, S. & Y. Sharvit (2014). Bound 'de re' pronouns and the LFs of attitude reports. *Semantics and Pragmatics* 7(3), 1–43. DOI: 10.3765/sp.7.3.
Church, A. (1954). Intensional Isomorphisms and Identity of Belief. *Philosophical Studies* 5(5), 65–73. DOI: 10.2307/4318201.
Coppock, E. (2018). Outlook-based semantics. *Linguistics and Philosophy* 41(2), 125–164.
Cresswell, M. J. & A. von Stechow (1982). De re belief generalized. *Linguistics and Philosophy* 5(4), 503–535. DOI: 10.1007/BF00355585.
Crimmins, M. (1992). Talk about beliefs. MIT Press.
Crimmins, M. & J. Perry (1989). The Prince and the Phone Booth: Reporting Puzzling Beliefs. *The Journal of Philosophy* 86(12), 685–711.
Deal, A. R. (2017). Shifty asymmetries: universals and variation in shifty indexicality. *Manuscript. Berkeley: University of California*.
Deal, A. R. (2020). A theory of indexical shift: meaning, grammar, and crosslinguistic variation. MIT Press.
Dorr, C. (2014). Transparency and the context-sensitivity of attitude reports. *Empty representations: reference and non-existence*. Ed. by M. García-Carpintero & G. Martí, 25–66.
Fodor, J. D. (1970). The Linguistic Description of Opaque Contexts. PhD thesis. Massachusetts Institute of Technology.
Giannakidou, A. & S. Yoon (2011). The subjective mode of comparison: Metalinguistic comparatives in Greek and Korean. *Natural Language & Linguistic Theory* 29(3), 621–655.
Goodman, J. & H. Lederman (2019). Perspectivism. *Noûs*.
Heim, I. (1992). Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9(3), 183–221. DOI: 10.1093/jos/9.3.183.
Heim, I. & A. Kratzer (1998). Semantics in generative grammar. Malden, MA: Blackwell.
Hintikka, J. (1962). Knowledge and belief: an introduction to the logic of the two notions. Ithaca, NY: Cornell University Press.
Kaplan, D. (1968). Quantifying in. *Synthese* 19(1-2), 178–214. DOI: 10.1007/BF00568057.
Kratzer, A. (2006). Decomposing attitude verbs. *Talk given at the Hebrew University of Jerusalem*.
Kripke, S. A. (1979). A puzzle about belief. In: *Meaning and use*. Ed. by Margalit A. Dordrecht: Synthese Language Library (Texts, Studies in Linguistics, and Philosophy), vol 3. Springer, 239–283.
Lasersohn, P. (2005). Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy* 28(6), 643–686. DOI: 10.1007/s10988-005-0596-x.
Lederman, H. (2021). Fine-grained semantics for attitude reports. *Semantics and Pragmatics* 14(1), 1–33. DOI: 10.3765/sp.14.1.
Lewis, D. (1979). Attitudes de dicto and de se. *The Philosophical Review* 88(4), 513–543.
Lewis, D. (1983). Individuation by Acquaintance and by Stipulation. *The Philosophical Review* 92(1), 3–32. DOI: 10.2307/2184519.
Maier, E. (2007). Quotation marks as monsters, or the other way around. In: *Proceedings of the Sixteenth Amsterdam Colloquium*. ILLC Amsterdam, 145–150.
Maier, E. (2014a). Mixed quotation: The grammar of apparently transparent opacity. *Semantics and Pragmatics* 7, 1–7.
Maier, E. (2014b). Pure quotation. *Philosophy compass* 9(9), 615–630.
Maier, E. (2020). Mixed Quotation. In: *Blackwell Companion to Semantics*. Ed. by L. Matthewson, C. Meier, & H. Rullman. Wiley.

Mates, B. (1952). Synonymity. In: *Semantics and the Philosophy of Language*. Ed. by L. Linsky. Urbana, IL: University of Illinois Press, 111–136.
Moffett, M. A. (2002). A note on the relationship between Mates' Puzzle and Frege's Puzzle. *Journal of Semantics* 19(2), 159–166.
Morzycki, M. (2011). Metalinguistic comparison in an alternative semantics for imprecision. *Natural Language Semantics* 19(1), 39–86.
Partee, B. H. (1973). The syntax and semantics of quotation. In: *A festschrift for Morris Halle*. Ed. by S. R. Anderson & P. Kiparsky. New York: Holt, Rinehart & Winston, Inc, 409–418.
Percus, O. & U. Sauerland (2003). On the LFs of Attitude Reports. In: *Proceedings of the Conference Sinn und Bedeutung 7*, 228–242.
Potts, C. (2007). The dimensions of quotation. In: *Direct compositionality*. Ed. by C. Barker & P. I. Jacobson. Oxford University Press, 405–431.
Quine, W. V. (1956). Quantifiers and Propositional Attitudes 24. *The Journal of Philosophy* 53(5), 177–187. DOI: 10.2307/2022451.
Richard, M. (1990). Propositional attitudes: An essay on thoughts and how we ascribe them. Cambridge University Press.
Rieppel, M. (2017). Names, masks, and double vision. *Ergo* 4(8), 229–257.
Schiffer, S. (1992). Belief ascription. *The Journal of Philosophy* 89(10), 499–521.
Schwager, M. (2011). Speaking of qualities. In: *Semantics and Linguistic Theory*. Vol. 19, 395–412.
Sellars, W. (1955). Putnam on synonymity and belief. *Analysis* 15(5), 117–120. DOI: 10.1093/analys/15.5.117.
Shan, C.-c. (2010). The character of quotation. *Linguistics and Philosophy* 33(5), 417–443.
Sharvit, Y. (1998). Individual Concepts and Attitude Reports. In: *Proceedings of SALT VIII*, 233–248.
Sudo, Y. (2014). On De Re Predicates. In: *Proceedings of the 31st West Coast Conference on Formal Linguistics*. Ed. by R. E. Santana-LaBarge. Somerville, MA: Cascadilla Proceedings Project, 447–456.
Tancredi, C. & Y. Sharvit (2019). Qualities and translations. *Linguistics and Philosophy*, 1–41.
Wiślicki, J. (2021). Quotation as a modality. *Unpublished manuscript*.
Yalcin, S. (2015). Quantifying in from a Fregean perspective. *Philosophical Review* 124(2), 207–253.