Linguistics and Philosophy
https://doi.org/10.1007/s10988-022-09346-1

ORIGINAL RESEARCH

Fregeanism, sententialism, and scope

Harvey Lederman[1]

Accepted: 3 November 2021

Abstract
Among philosophers, Fregeanism and sententialism are widely considered two of the leading theories of the semantics of attitude reports. Among linguists, these approaches have received little recent sustained discussion. This paper aims to bridge this divide. I present a new formal implementation of Fregeanism and sententialism, with the goal of showing that these theories can be developed in sufficient detail and concreteness to be serious competitors to the theories which are more popular among semanticists. I develop a modern treatment of quantifying in for Fregeanism and sententialism, in the style of Heim and Kratzer (1998), and then show how these theories can—somewhat surprisingly—account for "third readings" (Fodor, 1970) on the model of the "Standard Solution" from possible-worlds semantics (von Fintel and Heim, 2002). The resulting Fregean/sententialist proposal has a distinctive attraction: it treats data related to counterfactual attitudes (Ninan, 2008; Yanovich, 2011; Maier, 2015; Blumberg, 2018)—which have proven challenging to accommodate in the setting of possible worlds semantics—straightforwardly as third readings.

Keywords Attitude reports · Quantifying in · Frege's puzzle · Mates's puzzle · Binding · Sententialism · Fregeanism

1 Quantifying in and third readings

In his 1892 paper, "On Sense and Reference", Frege distinguishes between two aspects of the meaning of expressions, what they *refer to* and what they *express*. He proposes

This paper has benefited from many helping hands. Andrew Bacon, Mike Caie, Cian Dorr, Peter Fritz, and Daniel Rothschild gave oral comments on it, as did audiences at the New York Philosophy of Language Workshop and the Philosophy of Language and Linguistic Philosophy conference, especially David Chalmers, Simon Charlow, Jason Stanley and Una Stojnic. Kyle Blumberg, Jeremy Goodman, Ben Holguín, Dilip Ninan, Mike Rieppel, and two referees for this journal all gave probing and detailed comments on various drafts. Finally Seth Yalcin give thoughtful and careful suggestions both on the presentation of an earlier version, and on the problems presented in Sect. 3.2. As I hope will be clear throughout, this paper is deeply indebted to his.

✉ Harvey Lederman
Harvey.lederman@princeton.edu

[1] Department of Philosophy, Princeton University, Princeton, NJ 08544, USA

Published online: 23 April 2022

that when an expression occurs outside the scope of an attitude verb (as "Hesperus" does in an unembedded use of "Hesperus is bright"), it refers to its *ordinary referent* (in this case, the planet Venus), and expresses its *customary sense*. But when an expression occurs inside the scope of a single attitude verb (as "Hesperus" does in an unembedded use of "Plato believed Hesperus was bright") it does not refer to its customary referent, and instead refers to its customary sense.

This doctrine about "reference-shift" was an important part of Frege's theory of the semantics of attitude reports. Frege held that senses are non-concrete entities, and that the ordinary referents of attitude verbs like "believe" are relations between people and structured thoughts composed exclusively of these senses. He also held that what a use of a complex expression refers to is determined by what is referred to by the uses of the elementary expressions which compose it. These doctrines—about reference-shift, the metaphysics of senses, the meaning of "believe", and the composition of what expressions refer to—provide an elegant treatment of simple attitude reports. They predict, for example, that "Plato believed Hesperus was bright" is true if and only if Plato stood in the relation that is the ordinary referent of "believe" (henceforth, the "belief-relation") to the structured thought referred to by the use of "Hesperus was bright" in this sentence. Importantly, Frege's doctrine about reference-shift is key to this result. For if "Hesperus" did not refer to a sense in "Plato believed Hesperus was bright", then the sentence would be guaranteed to be false or undefined; no one stands in the belief-relation to entities which have non-senses as constituents.

But this elegant theory leads to a well-known problem. Consider:

1. There is a planet which Plato believed was bright.

Intuitively the sentence 1 could be true. But it is unclear how Fregeans can produce this result. The expression "is a planet" plausibly refers to a property, as does "which Plato believed was bright". The sentence is true if and only if something satisfies both of these properties. But Fregeans seem committed to denying that any entity could. On the one hand, the property referred to by "is a planet" is satisfied only by concrete entities, and hence not by any senses. On the other hand, it is natural to think that if an object satisfies the property expressed by "which Plato believed was bright", then there could be an expression α whose relevant occurrence referred to that object in a true use of ⌜Plato believed α was bright⌝. Given this natural thought, Fregeans would say that the property expressed by "which Plato believed was bright" is satisfied exclusively by senses. But then 1 must be false: no sense is a planet, and no planets could be the referent of α in a true use of ⌜Plato believed α was bright⌝.

This problem for Fregeans arises also for "sententialist" theories, according to which attitude verbs express relations to sentences (see Quine (1956)). On a simple-minded sententialist theory, "believe" means "would assent to a salient translation of", and the verb takes the expression which is its complement (e.g. the sentence "Hesperus is bright") as an argument. According to this theory, "Plato believed that Hesperus was bright" is true if and only if Plato would have assented to a salient translation of "Hesperus was bright".[1] A flatfooted development of this sententialist theory would yield the result that "which Plato believed was bright" denotes a property

[1] A popular version of the theory invokes a sentence-like notion of mental representation, and takes "believe" to mean "accepts a translation of", where "accept" is a relation to mental representations ("have in one's

of linguistic expressions, and is satisfied by those expressions α such that Plato would have assented to a salient translation of $\ulcorner \alpha$ is bright$\urcorner$. But this flatfooted idea leads straight to a problem with 1 similar to the problem it poses for Fregeans. If only linguistic expressions satisfy the property expressed by "which Plato believed was bright", then, since no linguistic expressions are planets, 1 would be false.

These problems are well known, and there are also well-known outlines of solutions to them (most importantly that of Kaplan (1968), cf. Kaplan (1986)). These outlines involve rejecting the "natural thought" in the case of Fregeans, and amending the "flatfooted development" in the case of sententialists, so that the meaning of expressions like "which Plato believed was bright" denote properties of objects like planets, not properties of senses or expressions. But as Yalcin (2015) has emphasized, no compositional semantics corresponding to these outlines has yet received wide acceptance. Fregeans and sententialists alike should hope not just for appropriate paraphrases of sentences like 1, but for a systematic theory of how the meanings of such sentences are produced from the meanings of their parts: they should hope for a compositional semantics.

The first part of this paper provides a new Fregean (and sententialist) semantics for quantifying in, building on ideas from Bigelow (1978). The proposal improves on the most prominent recent proposal for how to give such a Fregean semantics, due to Yalcin (2015), in several ways. Perhaps most significantly, it allows Fregeans to treat the *de se* and the *de re* in parallel, making good on a longstanding aspiration of Fregeans (expressed powerfully by Stanley (2011)). The proposal achieves these results while preserving some of the well-known attractions of Fregeanism and sententialism, which have made these theories so popular among philosophers. For instance, it provides a straightforward treatment of Frege's own puzzle about the substitution of intuitively co-referring names, as well as Mates' related puzzle (Mates, 1952), both of which cause trouble for standard more "coarse-grained" theories.

The second half of the paper turns to a second problem, that of accounting for "third readings" or "scope paradoxes" (Fodor, 1970; Bäuerle, 1983). To introduce the problem, consider first a simple example. Suppose that Plato believed that Hesperus was bright, believed that Phosphorus was bright, believed that Hesperus was not Phosphorus, and believed that no other planets were bright. Consider:

2. Plato believed exactly one planet is bright.

This sentence has a true reading in our scenario. This reading—on which the embedded quantifier or "determiner" "exactly one" is assessed from the speaker's perspective, not from the attitude holder's—is sometimes called a "*de re*" reading. Partly to avoid the usual associations of this term, I will call it a *transparent* reading (to be contrasted with an *opaque* reading, sometimes called "*de dicto*").[2] Fregeans and sententialists have a *prima facie* difficulty producing this transparent interpretation of 2. To the

Footnote 1 continued
belief box") and translations include translations into the mental language. (See, e.g. Field (1978) and now Field (2017).) I won't consider those details further here; it's easy to see that the problems I present generalize to that more nuanced version of the theory.

[2] There are in fact two distinctions which are run together by the traditional terminology, one regarding the scope of the quantifier relative to the attitude verb ("specific" vs. "unspecific"), and the other regarding what we might call the "attitudinal status" of the determiner phrase, i.e. whether it is understood as part

extent that we have an intuitive grip on what Fregean thoughts Plato is belief-related to, Plato is not belief-related to the thought expressed by "exactly one planet is bright"; he thinks there are two such planets. Similarly, for the sententialist: Plato would not have assented to the Attic translation of "Exactly one planet is bright".

Fregeans and sententialists have claimed to have an easy answer to this problem. They say that the true reading of this sentence does not have the syntax it appears to have; the sentence is to be regimented as "There is exactly one planet which Plato believed was bright." Call this the *simple movement strategy*: according to it, the determiner phrase "exactly one planet" is assumed to move, taking scope above the attitude verb "believe". The simple movement strategy allows Fregeans and sententialists to reduce any problem with 2 to the more familiar problem of quantifying in.

But a class of well-known "third readings" and "scope paradoxes" show that this strategy is not flexible enough. Consider:

Context (Based on Fodor (1970)) Ann, Bill, Carol and Dan are running in four different races. John believes two of them lost their respective races, but he doesn't know which of the two lost. As a matter of fact, Ann, Bill, Carol and Dan all won.

3. John believes two winners lost.

This sentence is intuitively true. But John does not stand in the belief-relation to the structured Fregean thought expressed by "two winners lost", nor would John assent to (a translation of) the sentence "two winners lost". Moreover, unlike in the case of 2, the simple movement strategy does not produce the correct transparent reading here, either. For in our scenario "there are two winners who John believes lost" is false: there is no particular winner that he believes lost (never mind two of them); he merely believes that two of these four people lost.

Fregeans and sententialists have yet to provide a systematic account of these data, which also preserves the distinctive advantages of their theories.[3] And, as I will discuss in detail in Sect. 3, there are reasons for thinking that their only hope for doing so rests on postulating an *ad hoc* and unprecedented form of syntactic movement. In the second half of the paper, however, I will show that these reasons are not decisive. I will develop a Fregean version of the most popular approach to scope-paradoxical readings in possible-worlds semantics—the "Standard Solution" (as von Fintel and Heim (2002, p. 102–109) call it; cf. Percus (2000))—which does not require a problematic form of syntactic movement. I conclude by discussing a new—and, it seems

Footnote 2 continued
of the content of the ascribee's attitude (transparent vs. opaque). Arguably, then, there are two kinds of transparent readings: ones in which the quantifier is read wide-scope with respect to the attitude verb, and ones in which it is read narrow-scope with respect to the verb. In this paper, I will set aside questions about whether this first kind of reading exists independently of the second, and will focus on treating transparent readings of the second kind. I use "opaque" and "transparent" because the traditional terminology of *de dicto* and *de re* is typically associated with the claim that transparent readings must be produced by a scope ambiguity, a that claim will be rejected by some theories I consider here.

[3] Zimmermann (2018) develops an account of these phenomena, which is sensitive to some Fregean assumptions, but which I believe does not account for Frege's puzzle (as in the contrast between 9 and 10 below) or Mates's puzzle (as in the contrast between 11 and 12 below). Yalcin (2015, n. 48) suggests a direction for such an account, which I develop in more detail in Appendix B.

to me, intriguing—payoff of Fregean and sententialist accounts of third readings. Recently, a good deal of attention has been focused on data related to counterfactual attitudes which, in possible-worlds semantics, can only be accommodated by mechanisms which go beyond those used to account for third readings (Ninan, 2008; Yanovich, 2011; Maier, 2015; Blumberg, 2018). Unlike these standard accounts, however, my Fregean account accommodates some of the core data straightforwardly, handling them in exactly the way that it handles third readings. This new account thus holds some promise for unifying two phenomena which, from the vantage of possible-worlds semantics, seem to require different treatments.

One word of caution, before we begin in earnest: I have described myself to this point, and will continue to describe myself throughout, as developing a "Fregean" theory. But my goal is not to offer an interpretation of the historical Frege, and indeed I will say almost nothing about the details of Frege's writings. My theory will respect Frege's claim that expressions are associated with two dimensions of meaning, a referent and a sense, as well as his claim that propositional attitudes are relations to structured thoughts composed of non-concrete senses. In calling my theory "Fregean", I mean only that it makes good on these two Fregean ideas. In fact, as I will discuss later on, there are some significant ways in which my theory diverges from Frege's own views about the semantics of attitude reports. If these divergences from Frege's views make the reader feel that the label "Fregean" is a misnomer, I am happy to give it up. My goal is to present a theory which captures some key attractions of Frege's, and which has intrinsic interest of its own in the current theoretical landscape. It is not important for these main aims whether the proposal can or cannot be justly called "Fregean".

Section 2 presents my Fregean approach to the problem of quantifying in, and describes some applications and advantages of it. Section 3 provides my Fregean account of third readings, and describes a new application of it. Section 4 concludes. Two appendices provide further material: Appendix A shows how to extend the basic Fregean proposal to account for an array of further data; Appendix B provides formal implementations of two alternative Fregean accounts of third readings.

2 A Fregean semantics

In this section I present a Fregean semantics of attitude ascriptions which handles quantifying in. The treatment builds on ideas from Bigelow (1978). But presenting these ideas in a way which conforms to the style of semantic theory developed in Heim and Kratzer (1998) requires some new ideas.

I will develop this proposal against a background Fregean metaphysics, using notation and model theory that builds on Yalcin (2015) (who is in turn building on Stanley (2011), in turn building on a vast array of other Frege-inspired treatments). In describing my model theory I will often use distinctively Fregean terms like "sense" and "sense-composition". But I will remain largely neutral on the nature of senses and sense-composition; in particular everything I say will be consistent with taking the sense of non-variable expressions to be the expressions themselves, and taking sense-composition to be expression-concatenation. So while I will use this Fregean idiom,

everything I say will be consistent with a sententialist reinterpretation of my model theory. Often in what follows, however, I will skip over detailed discussion of issues related to sententialism, leaving it to the reader to provide that reinterpretation.

2.1 Simple sentences

I will build up my Fregean semantics in stages, by working a series of increasingly complex examples. I start with the following simple one:

4. Hesperus is bright.

The semantics will assign a pair as the semantic value of every expression: for an elementary expression, the first coordinate of this pair will be its ordinary referent, and the second coordinate of the pair will be its customary sense.[4] I'll use expressions in smallcaps to name the customary sense of elementary expressions printed in that font. For instance, "HESPERUS" is a name of the customary sense of "Hesperus". As presaged above, this notation is consistent with taking the sense of an expression to be the expression itself, but it is also consistent with understanding senses in many other ways. Using v for the planet Venus, f for the ordinary referent of "is bright" (a function which takes entities like planets as arguments), and $[\![\cdot]\!]$ as usual as a function which maps expressions to their semantic values, we would have: $[\![\text{Hesperus}]\!] = \langle v, \text{HESPERUS}\rangle$, and $[\![\text{is bright}]\!] = \langle f, \text{IS BRIGHT}\rangle$. (For now we don't need assignment functions; I'll come back to those later on.)

In working with these pairs, it will be convenient to have a way of talking about one coordinate at a time. For this purpose, I'll use π_1 and π_2 for projection operations, so that $\pi_1(\langle v, \text{HESPERUS}\rangle) = v$ and $\pi_2(\langle v, \text{HESPERUS}\rangle) = \text{HESPERUS}$.

Frege held that sense determines reference. I'll formalize this idea using a partial function on senses, Δ, that maps senses which do have referents to their referents.[5] I'll assume for simplicity that the senses of all *elementary* expressions do have referents, and that for any such elementary expression α, $\pi_1([\![\alpha]\!]) = \Delta(\pi_2([\![\alpha]\!]))$: the customary sense of an elementary expression presents or denotes the ordinary referent of that expression. The system I develop in the remainder of this section is consistent with Δ being defined also on complex expressions in a natural way. In particular, it is consistent with there there being a Δ defined on $\pi_2([\![\alpha]\!])$ for every well-formed term of the fragment α, and such that if Δ is defined on $\pi_2([\![\alpha]\!])$ then $\pi_1([\![\alpha]\!]) = \Delta(\pi_2([\![\alpha]\!]))$.[6]

[4] Sometimes Fregeans invoke a hierarchy of senses to handle iterated attitude reports. I will set aside the question of whether they need such a hierarchy here. It's fairly straightforward (if a bit involved) to modify my setup to allow for a hierarchy if one wanted one, but it won't be needed for my main examples.

[5] The symbol "Δ" was used in a related way by Church (1946, 1951, 1973, 1974, 1993), though I use the symbol for a function on senses, where he used it for a functional relation between senses and their referents. Following Klement (2002, n. 18) I sometimes pronounce the symbol by saying that senses "present" the corresponding referents, though I also often speak (less exactly) of their "having" referents (if they have them). Note that Δ is a function in my (set-theoretic) model theory; I will not consider how to give semantics for fragments of English which are extended to include expressions corresponding to Δ.

[6] In Sect. 3 my semantics will no longer be consistent with such a Δ being defined on every well-formed term, but it will still be consistent with there being such a Δ for every well-formed sentence. I discuss this point in more detail in n. 30.

In the textbook semantics of Heim and Kratzer (1998), the semantic values of predicates like "is bright" are functions, which have the semantic values of names like "Hesperus" in their domain, and the semantic value of the complex expression "Hesperus is bright" is computed by applying the function denoted by "is bright" to the denotation of "Hesperus". But in our setting, this form of straightforward composition by function application is not available: the pair $\langle f, \text{IS BRIGHT}\rangle$ is not itself a function (even though f is) and even if we could look past this point, the pair $\langle h, \text{HESPERUS}\rangle$ is not in the domain of the function f (even though h is). But there is a natural way around this difficulty. We modify the "Function Application" composition rule to require that we compute the first coordinate of the semantic value of 4 by applying *the first coordinate* of the semantic value of "is bright" to *the first coordinate* of the semantic value of "Hesperus", to deliver the result that $\pi_1([\![\text{Hesperus is bright}]\!]) = f(v)$. This idea can be stated formally as follows:

Function Application (to be revised) If α is a branching node with daughters β, γ, then if $\pi_1([\![\gamma]\!])$ is in the domain of $\pi_1([\![\beta]\!])$, then $\pi_1([\![\alpha]\!]) = \pi_1([\![\beta]\!])(\pi_1([\![\gamma]\!]))$.

This rule governs the computation of the first coordinate of the semantic value of 4. But we also need a rule for calculating its second coordinate. To do this, we assume a dedicated operation of composing senses. Just as I will aim to be neutral on the nature of senses, I will also aim to be neutral on the nature of this operation, which I will notate "$\oplus$". In particular, as presaged earlier, everything I say will be consistent with treating this sense-composition operation as the concatenation operation on expressions.[7] Using this notation, our goal is to write a composition rule that produces IS BRIGHT$\oplus$HESPERUS as the second coordinate of the semantic value of 4. In taking this to be the goal, I'm assuming that the order of the *words* in this sentence is not important for the structure of the relevant complex sense; only the order of application of semantic values is. But nothing will turn crucially on this assumption; it is just made here for the sake of concreteness. The following rule gives us the right results for our sentence:

Sense Composition (to be revised) If α is a branching node with β and γ as daughters, then: if $\pi_1([\![\gamma]\!])$ is in the domain of $\pi_1([\![\beta]\!])$, then $\pi_2([\![\alpha]\!]) = \pi_2([\![\beta]\!]) \oplus \pi_2([\![\gamma]\!])$.

Together with Function Application, this rule delivers the desired result that $[\![\text{Hesperus is bright}]\!] = \langle f(v), \text{IS BRIGHT}\oplus\text{HESPERUS}\rangle$.

The word "compositional" is sometimes used to mean "functionally compositional", where a system is functionally compositional, very roughly, if and only if the semantic values of complex expressions are produced by function-applying the semantic values of simpler expressions to one another. When I say that my goal is to provide Fregeans with a compositional treatment of quantifying in, I will not be using "compositional" in this restrictive way. My goal is to give Fregeans a systematic

[7] The question of how to understand sense-composition among Fregeans is intimately connected to the controversial question of how to understand the senses of concepts (i.e. properties). What I say below will be consistent with a variety of views. To mention just a few salient ones, it will be consistent with taking senses which present concepts to be sense-functions (as e.g. Church (1946, 1951) does), with taking senses which present concepts to be objects (as e.g. Dummett (1981, pp. 293–294) does), or with taking senses which present concepts to be distinctively incomplete senses (as, e.g. Klement (2002, pp. 65–76) does).

theory which meets a standard informally employed in most of modern semantics, and paradigmatically embodied by the textbook of Heim and Kratzer (1998), but which does not amount to requiring functional compositionality. Both of my rules above are not functionally compositional, but they are clearly sufficiently systematic to count as "compositional" in this informal and vague sense. Those who hold that the conceptual motivations for compositional semantic theories require only that an adequate semantic theory be systematic in this vaguer way should thus find the theory here satisfactory on its own. By contrast, those who hold that the conceptual motivations for compositional semantic theories require that an adequate semantic theory must in particular be functionally compositional will see this paper as taking only a first step toward a fully satisfactory theory. Whichever way one understands my theory, it will turn out to be challenging enough to develop a theory which is compositional in the informal sense, and I will content myself with meeting this challenge in this paper.[8]

I've notated the property *being bright* as f and the first coordinate of the denotation of the sentence as $f(v)$. But it will be useful in what follows to have a concrete story about what these entities are. Frege himself thought that the ordinary referents (i.e. first coordinates) of predicates like "is bright" were extensions (i.e. functions from entities to truth-values, which I'll notate "1" and "0") and that the ordinary referents of sentences were truth values. In the present setting we could impose this assumption, too. But in what follows I'll take a slightly different tack; I'll assume that the ordinary referents of predicates like "is bright" are (what I'll call) *intensional properties*: functions from individuals to functions from worlds to truth-values. Accordingly, I'll assume that the ordinary referents of sentences are (what I'll call) *intensional propositions*: functions from worlds to truth-values. While this choice (and the details of my implementation of it) will of course determine some of the development below, I do not believe it is essential to the proposal I will present in this section (the choice will matter more to my proposal in Sect. 3.2). To go further into the details we just need to make *some* choice here.[9]

[8] In fact, modifying the theory to this point to make it functionally compositional is not formally particularly hard, but it raises conceptual questions about Fregeanism which are orthogonal to my main concerns here.

To see this, note that my rule of Sense Composition does not live up to a strong form of functional compositionality as it stands because it specifies a nonfunctional rule for combining the second coordinates of expressions. We can fix this issue (if it is one), by modifying the system so that, instead of taking the second coordinate of expressions like "is bright" to be the sense of "is bright", we take it to be that function which takes every sense of an expression which denotes an individual, m, to the result of sense-composing it with the sense of "is bright". (Pickel (2017) develops a version of this strategy in a different context, and also provides interesting discussion of methodological issues surrounding functional compositionality.) This modification raises important questions about whether Fregeans should understand the relevant aspect of the meanings of predicates to be senses or merely related to senses, and if the former, whether senses compose by function application or some other mechanism. To give a sense of the array of positions here: Dummett (1981, pp. 293–294) suggests that senses of concepts are not themselves functions but that occurrences of predicates in indirect contexts have functions on senses, not senses as their referents (so that we'd take our second-coordinates to be such functions); Church (1946, 1951) understands the senses of concepts to be functions, so that they can compose by function-application directly, requiring no change to what is written in the main text; Klement (2002, pp. 66–76) holds that senses of concepts are a distinctive kind of incomplete sense, requiring a special form of (non-functional) composition.

[9] One motivation for making this choice is to abstract from issues about alethic modalities which are orthogonal to the questions which concern me here. Neo-Fregeans face a choice about how to handle alethic modalities like "necessarily": either they will say that such expressions are like attitude verbs in that

2.2 Adding attitude verbs

We now turn to the treatment of simple attitude reports, which I will introduce using the following example:

5. Plato believed Hesperus was bright.

The previous section described how to compute the semantic value of the prejacent of "believe" in this sentence; the goal of this section will be to add a lexical entry for "believe" and extend our composition rules to accommodate its composition with this prejacent.

Fregeans hold that there is an important relation between people and structured thoughts which underwrites the semantics for attitude verbs; in the case of "believe" I'll write this relation as "BEL". I will assume that for a person x and thought p $xBELp$ is an intensional proposition, a function from worlds to truth-values. The simplest proposal for the ordinary referent of "believe" is:

Believe $\pi_1([\![\text{believe}]\!]) = \lambda p : p \in D_{m_p}.\lambda x : x \in D_e.xBELp.$

This entry requires some comments on the model theory I will be assuming in the background throughout the paper. I'll use e for the syntactic type of a name or pronoun referring to an individual, and p, for the syntactic type of a sentence, with $e \rightarrow p$ the syntactic type of an adjective or verb-phrase, and so on from there. (I'll use "p" (mnemonically, "propositions") as opposed to "t", since the latter is often used for truth-values and this could cause some confusion.) I will assume that for every syntactic type σ, there are two domains—D_σ and D_{m_σ}—the first corresponding to the ordinary referents of standard expressions of this syntactic type, and the second corresponding to their customary senses. I'll assume that D_σ and D_{m_σ} are disjoint, and also that if $\sigma \neq \tau$ then D_σ is disjoint from both D_τ and D_{m_τ}.[10] Looking back at our entries in the previous section, we can say now that the first coordinate of the

Footnote 9 continued
they operate on thoughts composed of senses, and that expressions which occur in their scope change what they refer to; or they will say that these expressions are like "is bright" in that they operate not on senses but on something computed from the ordinary referents of the expressions contained in their complement clauses. If I had made the former choice, I would have owed the reader some account of these alethic modals in this paper; by contrast the implementation I've opted for lets me straightforwardly adopt the usual possible-worlds treatment of them. Of course some Fregeans and sententialists may be driven to their position by resistance to appeal to possible worlds, so that for them this choice will look bizarre. I'll come back to this issue at the end of Sect. 3.2 below.

[10] In making this assumption, I am suspending an important commitment of Frege's own, namely, that senses are individuals, i.e. elements of D_e. This Fregean commitment is also very natural for sententialists: for them, the second coordinates of expressions' semantic values are expressions, and hence individuals, i.e. elements of D_e. I won't make this Fregean assumption here because, while I don't have a proof that it would be inconsistent given the assumptions I have made explicitly, reasonable ways of extending my assumptions would be subject to a version of the Appendix B paradox of Russell (1903). (See Klement (2001) and Klement (2002, Ch. 5-7), for extensive discussion of versions of this inconsistency in a variety of Fregean systems.) I will thus use the system with type-distinctions among senses, since I am more optimistic it can be extended in a natural, consistent way, and in any case it is still rich enough to allows me to engage with the main questions I want to address. If a natural system where structured senses are elements of D_e were shown to be consistent, one could transform my system into a version of that one by simply taking D_e to contain my D_{m_σ} for all σ.

semantic value of "Hesperus" is an element of D_e (an individual), and the second coordinate is an element of D_{m_e} (a sense).[11]

The entry above says that the first coordinate of the semantic value of "believe" takes *senses* of sentences, i.e. structured thoughts or elements of D_{m_p}, as its arguments. The result is then a function which takes an individual as an argument and produces an intensional proposition as its value, i.e. a function which returns true at a world if and only if the relevant individual is related by BEL at that world to the relevant structured thought.[12]

Our goal is for the first coordinate of the semantic value of 5 to be the intensional proposition which is true at worlds where Plato is related by BEL to the Fregean thought IS BRIGHT⊕HESPERUS. Our lexical entry paves the way for this result, but our initial rule for Function Application is not sufficiently flexible to get us there. That initial rule only allows us to calculate the first coordinates of semantic values of complex expressions by applying the first coordinates of simpler expressions to the first coordinates of other simpler expressions. It would thus make the first coordinate of the semantic value of 5 undefined: the first coordinate of the semantic value of "believe" cannot compose with the first coordinate of the semantic value of "Hesperus is bright", since the latter is an intensional proposition (not a structured Fregean thought), and is not in the domain of the former. To resolve this problem, we modify the rule for Function Application as follows:

Function Application If α is a branching node with daughters β, γ:

(a) if $\pi_1([\![\gamma]\!])$ is in the domain of $\pi_1([\![\beta]\!])$ then $\pi_1([\![\alpha]\!]) = \pi_1([\![\beta]\!])(\pi_1([\![\gamma]\!]))$;
(b) if $\pi_2([\![\gamma]\!])$ is in the domain of $\pi_1([\![\beta]\!])$ then $\pi_1([\![\alpha]\!]) = \pi_1([\![\beta]\!])(\pi_2([\![\gamma]\!]))$.

The first clause, (a), simply repeats the initial rule for function application; when it is possible, we want to apply the first coordinate of the semantic values of an expression to the first coordinate of its argument. But now the rule is more flexible: (b) tells us that when it is possible we apply the first coordinate of the semantic value of an expression to the *second* coordinate of its argument. In the case of 5, the first coordinate of the semantic value of "Hesperus is bright" will be an element of D_p, and so not in the domain of $\lambda p : p \in D_{m_p}.\lambda x : x \in D_e.xBELp$. But the *sense* of the sentence (the second coordinate of its semantic value) will be an element of D_{m_p} and thus in the domain of the relevant relation. Our second clause (b) allows the first coordinate of

[11] To guarantee that expressions of higher syntactic types receive lexical entries which allow them to compose in the right way, I'll make two further assumptions: first, for any types σ and τ, $D_{\sigma\to\tau}$ is the set of functions from D_σ to D_τ; second, if $x \in D_{m_{\sigma\to\tau}}$ and $y \in D_{m_\sigma}$, then $x \oplus y \in D_{m_\tau}$. The first of these assumptions ensures that typical first-coordinates can compose by function-application to produce an entity in the right domain, and the second similarly ensures that when senses compose they produce a sense in the right domain. For instance, it guarantees that IS BRIGHT⊕HESPERUS is an element of D_{m_p}, the sense of a sentence, which I'll also call a "thought".

[12] The semantic value of "believe" does not correspond straightforwardly to the pattern we saw above for "Hesperus" or "is bright", where an expression of syntactic type σ receives a lexical entry in $D_\sigma \times D_{m_\sigma}$. The syntactic type of "believe" is $p \to (e \to p)$, but its first coordinate is not an element of $D_{p\to(e\to p)}$; it is instead an element of $D_{m_{p\to(e\to p)}}$. I assume that Fregeans will want to articulate a general principle linking syntactic type to semantic type, which allows this case while also ruling out more outlandish, but I won't take a stand on how to state such a general principle here.

the semantic value of "believe" to find this argument, producing our desired result for the first coordinate of the semantic value of the sentence as a whole.[13]

This extension of Function Application also requires an extension of our earlier rule of Sense Composition. The basic idea of Sense Composition was for the second coordinate of the semantic value of a complex expression to result from sense-composing any constituents which were applied by Function Application in producing the first coordinate. Our earlier rule for Sense Composition did this in the simple case where the first coordinate of the semantic value of a complex expression was produced by applying the first coordinate of the semantic value of one of its constituents to the first coordinate of the semantic value of another. But it is now possible that first coordinates of the semantic values of complex expressions can be computed by applying the first coordinate of the semantic value of a constituent expression to the second coordinate of a constituent expression, and we need to extend our rule for Sense Composition to handle this case as well:

Sense Composition (still to be revised) If α is a branching node with β and γ as daughters, and if $\pi_1([\![\gamma]\!])$ or $\pi_2([\![\gamma]\!])$ is in the domain of $\pi_1([\![\beta]\!])$, then $\pi_2([\![\alpha]\!]) = \pi_2([\![\beta]\!]) \oplus \pi_2([\![\gamma]\!])$.

These rules now produce the desired result that

$$[\![\text{Plato believed Hesperus is bright}]\!] = \langle \text{Plato BEL } (\text{IS BRIGHT} \oplus \text{HESPERUS}), (\text{BELIEVE} \oplus (\text{IS BRIGHT} \oplus \text{HESPERUS})) \oplus \text{PLATO} \rangle.$$

This basic theory handles simple attitude reports neatly. The theory does not formally mark Frege's idea that what an expression refers to (and expresses) changes depending on whether it is embedded under an attitude verb: regardless of the context in which it occurs, the first coordinate of the semantic value of an elementary expression will always be its ordinary referent, and the second coordinate will always be its customary sense. Since the theory does not mark what an occurrence of an expression refers to (as distinct from the ordinary referent of the expression), it also does not allow us to formalize Frege's requirement that what an occurrence of a complex expression refers to is determined by what the occurrences of its constituent expressions refer to. In a note, I discuss how my theory might be brought more into line with these views

[13] In the system of this paper, it will turn out that (a) and (b) are never in conflict. Partly, this is due to my stipulation that the domains of senses are always distinct from the domains of referents. For instance, if senses were taken to be elements of D_e, then "John runs" would fall under the ambit of both parts of the rule: (a) would produce the reasonable reading; while (b) would produce an unreasonable one on which JOHN (the sense/expression) is said to run. This problem could however be avoided by adding an extra condition to the beginning of (b), which states that it comes into effect only if (a) cannot be used. This change would require corresponding changes to Sense Composition (below) and other rules of the system, but those changes are straightforward and I won't discuss them further.

of Frege himself.[14] But I do not view the question of whether it can be as central to my project. As I have said, my goal is to develop a usable theory inspired by Frege, not to vindicate every aspect of Frege's own view. My theory *is* compositional at the level of semantic values. And it is Fregean in the sense I advertised at the outset, since it makes good on two important ideas of Frege's: that meanings have two dimensions – reference and sense; and that propositional attitudes are relations between people and structured thoughts composed of senses.

Instead of seeing the fact that the theory does not mark a change in what expressions refer to as a limitation, I prefer to see the theory as offering a new perspective on the motivations for the doctrine of reference-shift. This doctrine is motivated by the desire to have different occurrences of a single expression contribute different aspects of the meaning of the expression to the meaning of the complex expressions in which they occur. Typically, this idea is implemented by requiring that different occurrences of an expression can in an important sense have different semantic values, or even meanings. But my theory shows that there is a different way forward. We can accommodate the idea that different occurrences of a single expression contribute differently to the meaning of complex expressions without altering the semantic values of the occurrences, provided we utilize a sufficiently flexible composition rule. Such a composition rule can allow us to call on different aspects of an expression's meaning in different cases as needed, without changing the semantic values of occurrences of the expressions. The present system does just this: we do not alter what occurrences of expressions refer to or express, but instead allow different occurrences of a single expression to compose with other expressions in different ways, depending on the Fregean "context" in which they occur.

2.3 Quantifying in

With these pieces of the theory in place, we turn at last to the problem of quantifying in, exemplified by 1, which I repeat here for ease of reference:

1. There is a planet which Plato believed was bright.

In giving a treatment of this sentence, I assume that the quantifier "there is" has a syntactic type which takes two expressions of predicate-type and produces a sentence, i.e., that the sentence can be regimented as: (∃ (planet)) (7 Plato believed t_7 was bright). Here the "7" indicates that there is to be Predicate Abstraction or λ-abstraction, which

[14] We can mark what expressions refer to as follows. If an occurrence of an expression is not embedded under any attitude verb or similar operator, we say that it refers to the first coordinate of its semantic value. Otherwise, we say that it refers to the second coordinate of its semantic value. (I am setting aside issues to do with a hierarchy of senses here (see n. 4), but it's not hard to incorporate a hierarchy if one wants to.) If we mark what expressions refer to in this way, then what occurrences of complex expressions refer to will be compositionally determined by what their constituent occurrences of simpler expressions refer to. These moves bring us a significant step closer to Frege himself. But they do not get us all the way there: what complex expressions refer to will not be determined *via functional composition alone* from what simple expressions refer to. Most obviously, what occurrences of expressions embedded under attitude verbs refer to will compose with one another by sense-composition, not by function application. But it is not clear that this is an idiosyncratic feature of my treatment, since there are in principle reasons for thinking that this goal can't be achieved while preserving other important views of Frege's. See Pickel (2021) for helpful discussion and citations.

binds variables indexed by "7" within its scope. On this regimentation of the syntax of our sentence, the key question is how to handle this Predicate Abstraction.

Central to my treatment of Predicate Abstraction will be a special treatment of variables. I will assume that there are two assignment functions supplied by context, g and h. The first of these, g, is an ordinary assignment function which maps indices to elements of D_e. The second, h, maps individuals to senses, i.e. elements of D_{m_e}. Russell (1905) famously says that there is no "backward road" from reference to sense. But in our setting, in context, h will provide just such a backward road, choosing one distinguished sense for each referent. Given this backward road in context, we interpret free pronouns by the rule:

Traces $[\![t_i]\!]^{g,h} = \langle g(i), h(g(i))\rangle$

and use the following rule for abstraction:

Predicate Abstraction If α is a branching node with daughters β and γ, where β dominates only a numerical index i of type e, then for any assignment functions g and h,

(a) $\pi_1([\![\alpha]\!]^{g,h}) = \lambda x : x \in D_e.\pi_1([\![\gamma]\!]^{g[x/i],h})$;
(b) $\pi_2([\![\alpha]\!]^{g,h}) = \text{LAMBDA- I} \oplus \pi_2([\![\gamma]\!]^{g,h[\text{I}/g(i)]})$.

Here and throughout, I will use $g[x/i]$ for the function such that $g[x/i](i) = x$ and $g[x/i](j) = g(j)$ for all $j \neq i$ in the domain of g. This applies not only for the case where a function has indices as its domain, as g does, but also for functions like h, which have other domains. In clause (b), for example, $h[\text{I}/g(i)]$ is the function such that $h[\text{I}/g(i)](g(i)) = \text{I}$ but for all $y \neq g(i)$ $h[\text{I}/g(i)](y) = h(y)$.

This rule allows us to compute the semantic value of "which Plato believed was bright", while, crucially, still allowing the first coordinate of that semantic value to be a function on individuals, as I will now show. By Predicate Abstraction, we have that: $\pi_1([\![7 \text{ Plato believed } t_7 \text{ was bright}]\!]^{g,h}) = \lambda x : x \in D_e.\pi_1([\![\text{Plato believed } t_7 \text{ was bright}]\!]^{g[x/7],h})$. Since BEL isn't defined on first arguments of semantic values, computing this semantic value requires that we use clause (b) of Function Application; the expression can be simplified to: $\lambda x : x \in D_e$. Plato BEL $(\pi_2([\![t_7 \text{ was bright}]\!]^{g[x/7],h}))$. To see the mechanics of this expression consider first the semantic value of the trace t_7 relative to the assignment functions $g[x/7], h$. Since in general $[\![t_i]\!]^{g,h} = \langle g(i), h(g(i))\rangle$, here we have $[\![t_7]\!]^{g[x/7],h} = \langle g[x/7](7), h(g[x/7](7))\rangle = \langle x, h(x)\rangle$; the pronoun is assigned an individual by the shifted first assignment function, and also a sense of that individual by the (unshifted) second assignment function. Crucially, the variable is assigned a sense as its second-coordinate even though the function denoted by the λ-term is a function on individuals, not on senses. The clause for variables thus allows abstraction to produce a property of individuals like planets (elements of D_e) because it only shifts the first assignment function, while nevertheless allowing us to compute the semantic value of the whole clause, because the second assignment function assigns the trace an appropriate sense. This is the trick that solves the problem of quantifying in; from here it's easy to see that things will work out. The whole λ-term computes to: $\lambda x : x \in D_e$. Plato BEL (WAS BRIGHT$\oplus$ $h(x)$). And thus if the value of h on the planet Venus is either HESPERUS or PHOSPHORUS, the sentence will come out true. There will be an object, namely Venus, which is a planet,

and whose value under h is a sense m such that Plato is belief-related to the result of composing m with the sense of "is bright".[15]

Following Bigelow (1978), the key move here is to assume, in context, a backward road from reference to sense, or, in the sententialist's idiom, a distinguished expression for each individual. The idea that objects which are relevant to a conversation are associated with a salient way of thinking about them seems quite plausible. The formal object h generalizes from this central case to a function from all objects to senses of them. The assumption that context supplies such a function does not seem problematically more demanding than the standard assumption that context supplies a normal assignment function, g, which maps numbers to objects. Natural stories about what it is for context to provide such a g generalize straightforwardly to stories about what it is for context to provide such an h.

The second coordinate of an abstraction—represented by clause (b) here—is a little more abstruse than the first coordinate, but it is conceptually quite important. First, we need a way of computing appropriate senses for binding that occurs within the scope of an attitude verb, for instance on the most salient reading of "Plato believed that some of the gods did not love their children". On this reading, the fact that "their" is bound by an abstraction below "some of the gods" must be reflected in the sense of the complement clause of "believed" so that the whole sentence says that Plato was belief-related to the thought that some of the gods did not love *their own* children. The mechanism for preserving this binding structure must also work when there are multiple embeddings in a sentence, and when abstraction binds across an attitude verb, as in "John learned that there is a planet which Plato believed was bright." In fact, this second coordinate is important even in unembedded uses of 1. For Fregeans will naturally hold that the content asserted in an assertion is what the participants in a conversation come to believe if they accept that assertion, and since they hold that structured Fregean thoughts are the objects of belief, they will hold that such assertions have as their asserted content the second coordinate of their semantic value, which must therefore reflect the binding structure of the sentence as well.

The way in which our clause for the second coordinate achieves these results requires a few further assumptions, but is formally straightforward. We assume that for every variable-index i of type e, there is a sense LAMBDA- I of type $m_{p\rightarrow(e\rightarrow p)}$; this means that sense-composing this sense with one of type m_p yields a sense of type $m_{e\rightarrow p}$.[16] We also assume that corresponding to each such sense, there is a special

[15] Heim and Kratzer's rule for Predicate Abstraction is not compositional (see e.g. Rabern (2012), Yli-Vakkuri (2013)), and my rule inherits this feature of it. As I said earlier, my goal is to meet the standard of Heim and Kratzer on behalf of the Fregean, not to exceed it. To go beyond Heim and Kratzer here, my style of Fregean could use standard ways of producing a (functionally) compositional theory, for instance, by taking expressions' semantic values in general to be functions on assignment functions (e.g. Janssen and Partee (1997), with discussion in Pickel (2017, §7)). But it is unclear whether the resulting theory will really be Fregean, for reasons somewhat related to those mentioned in n. 8, concerning whether Fregeans can allow the relevant dimension of expressions' meaning to be (for example) functions from assignment functions to senses rather than senses themselves. As with the earlier issues, I will leave these hard questions to others, and will be content here to produce a theory that lives up to the informal standard of compositionality alluded to earlier, and embodied by Heim and Kratzer themselves.

[16] These ideas could be extended to variables of types other than e, and abstraction over nodes of types other than p, but we won't need the general versions of the ideas here.

sense of type m_e, I, understood to mark the place where the binder would bind. Given these assumptions, clause (b) ensures that the second coordinate of sentences which include an abstraction will display the binding structure by the coordination of the same numeral in a sense of a λ-term as occurs in the place where a trace would be required, for instance, a relevant occurrence of "LAMBDA- 7" coordinated with relevant occurrences of "7".[17]

Claiming that ordinary people stand in the belief-relation to thoughts containing a sense notated LAMBDA- 7 might seem to require, absurdly, that everyone understands the λ-calculus as used in semantics. But my notation should not be understood in this way. People do think thoughts which differ in ways corresponding to what our formal system represents as differences in binding structure. Everyone, Fregeans included, must hope for some theory of how the contents of such thoughts differ. My rule for the second coordinate of Predicate Abstraction is stated using a particular way of conceptualizing those differences in content; it assumes that binding is represented in thought in a way that structurally mirrors the way it is notated in formal semantics. But nothing in my system depends on the details of this toy theory of how binding is represented in thought. Essentially any reasonable theory could be substituted here; we can think of the rule for the second coordinate schematically as saying: "deliver a thought which represents the binding structure appropriately, while preserving other aspects of the relevant sense".[18]

[17] Our assumptions to this point leave open an odd possibility. Consider again the sentence "Plato believed some of the gods did not love their children". It is standard to allow for infinitely many distinct syntactic regimentations of the intended reading of this sentence, varying by which index is assigned to relevant traces. The syntax corresponding to "Plato believed (some of the gods) 1 (t_1 did not love t_1's children)" is distinct from that corresponding to "Plato believed (some of the gods) 2 (t_2 did not love t_2's children)", and so on. These distinct syntaxes should not generate *semantically* distinct readings: the multiplicity here is simply a byproduct of flexibility required for sentences with multiple binders or free pronouns. But our system to this point leaves open the possibility that the different syntaxes *would* lead to semantic differences. For the first will be true if Plato was belief-related to a thought which consists in part of the senses LAMBDA- 1 and 1, while the second will be true if Plato was belief-related to a thought which consists in part of the senses LAMBDA- 2 and 2, and these are, as far as we have said, different senses.

A Fregean or sententialist might respond to this observation in three different ways, the third of which I'll take as part of the official position of the paper. The first (in my view the most plausible position for a sententialist) is to hold that the standard formal system for variable binding is at fault, and must be altered. While this system is fine for most purposes, the sententialist might argue, when we look closely at belief-reports, we see that notation matters, and we should hope for a system which generates only syntaxes that align more closely with the semantic facts, perhaps the wire diagrams of Quine (1940/1981) or something else (see e.g. King (2007, Appendix) and Higginbotham (1991) for thoughts along these lines). A second response (only available to the Fregean) is to hold that however the notation of binding works, the thought expressed by a sentence regimented according to syntax s is identical to the thought expressed by a sentence regimented according to syntax s' if s' differs from s only by appropriate relabeling of bound variables (i.e. to any α-equivalent sentence). (Bigelow (1978) has a nice way of producing this result systematically (see pp. 124–125 for discussion).) A third response (available to both the sententialist and Fregean) is to hold that the BEL relation is such that, necessarily, if a person bears it to the thought expressed by a syntax s in context, they bear it to any thought which is expressed by a syntax s' which differs from s only by the appropriate relabeling of bound variables (i.e. to any α-equivalent sentence). As I have said, I will assume in what follows that Fregeans and sententialists adopt this third response, mainly because it can be endorsed by proponents of both positions.

[18] A different concern is that, unlike "Hesperus", "λ" is not an expression which is assigned a denotation; it is not a term of our language at all, but a syncategorematic marker of binding structure, and thus, LAMBDA- 1 is a sense which does not have a corresponding denotation. This issue is subtle and requires more discussion

2.4 Situating the account

This completes the presentation of my basic proposal for a Fregean/sententialist semantics which handles quantifying in. In this section and the next, I will situate this account. In this section I note three contrasts between the view developed here and the most prominent recent proposal for how Fregeans should handle quantifying in, that of Yalcin (2015). In the next, I offer some rough comparisons between the present account and standard accounts of attitude reports in possible-worlds semantics. In Appendix A I further show how the account can be extended to handle a wider array of data, and to produce truth-conditions featuring existential quantification over senses, which many Fregeans have aimed to produce. I do not present these details in the main text, since they will not be relevant to the main thread of development, but I think they are an important component of a full Fregean theory.[19]

A first difference between the present account and Yalcin's concerns sentences like "Someone believes they are a spy" and "Some boy's mother believes he is a spy", in which a binder binds a pronoun both outside the scope of an attitude verb and inside the scope of an attitude verb. Yalcin's theory predicts either that these sentences are undefined or that they are false.[20] But intuitively they can be true. The present proposal allows these true readings.

A second difference concerns the logic of numerical quantifiers. Recall that there are eight planets. Yalcin's proposal allows the following to be true in a single context:

6. There are eight planets which Plato knows are bright, and eight planets which Plato does not know are bright.

The present proposal, by contrast, predicts that this sentence cannot be true in a single context. While I won't develop an account of how Fregeans should think about validity or entailment, the fact that the present approach predicts that the sentence above is true in no context at all suggests that it is in a better position to vindicate intuitive entailment relations among sentences featuring numerical quantifiers (e.g. it vindicates the claim that in every context the inference from "there are eight planets which Plato knows are bright" to "there are no planets which Plato does not know are bright" preserves truth).

A third contrast concerns the *de se*. The present proposal, unlike Yalcin's, allows Fregeans to handle the *de se* by analogy to the *de re*. This last point is perhaps particularly conceptually significant. One of Yalcin's central targets in his paper is the Fregean, intellectualist treatment of "know-how" ascriptions presented by Stanley

Footnote 18 continued
than I can give it here. But I will note that Frege also held that some senses (e.g. the customary sense of an empty definite description) do not have a corresponding referent, so the mere fact that these senses have no referent is not problematic in itself. Moreover, the limitation does not affect the prospects of assigning denotations to complex senses which *are* in the domain of Δ; it is consistent for instance with Δ mapping the second coordinate produced by Predicate Abstraction to the first. So, as I claimed earlier, the second coordinates of well-formed expressions can still be mapped to the first by Δ.

[19] The points I make in this section in comparing my account to Yalcin's hold equally for the theory in Appendix A.

[20] Whether they are false or undefined depends on whether the sense-domains in his appendix are defined to be disjoint from the domain of individuals, and whether sense-composition is defined on the latter.

(2011) (cf. Pavese (2017, Appendix C, D)). Yalcin argues that, given Yalcin's own semantics, Stanley's claims that Fregeans can handle the *de se* in parallel to the *de re* are false. But, as I show in the attached note, the present proposal allows us to make good on Stanley's basic idea within a Fregean framework.[21]

2.5 Logical omniscience and Mates' puzzle

The present theory thus compares favorably to Yalcin's. But, one might wonder, is there linguistic motivation for being a Fregean or sententialist in the first place? In this section I rehearse some standard reasons for favoring Fregeanism/sententialism by comparison to the benchmark semantics of Hintikka (1962), as well as some well-known modifications of it.

In the Hintikka semantics, $\ulcorner S$ believes $\varphi\urcorner$ is true in a context if and only if φ expresses an intensional proposition in that context which is true at every (metaphys-

[21] The formal ideas I'll discuss here are due to Santorio (2014), but I will rehearse them in our present Fregean context.

There is a standard contrast between:

7. Bekele wants to win.
8. Bekele wants himself to win.

(The examples are Yalcin's.) If Bekele sees himself running a race on TV, but doesn't recognize himself, one can use 8—but not 7—truly to describe the situation. The question—one which is central to the account of know-how in Stanley (2011)—is whether Fregeans can provide a semantics that predicts this contrast. Following in a distinguished tradition, Stanley (2011) hopes that if Fregeans postulate a distinguished class of senses—"first personal senses"—they can restrict the quantification over senses introduced in their semantics for the *de re* to first-personal senses in the case of the *de se*. But Yalcin shows that if Fregeans adopt Yalcin's approach to quantifying in, this hope will be dashed. A standard syntax for 7 has a covert pronoun PRO occurring inside the scope of the attitude verb, which is bound above the attitude verb:

9. Bekele [1 t_1 wants [PRO_1 to win]].

Yalcin (2015, Section 11) argues that, given this syntax and Yalcin's semantics, we cannot simply restrict the quantification over senses to "first-personal" modes of presentation of Bekele.

But in the present system it is straightforward to impose the requirement that the quantification over senses be restricted to first-personal ones for obligatorily *de se* pronouns. We assume that the variable corresponding to PRO carries extra information; for instance, in addition to having an ordinary numerical index, it also has an extra index $*$. We then alter the Traces rule so that: $[\![t_i]\!]^{g,h} = \langle g(i), h(g(i))\rangle$ and $[\![t_{i,*}]\!]^{g,h} = \langle g(i), h(\langle g(i), *\rangle)\rangle$; assume that every sense-assignment function h is defined not only on individuals, but also on pairs consisting of an individual and $*$; and assume that its value on such pairs is a first-personal sense of the relevant individual.

Given these assumptions, 7 will receive the following truth-conditions:

10. Bekele is desire-related to WINNING$\oplus m$, where m is the salient first-personal sense of Bekele.

Using the extension of the present proposal described in Appendix A, we can go further, to produce almost exactly the truth-conditions Stanley (2011, p. 88ff.) hoped for (although our syntax for these sentences is a different, more usual one, than the one he postulates). In the extended system, we require that if t is an element of $f(x)$, then t maps first-personal senses of x to first-personal senses of x. The truth conditions for 7 will then be:

11. Bekele is such that there is a salient first-person sense m of Bekele and a salient sense f whose denotation is the property of winning, such that Bekele is desire-related to $f \oplus m$.

The present proposal therefore provides at least partial vindication of Stanley's ambitions for a Fregean semantics that handles the *de se* straightforwardly.

ically/logically) possible world consistent with *S*'s beliefs. This theory thus renders the following two ascriptions equivalent:

7. John believes that the axioms of Peano arithmetic imply that the biggest prime number is greater than five.
8. John believes that the axioms of Peano arithmetic imply that there are positive integers a, b, c and $n > 2$ such $a^n + b^n = c^n$.

The complement clauses of both ascriptions are true at exactly the same (metaphysically/logically) possible worlds (i.e. at none), and hence the Hintikka semantics predicts that the first is true if and only if the second is. This is a bad prediction, since intuitively the first might be false, while the second is true. The present theory avoids this bad prediction, since the theory is consistent with holding that the senses of the complement clauses are distinct, and hence that the sentences have different truth values in a single context.

As has been much discussed (e.g. Lewis (1970), Cresswell (1975), Cresswell and von Stechow (1982), Cresswell (1985)), one can accommodate the difference between 7 and 8 without being a Fregean or a sententialist. One idea (put very roughly) is to hold that the semantic values of complex expressions are syntactic structures decorated with the semantic values of the simple expressions which are their terminal nodes, and to hold that these semantic values are (roughly) the standard ones the expressions would have in possible-worlds semantics. (So "Hesperus" is associated with just an individual, "is bright" with an intensional property, and so on.) While this approach faces many technical difficulties, which have as yet received no widely accepted resolution (see von Stechow (1985) for discussion of some in Cresswell's theory), it offers the promise that one could accommodate a difference between 7 and 8 without endorsing Fregeanism or sententialism. Since the two complement clauses of "believe" in these sentences have different syntactic structures, the kind of "structured" theory I have just described will guarantee that they have distinct semantic values, allowing for a difference in truth-value between the two reports.

But as has also long been recognized, even if this more flexible theory can be developed in a consistent way, it is plausibly still not flexible enough to account for all related data. A structured theory of propositions, plus a Millian theory of proper names—on which the semantic value of a name is determined by the object it names—predicts that the following are equivalent:

9. Plato believed that Hesperus rises in the evening.
10. Plato believed that Phosphorus rises in the evening.

But intuitively the first could be true in circumstances where the second is false. The problem for the structured approach just described is that on their face these sentences have the same syntactic structure, and also have the same semantic values in each element of the structure (since we are assuming a Millian theory of proper names) so that a structured theory on its own is not sufficiently general to produce is the intuitive contrast. The theory I have developed in this paper, however, produces these results straightforwardly, since the sense of "Hesperus" is presumed to be distinct from the sense of "Phosphorus".

One might think that the foregoing problem arises only for a particular theory of the meaning of *names*. But crucially it is much more general. Mates (1952) presented

essentially the following problem as a counterexample to an early version of the above ideas found in Carnap (1947):[22]

Context Barbara, a monoglottal English speaker, thinks that groundhogs are blind like mole rats, while woodchucks are sighted and are often seen above ground. In fact, woodchucks just are groundhogs. There is a woodchuck/groundhog who lives in Barbara's neighborhood, who is known to Barbara and other locals as "Alonzo". Barbara recognizes that Alonzo is a woodchuck, but she thinks he is not a groundhog; in fact she believes she's never seen a groundhog:

11. Barbara believes that Alonzo is a woodchuck.
12. Barbara believes that Alonzo is a groundhog.

Since "woodchuck" and "groundhog" express the same intensional property, the structured theory sketched above will predict that 11 and 12 are equivalent. But intuitively they are not: 11 is acceptable in describing the story, while 12 is not. Once again our Fregean has the resources to make the correct prediction here, provided (as is typically assumed by Fregeans) that the sense of "woodchuck" is not the sense of "groundhog".[23]

There are many further questions to be asked about the empirical coverage of the account. I address some of these in Appendix A.[24] But my goal here is not to offer anything like an exhaustive treatment. This brief discussion is just meant to give context for why one might hope that a Fregean or sententialist theory will indeed work out in the end.

One might think that all the points I have made are moot, because the present theory is so much more complex than standard possible-worlds-based theories, and hence not worth exploring, regardless of its other putative advantages. This prejudice against Fregean theories seems to me based on an important mistake. The basic Hintikka semantics covers none of the data discussed in this subsection, and it is not fair to compare a theory which can accommodate these data to one which cannot. Moreover,

[22] See also Partee (1973) and Higginbotham (1991), the latter of whom directly applies this problem to the theory of Cresswell and von Stechow (1982).

[23] The intuitive contrast between 9 and 10, and between 11 and 12 can be explained by a straightforward extension of a coarse-grained theory, which assumes that attitude reports are context sensitive, and that the sentences are naturally interpreted in different contexts (for discussion see Schiffer (1979), Crimmins and Perry (1989), Dorr (2014), and now Goodman and Lederman (2021)). But natural versions of this contextualist approach cannot explain a similar contrast for slightly more complex sentences, e.g. the contrast between:

(i) Plato believed that Hesperus shares its orbit with Phosphorus; and
(ii) Plato believed that Hesperus shares its orbit with Hesperus.

See Goodman and Lederman (2021, §9) and Lederman (2021, §2) for discussion. So, while going versions of this kind of contextualism can do significantly better than non-contextualist theories in accommodating some relevant data, they still do not accommodate all the data the Fregean can handle.

[24] Some I do not address there, but which I believe the account handles straightforwardly, are the subtle data discussed in Tancredi and Sharvit (2020) and now Soria Ruiz (forthcoming). A more challenging set of data that has been much discussed are the *de qualitate* reports of Schwager (2009). It would take me too far afield to discuss these here, but suffice it to say that I think the reports she discusses are either examples of Mates's puzzle (as in her "buyers' intentions" and "Burj Dubai" cases), or should be handled by a mechanism analogous to that discussed in Abreu Zavaleta (2019) and Blumberg and Lederman (2020) ("Foyle's investigation"). For a similar take on the first class of cases, see Kusliy and Vostrikova (2019).

if we compare the present theory to a genuine competitor, which has similar empirical coverage, it is not at all clear that the present one is more complex. For instance, theories which build on Percus and Sauerland (2003) (see Anand (2006), Charlow and Sharvit (2014)) to give an account of 9 and 10 must modify standard syntax to include variables over concept-generators and must postulate that the relation expressed by "believe" is multigrade (or else that the verb is ambiguous between countably infinite meanings) (Baron, 2015). And in fact, as they stand these theories still do not account for data closely related to 9 and 10 (see n. 23 and Lederman (2021)). Alternatively, theories like Ninan (2012) or Rieppel (2017) (which can account for these further data) make attitude verbs operate on assignment functions in a way that is at least as unfamiliar as the distinctive second coordinates used here. Most importantly, neither of these theories on their own have their resources to accommodate the contrast between 11 and 12. Recent extensions to handle such further data all involve a great deal of extra machinery and complexity, which makes them look at least as complex as the present theory (Tancredi and Sharvit, 2020; Soria Ruiz, forthcoming).

3 A Fregean account of third readings

3.1 The problem

I now turn to the second problem of the paper. To recall, the goal is to provide a Fregean account of "third readings" and "scope paradoxes" (Fodor, 1970; Bäuerle, 1983). These readings can be illustrated by the following scenario, repeated from the introduction:

Context (Based on Fodor (1970)) Ann, Bill, Carol and Dan are running in four different races. John believes two of them lost their respective races, but he doesn't know which of the two lost. As a matter of fact, Ann, Bill, Carol and Dan all won.

3. John believes two winners lost.

This sentence has a true reading. But the true reading cannot be the usual opaque one: John does not stand in the belief-relation to the structured Fregean thought expressed by "two winners lost". And it also cannot be the reading produced by interpreting "two winners" outside the scope of "believes" in line with the simple movement strategy; in this scenario "there are two winners who John believes lost" is false. There is no particular winner that John believes lost; he merely believes that two of these four people lost.

How might Fregeans account for this datum? It is natural to start by considering existing accounts in possible-worlds semantics, to see whether Fregeans can adapt them to their purposes. In possible-worlds semantics, two different kinds of approach have received most attention.[25] A first approach—which I will call "the sophisticated movement approach"—postulates a new form of movement. On a simple version of this approach, for instance, it is postulated that a predicate (in our case "winners")

[25] Von Fintel and Heim (2002, Ch. 8) and Keshet and Schwarz (2019, §5) both provide excellent surveys, along with summaries of considerations for and against different ways of handling these data.

moves outside the scope of the relevant attitude verb ("believes") and, as a result is interpreted differently than it otherwise would be. In 3, the relevant interpretation could be paraphrased very roughly as "the winners are such that John believes two of them lost", which is intuitively the correct reading.[26]

There are some challenges in the details, but it is fairly straightforward to develop a Fregean version of this approach, as was suggested already by Yalcin (2015, n. 48). I present such an account in Appendix B. Although doing so is not technically difficult, it is significant that it can be done. If Fregeanism had no account of third readings, this would be a powerful reason to reject the theory.

But there would still be reason to be dissatisfied if this were all that Fregeans could achieve. While the sophisticated movement approach can deliver the correct truth-conditions for third readings, it remains controversial whether this approach is consistent with a reasonable syntax. I will not rehearse these considerations here, since they have been well discussed in other places (see von Fintel and Heim (2002, Ch. 8.4), Keshet and Schwarz (2019, §5.4)). But doubts about the plausibility of the syntactic movement postulated by versions of this approach have been an important factor leading many semanticists to reject it, and to prefer instead a different approach to third readings, which does not require syntactic movement of the problematic kind. Indeed, this second approach has become the benchmark approach to third readings, so much so that in their textbook von Fintel and Heim (2002, pp. 102–109) call it the "Standard Solution". Given this situation, if Fregeanism and sententialism could *only* offer a movement-based account of the scope paradoxes, these theories would be "hostage to fortune". They would depend on controversial views about syntactic movement, and hence depend on how delicate issues in syntax are resolved. By contrast, since standard possible-worlds theories can invoke the Standard Solution, such theories do not depend on how these syntactic questions are settled, and thus are not hostage to fortune in this way.

At first sight, this problem for Fregeanism and sententialism might seem to be inescapable.[27] A guiding idea behind both theories is that lexical material in the complement clause of an attitude verb should receive a distinctive semantic treatment, indicative of the fact that it contributes to the content of the reported attitude. As a result, whether an expression is interpreted as contributing to the content of a reported attitude or not—what we might call the *attitudinal status* of the expression—is determined by its scope with respect to relevant attitude verb. This tight connection

[26] Different forms of movement can produce the same result, for example, if the quantifier moves but leaves a trace of a different type (usually $(e \to p) \to p$). I discuss this second kind of movement in more detail in Appendix B.

[27] For the following points, see Yalcin (2015, pp. 240–241).

between syntactic facts about relative scope and semantic facts about attitudinal status seems baked in to Fregeanism and sententialism. And thus it might seem that syntactic movement, which changes the relative scope of key expressions, is the only hope these theories have for producing third readings.

The prospects for Fregeans and sententialists to develop a non-movement-based account of third readings initially seem bleak. And when we look at the details of the Standard Solution—the leading non-movement-based account in possible-worlds semantics—the prospects for such an account seem, if anything, even bleaker. On the Standard Solution, it is postulated that world- or situation-pronouns appear in the syntax of English, and different patterns of binding for these pronouns are used to explain the contrast between opaque and transparent readings, by shifting whether the relevant lexical material is interpreted at the worlds of an ascribee's attitudes, or at different worlds (Percus, 2000). It is hard to see how Fregeans or sententialists could imitate these results. Neither Fregeans nor sententialists have a notion of "belief-world", so it is unclear how they could use world-pronouns to shift what I earlier called the "attitudinal status" of any lexical material.

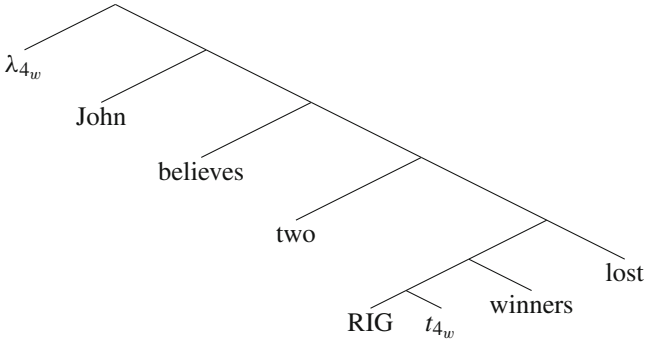
But in the rest of this section I will show that, in spite of these *prima facie* challenges, Fregeans and sententialists *can* develop a version of the Standard Solution, and thus that neither of these views is "hostage to fortune" in the manner described above. The Fregean account of third readings that I will present will be somewhat intricate. In my view, the intricacies are illuminating: they shed light on subtle conceptual issues for Fregeans, as well as on the mechanics of the Standard Solution itself. But some may be satisfied with knowing that Fregeans can provide a non-movement-based theory of third readings, without needing to understand in detail how exactly this is done. Those readers may skip to Sect. 3.3.

3.2 Detailed implementation

On the most familiar versions of the Standard Solution, every expression of the language takes a world-pronoun as its argument (Percus, 2000). I will start from a slightly different approach, on which only special expressions have world-pronouns as arguments.[28] The main reason for this choice is to simplify the technical presentation. Everything I'll do below could have been done starting from a more standard version of the Standard Solution; it just would have been more complex.

I will assume that the intended reading of 3 arises on the following syntax:

[28] I take inspiration here from Schwarz (2012), and some of my technical development is indebted to his. But Schwarz's goals are much more ambitious than mine; he wants to use a restriction on which expressions take world-pronouns to predict the distribution of these pronouns and rule out unwanted readings. My approach won't achieve those more ambitious aims; it will allow some unobserved readings, and (like the more standard Standard Solution) must rely on further assumptions about syntax to rule them out. I'll return to this point later on.



I will build up to my interpretation of the whole sentence starting from the key expression "RIG t_{4_w} winners".

Let's begin with t_{4_w}. Following the Standard Solution, I will assume that there are world-pronouns which occur in the syntax of natural language, and which belong to a new syntactic type s. I will assume moreover that these world-pronouns are associated with special indices of the form n_w for natural numbers n and that the functions g and h are extended to assign worlds (elements of a new D_s) and senses of worlds (elements of a new D_{m_s}) to these indices as before, i.e. for $[\![t_{n_w}]\!]^{g,h} = \langle g(n_w), h(g(n_w))\rangle$. I'll bracket for the moment any worries about whether Fregeans should accept that there are such pronouns, or that there are senses of worlds at all; I'll return to these questions at the end of the section.

Next we have a new expression, "RIG", for "rigidifier". The lexical entry for the rigidifier is as follows:

RIG

(a) $\pi_1([\![RIG]\!]^{g,h,\kappa}) = \lambda w.\lambda F.\lambda x.\lambda w'.F(x)(w)$;
(b) $\pi_2([\![RIG]\!]^{g,h,\kappa}) = \kappa$.

The first coordinate of the semantic value of the rigidifier is a function which takes a world w and an intensional property F to a rigid intensional property which, for any individual, returns true at a world w' if and only if the individual is in the extension of F at the specified world w. I'll discuss the second coordinate of the semantic value in detail in a moment. For now, I simply want to note that the second coordinate is determined by context as follows. The denotation function will from now on be understood to be relativized to a new parameter, "κ", in addition to the two assignment functions, and the second coordinate of the semantic value of the rigidifier is whatever the value of this parameter is, in the given context.

We can begin to illustrate the mechanics of the rigidifier by computing the first coordinate of the complement clause of 3 in the syntax above. Since the first coordinate of t_{4_w} relative to an assignment function g will as usual be $g(4_w)$, the first coordinate of the semantic value of $[\![RIG\ t_{4_w}\ \text{winners}]\!]^{g,h,\kappa}$ will be: $\lambda x.\lambda w'.\ x$ is a winner at $(g(4_w))$. Using a standard entry for "two", the whole complement clause will then compute to: $\lambda w'.$ there are two winners at $g(4_w)$ who lost at w'. If $g(4_w)$ is the actual world, this can be paraphrased as "two people, who in fact won, lost". This first coordinate is exactly what the Standard Solution would predict in a possible-worlds setting as the whole semantic value of the complement clause. We could have achieved this result in a

simpler way (for example, as on the Standard Solution itself), but the extra complexity we have introduced here will be needed to produce the correct second-coordinate for the complement clause. And for the Fregean, that is where the real action is.

To motivate my approach to this second coordinate, let us first consider why a flatfooted extension of our earlier system will not work here. The flatfooted extension of our account so far would be to sense-compose the senses of "two", "RIG", "t_{4_w}", "winners" and "lost", and assume that the resulting thought denotes the intensional proposition just described. This approach will not work because Fregeans will naturally say that in this situation, John does *not* stand in the belief-relation to any relevant thought which has the sense of "winner" as a constituent. John certainly does not stand in the belief-relation to a thought that *he* would express as "two people who are in fact winners lost". And more generally, since John does not think that the relevant people *are* winners, it is unclear how there could be any relevant thought which both has the sense of "winner" as a constituent, and to which John stands in this relation.

I will go beyond the flatfooted treatment by introducing a way of replacing the sense of "winner" with a sense which is a constituent of a relevant thought that John stands in the belief-relation to. In particular, I will assume that the second-coordinate of the semantic value of "RIG" is not itself a sense but instead a function from senses to senses. To make way for this possibility, I will first extend the earlier rule for Sense Composition, so that it allows for second-coordinates which are functions:

Sense Composition If α is a branching node with β and γ as daughters, then for any g:

(a) if $\pi_2([\![\gamma]\!]^{g,h,\kappa})$ is in the domain of $\pi_2([\![\beta]\!]^{g,h,\kappa})$, then $\pi_2([\![\alpha]\!]^{g,h,\kappa}) = \pi_2([\![\beta]\!]^{g,h,\kappa})([\![\gamma]\!]^{g,h,\kappa})$;

(b) if not, then if $\pi_1([\![\gamma]\!]^{g,h,\kappa})$ or $\pi_2([\![\gamma]\!]^{g,h,\kappa})$ is in the domain of $\pi_1([\![\beta]\!]^{g,h,\kappa})$, then $\pi_2([\![\alpha]\!]^{g,h,\kappa}) = \pi_2([\![\beta]\!]^{g,h,\kappa}) \oplus \pi_2([\![\gamma]\!]^{g,h,\kappa})$.

The new addition here, in clause (a), says that if it is possible (and here this will only arise in connection to κ, the second coordinate of the semantic value of "RIG") we apply one second coordinate to another. Clause (b) tells us that if (a) is not applicable, we use the original clause for Sense Composition.

This brings us at last to the second coordinate of the semantic value of "RIG", and our new parameter κ. As promised, I will assume that the κ supplied by context for 3 above is a function which takes the sense of a world to a function which replaces the sense of "winner" with a sense of the rigid property which is true of an individual at a world if and only if they are winners at the world $g(4_w)$. More formally, recalling that Δ maps senses to what those senses present, we say that a function $k : D_{m_s} \rightarrow (D_{m_{e\rightarrow p}} \rightarrow D_{m_{e\rightarrow p}})$ is a *sense-rigidifier* if and only if $\Delta((k(m_w))(m_F)) = \lambda x.\lambda w'.\Delta(m_F)(x)\Delta(m_w)$, that is, if and only if k takes a sense of a world w and a sense of an intensional property F, and returns a sense of the rigid intensional property given by the extension of F at w. Since there may be many senses of a given intensional property, there are many sense rigidifiers. We assume that every context supplies one of them.

With these assumptions in place, we can compute the second coordinate of the semantic value of the clause below the world-abstractor in the syntax for 3 presented above. We apply κ to the second coordinate of $[\![t_{4_w}]\!]^{g,h,\kappa}$ and apply the resulting

function to WINNERS, yielding a sense F, which (regardless of the value of κ in context) is guaranteed to be a sense of the rigid property which returns the set of winners at the world $g(4_w)$. We then have $\pi_2([\![\text{two RIG } t_{4_w} \text{ winners lost}]\!]^{g,h,\kappa}) = \text{TWO} \oplus \text{F} \oplus \text{LOST}$. If $g(4_w)$ is the actual world, and κ in our context returns (when applied to the salient sense of the actual world m_w, and to WINNERS), a sense corresponding to the list "Ann, Bill, Carol and Dan", then the sentence will plausibly be true. John does think precisely that two of Ann, Bill, Carol and Dan lost. At least in the case where $g(t_{4_w})$ is the actual world, our treatment promises to deliver the correct result for the clause below the world-abstraction.

This is a big step in the right direction, but we are not quite done. We have not yet said how to handle the crucial abstraction over worlds at the head of the sentence, and in fact this abstraction will turn out to pose a more difficult challenge than any we have dealt with so far.

Before discussing how this abstraction is challenging for Fregeans, let us first step back to consider why it is needed in the first place. In 3 we can plausibly produce a correct interpretation of the sentence by considering only the value of "winners" at the actual world—i.e. the set consisting of Ann, Bill, Carol and Dan. But it is easy to see that this simple way of handling such sentences does not work in general: in slightly more complex examples we must have a way of varying which set of people is chosen. For instance:

Context John believes two of Emma, Frank and Grace lost. His beliefs are true, but they were formed in an unreliable way; if Emma, Frank and Grace had all won, as could easily have happened, John still would have thought they lost. So:

13. It could have been that John thought two winners lost.

The correct reading of this sentence cannot be produced if the only available interpretations are (a) an opaque reading on which John must think to himself "two winners lost" or (b) a transparent reading on which the predicate "winners" is interpreted with its extension in the actual world. For in the relevant scenario, we are not considering a possibility where John changes his beliefs; rather we are considering the possibility that Emma, Frank and Grace all won, and in that case (where John had the same beliefs he actually has), John would have thought that two winners lost. The example motivates allowing the modal "it could have been that" to vary the world relative to which the extension of "winner" is fixed. Abstracting over worlds gives us a way of doing just this: if the world-pronoun occurring inside the attitude report is bound by an abstractor over worlds which occurs inside the scope of "it could have been that", then given that the first coordinate of the semantic value of this modal operates on the first coordinate of its prejacent, we can produce the appropriate intensional proposition as the first-coordinate of the whole sentence.

So we need this abstraction over worlds. But it might seem that my Fregean can handle it straightforwardly, by extending the clause for abstraction over individuals presented in Sect. 2.3 to the case of worlds. This is half right. The first coordinate of the abstraction can indeed be handled in this way: just as in the case of individuals, we use g to shift the value of a world-pronoun and h to ensure that the relevant sense is a sense of the shifted world. But it is only half right, because the second coordinate of

the abstraction poses a new and distinctive challenge. As discussed earlier, Fregeans will naturally hold that the *asserted content* of a sentence in context is something which hearers believe, and thus that it is a thought; in our setting it will be the second coordinate of the semantic value of the sentence. Fregeans typically hope that the second coordinate of a sentence in general will reflect the syntactic structure of the sentence; in our case, that the relevant thought reflects the fact that it was the result of a third reading. But this desideratum presents us with a serious challenge. A moment ago we revised the rule of Sense Composition to allow function application in the computation of second coordinates. These revisions were needed because we wanted the sense of "winners" to be replaced with something more appropriate in the thought which John is said to be belief-related to. The revisions were designed so that when the contextually supplied κ applies to the sense of a world-pronoun, and when the result is applied to the sense of a predicate, the sense of the predicate in question ("winners") disappears, as do the functions which are used to make it disappear. But this tactic delivers the wrong result for 3 as a whole; using the κ above it would yield:

LAMBDA-W-1$\oplus$(BELIEVES$\oplus$(TWO$\oplus$ANN, BILL, CAROL AND DAN$\oplus$LOST))$\oplus$JOHN

Here, the world-pronoun has been lost, and there is no evidence that the reading is the result of a third-reading. This is bad enough. But the same strategy applied to 13 yields even worse results. It would predict that the relevant sense would not preserve the structure of binding between the abstraction over worlds and the world-pronoun beneath the modal "it could have been that", since it would predict that the sense of the world-pronoun disappears in the thought expressed by the whole sentence.

This problem might seem inevitable. It might seem impossible to both (i) use function application in computing the second coordinate of the prejacent in 3 (to replace the sense of "winner" with something more suitable), and, at the same time, (ii) ensure that the sense of the whole sentence preserves the fact that it resulted from a transparent reading. But, as I will now show, in fact Fregeans and sententialists can manage to have their cake and eat it too. The key is in the second coordinate of a new rule for World Abstraction:

World Abstraction If α is a branching node with daughters β and γ, where β dominates only a numerical index i_w, then for any assignment functions g and h, and any κ,

(a) $\pi_1([\![\alpha]\!]^{g,h,\kappa}) = \lambda w.(\pi_1([\![\gamma]\!]^{g[w/i_w],h,\kappa})(w))$;

(b) $\pi_2([\![\alpha]\!]^{g,h,\kappa}) = \text{LAMBDA-W-I} \oplus \pi_2([\![\gamma]\!]^{g,h[\text{W-I}/g(i_w)],h(\kappa)})$.

(Here we assume world-abstraction only occurs over nodes of type p and we assume that for every i LAMBDA-W-I is a sense of type $m_{p\to p}$. We also assume that h is extended to provide senses to entities which are the type of κ.)

The final application to w in clause (a) is a technical condition which fixes a type-mismatch that would otherwise arise. This results from the special version of the Standard Solution I am using, not from any aspect of Fregeanism, so I discuss it only

in a note.[29] The key conceptual idea is in clause (b). As one would have expected from the parallel clause in Predicate Abstraction, World Abstraction alters the sense of the relevant world-pronoun to w- i: this sets the stage to allow the binding structure to be preserved in the sense of the whole sentence. But crucially it *also* alters the value of κ to a sense $h(\kappa)$. Since $h(\kappa)$ is a sense and *not a function on senses*, it no longer triggers clause (b) of Sense Composition, but instead composes with other constituents via clause (a). So this sense of κ will be preserved in the structure of the sense *of the sentence*, even though it disappears from the thought which the sentence says John is belief-related to.

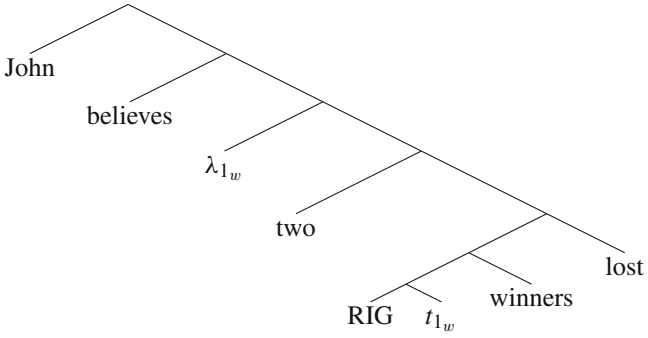
The system as a whole thus allows us to do two things which might have seemed incompatible. First, when we compute the second coordinate of the prejacent of "believes" in 3, κ disappears, and also replaces the sense of "winners" with something more suitable. This allows the first coordinate of the semantic value of the whole sentence to be an appropriate intensional proposition, which can be true in our scenario, even though John does not stand in the belief-relation to a relevant thought which has the sense of "winners" as a constituent. But, second, at the same time, the second coordinate of the semantic value of the whole sentence re-instates a sense of κ, and also has the sense of "winners" as a constituent, recording the fact that the first coordinate is the result of a transparent reading.[30]

In concluding this discussion, I want to highlight three features of the proposal. First, the system allows us to interpret not only (i) the usual opaque reading of 3 (with no world-pronouns) and (ii) the transparent reading of this sentence (with a world-pronoun bound by a binder taking scope outside of "believes" in the sentence), but also (iii) a reading where "RIG t_{1_w}" occurs above "lost" instead of above "winners", yielding a bizarre interpretation paraphrasable as "John believes two winners are the actual losers"; and (iv) an alternative syntax on which the world-pronoun is bound inside "believes", as follows:

[29] The mismatch would arise because, for the relevant γ, $\pi_1([\![\gamma]\!]^{g[w/i_w],h,\kappa})$ is an intensional proposition, so that $\lambda w.(\pi_1([\![\gamma]\!]^{g[w/i_w],h,\kappa}))$ (which one might have expected in clause (a)) is not an intensional proposition, but instead a function from worlds to intensional propositions. To produce an intensional proposition, then, we apply the intensional proposition that would be the value of this latter function at a world w, to w itself, thereby producing a function from worlds to truth-values, as desired. (For more details, see Schwarz (2012, p. 446).) If we had assumed that every expression has a world-pronoun as its argument, and altered the types of the first coordinates of expressions accordingly, this complication would not have been needed (though, as I said earlier, dealing with these worlds would have created a number of other needless distractions).

[30] In the system of this section Δ is not defined on the second coordinate of semantic values of all well-formed terms. In fact Δ is not even defined on the second coordinates of some elementary expressions, since Δ is not defined on relevant κ, the second coordinate of the semantic value of the elementary expression RIG. (κ is a function from senses to senses, but is not itself a sense, while Δ is defined only on senses.) Δ is also not defined on the result of applying κ to the sense of a world, since we have assumed that the results of these applications are functions on senses, not senses. Even so, at least in the case studied here, it is consistent to require that when Δ is defined for the second coordinate of the semantic value of an expression $\Delta(\pi_2([\![\alpha]\!]^{g,h,\kappa})) = \pi_1([\![\alpha]\!]^{g,h,\kappa})$.

Note also that the system can be characterized as compositional (though not functionally compositional) in terms of what expressions refer to (if we label what expressions refer to in the way described for the system of Sect. 2 in n. 14). Or, at least, it can be characterized in such a way that it is only as un-compositional as standard treatments of binding are (see n. 15).



The basic Standard Solution in possible-worlds semantics also predicts a reading analogous to (iii), and there has been much discussion of how it might be ruled out (Percus, 2000; Keshet, 2008; Schwarz, 2012). This problem is not a special problem for the Fregean, and the Fregean can adopt an existing strategy to handle it. But the problem with (iv) is distinctive for the Fregean. In standard possible-worlds semantics, this reading would be semantically equivalent to the opaque reading. But in our Fregean setting it will not be: it will require that John stands in the belief-relation to a possibly distinct thought which has the sense KAPPA, and the sense of a world, as constituents. What should Fregeans say about this alternative reading? There are a variety of options; I'll mention two. The simplest way to avoid this issue would be to stipulate that the world-pronoun must be bound *outside* at least one intensional operator in whose scope it occurs. This kind of stipulation is not dramatically different from various principles which have been proposed as constraints on where world-pronouns can be bound, (Percus, 2000; Keshet, 2008) and so perhaps not objectionably *ad hoc*, insofar as the Fregean of this section is simply aiming to achieve parity with the Standard Solution. Alternatively, Fregeans might hold that while not unavailable, the interpretation corresponding to the syntax above is outcompeted by the much more plausible interpretation which does not involve the locally bound world-pronoun. Alternatively still, Fregeans might try to develop a not-fully-structured conception of senses, holding that the sense of "λ_{1_w} two t_{1_w} RIG winners lost" is identical with the sense of "two winners lost". This last option might seem initially most attractive, but much more work would need to be done to see whether it is consistent in general.

Second, the proposal requires that, when a person's beliefs can truly be reported using a transparent attitude ascription, the person must stand in the belief-relation to a thought composed in part of a sense of a rigid property. For instance, I assumed that John stood in the belief-relation to a thought composed in part of the sense of the list "Ann, Bill, Carol and Dan". This assumption was natural in our setting. But one might wonder whether in *every* case where such reports are available, the ascribee can reasonably be said to stand in the belief-relation to a relevant thought. In response to this concern, I note that a parallel assumption is also built into all competing treatments I am aware of. In the possible-worlds setting we cannot speak of "constituents" of the content of individuals' beliefs, since there is not a clear notion of a "constituent" of an intensional proposition. But we can still speak of an expression's semantic value relative to an assignment, and on the possible-worlds version of the Standard Solution,

in transparent attitude reports the restrictor of a determiner will in effect be assessed as a rigid property. The same point holds for possible-worlds versions of the sophisticated movement approaches, as well as for the Fregean versions of them I develop in in Appendix B.

Third, whereas my assumption that intensional propositions are the first coordinate of the semantic values of sentences could be seen as simply a choice for the sake of concreteness in the context of Sect. 2, the theory of this section depends more heavily on this assumption. In fact, the approach of this section does not just depend on the use of intensional properties and propositions, it also appeals to world-pronouns—which refer to worlds—and even senses of worlds. Both of these facts may give pause to some Fregeans and sententialists, who may reject the idea that intensional propositions are any aspect of the meaning of sentences, and may even reject the very idea of intensional propositions. In response to these concerns, I want to observe, first, that my use of world-pronouns is less extensive than it might seem at first sight, and, second, that the way they are used seems to correspond to a general demand that any adequate theory must have a way of meeting. For the first point, the system does not predict that any senses of worlds are constituents of thoughts attributed by attitude ascriptions, or in fact even of thoughts expressed by whole sentences. When we compute the second-coordinate of the complement clause of 3 on its intended reading, κ makes the sense of the world disappear, so that the thought to which John is said to be belief-related does not itself have the sense of a world as a constituent. And the second coordinate of the semantic value of the whole sentence also does not contain the sense of a world; it contains only a special bound-world sense, which indicates that the sense of the rigidifier is to be interpreted along with the binder that occurs at the head of the sentence. These observations about how exactly world-pronouns are used pave the way for my second point: that the use of world-pronouns to record binding structure corresponds to a very general demand that everyone must have an account of. Everyone needs some way of explaining how a modal can shift the reading of distant embedded material in third readings of sentences like 13. The approach I've presented here articulates these shifts in terms of the comparatively well-understood machinery of world-pronouns, but if Fregeans have an alternative way of accounting for such shifts, then since my account only uses world-pronouns to capture these shifts, they should be able to plug it in here as well.

The proposal of this section gives Fregeans a non-movement-based account of transparent readings of determiners embedded in intensional operators. I have not argued that this approach is the best available Fregean approach. In fact, if the syntactic movement postulated by the strategies which I develop in Fregean terms in Appendix B is acceptable, then I believe that Fregeans should adopt some version of those comparatively simpler accounts. Here, I have simply shown that, even if those accounts are ruled out on syntactic grounds, this fact alone would not doom Fregeanism or sententialism. Contrary to appearances, these theories too are compatible with a non-movement-based account of third readings.

3.3 A puzzle about counterfactual attitudes

In closing this discussion I want to shift gears a little, to consider an empirical payoff of a Fregean/sententialist approach to third readings.

Consider the following case due to Blumberg (2018) (cf. Ninan (2008); Yanovich (2011); Maier (2015)):

Context It's Saturday morning, and Bill wakes up to find that the window of the back-door of his house has been smashed, with a trail of muddy footprints leading to his study. Fearing the worst, he runs to his study to check on his safe. He discovers the safe door open, and the safe emptied of its contents. His valuable collection of silverware is nowhere to be found. Given all of the evidence, Bill is quite certain that he's been burgled, and that the perpetrator acted alone. As it happens, Bill wasn't robbed. His wife removed the silverware from the safe so that it could be cleaned; a confused bird flew into the window pane and smashed it; and the muddy footprints belonged to Bill—he made them unknowingly the night before. After calling the police, Bill sits at his kitchen table with his head in his hands and says 'I wish that the person who robbed me had never robbed anyone'.

14. Bill wishes that the person who robbed him had never robbed anyone.

This sentence has a true reading in this scenario. But as Blumberg notes, this true reading cannot involve an opaque reading of "the person who robbed him": Bill does not wish that there is a robber who does not rob. In our Fregean idiom, Bill does not stand in the wish-relation to the Fregean thought expressed by ordinary uses of "the person who robbed Bill never robbed anyone".

Blumberg also argues that the reading cannot arise from a transparent reading. On this transparent reading, the determiner phrase "the person who robbed him" would be evaluated at the world of evaluation of the whole sentence, i.e. from the speaker's perspective. Since the speaker knows that no one robbed Bill, the intensional proposition expressed by "the person who (actually) robbed him had never robbed anyone" would be the proposition which is true at no worlds at all. Since Bill's wish is again not the trivial one that an arbitrary contradiction be true, this transparent reading would not be appropriate, either.

This argument against the appropriate reading of the sentence being a transparent one relies essentially on the possible-worlds framework. The argument shows that the *intensional proposition* expressed by the complement of "wish" on the transparent reading is the proposition which is true at no worlds. If "wish" is given a lexical entry which operates on such intensional propositions, then it is hard to see how the sentence would have a reasonable true reading. But in our Fregean setting "wish" will not operate on intensional propositions; it will operate on Fregean thoughts (or, in the sententialist's idiom, on sentences). And there are many Fregean thoughts (and sentences) which determine the trivially false intensional proposition. We saw this already in connection to 7 and 8: in those two sentences the complement clause of the relevant attitude reports both expressed the trivially false intensional proposition. But the senses associated with the complement clauses (and the complement clauses themselves) were different, so that the Fregean and sententialist could straightforwardly account for differences in the truth or falsity of the relevant sentences.

Merely making this observation—that the Fregean setting opens the possibility for a true transparent reading of 14—is of course not enough to count as a "solution". One needs to say at least what the Fregean thought/expression is that Bill stands in the wish-relation to. But our theory of third-readings provides an answer to this question.[31] On a transparent reading of the determiner phrase with respect to the attitude verb "wish", the relevant Fregean thought will be composed of the sense of "the", a sense of the rigidification of the property expressed by "person who robbed Bill", and the sense of "never robbed anyone". Provided the salient sense of the relevant rigid property is something like "person who in fact robbed Bill", Bill's wish will not be contradictory: he will wish, precisely, that the person who in fact robbed him never robbed anyone. There is nothing contradictory in wishing that a person who in fact did something had done something different. If Bill had been robbed, he would have wished exactly this. Of course, because Bill was not robbed there is no person who in fact robbed him, and the intensional proposition corresponding to this Fregean thought will be the trivially false one. But *that does not make his wish trivial*, since his wish is a relation to a Fregean thought, not to an intensional proposition.

The Fregean and the sententialist thus have a natural and intuitive account of the object of Bill's wish in this case. But their theory does have an important lacuna. Fregeans and sententialists owe us a fuller theory of how κ is determined by features of context so that it delivers the intuitively natural object of Bill's wish. This question is an urgent one for Fregeans and sententialists. But it is a question that arises for the theory independently of its approach to the data of this section, since it arises for the treatment of third readings in the previous section as well. This observation underscores a key point: whereas standard possible-worlds-based theories (as the cited authors have argued) cannot handle 14 using machinery already in place to handle 3, Fregeans and sententialists can. Unlike the possible-worlds-based theories, then, Fregeanism and sententialism have some hope of giving a unified treatment of third readings and the general class of examples like 14. If they are able to offer such a unified treatment in general, this would seem to give them a theoretical advantage over their competitor theories.

At present, this advantage is merely a hope for Fregeans and sententialists. It may be that Fregeans and sententialists cannot handle all data which are prima facie similar to 14 by the mechanism I have just described. But the fact that, as I have shown, the tools developed for third readings deliver this result even for 14 itself is interesting, and I hope it will spur further investigation of Fregean and sententialist treatments of this kind.

4 Conclusion

In the first half of this paper, I developed a new Fregean/sententialist semantics for attitude reports. My semantics builds on an idea from Bigelow (1978), but presenting this thought in the format of Heim and Kratzer (1998) requires some new ideas. The

[31] This answer does not depend on the specific implementation of the previous subsection, as opposed to the treatments presented in Appendix B, but it does depend on the details of Fregeanism specifically.

resulting account improves in several ways on the prominent recent Fregean treatment of Yalcin (2015), perhaps most notably by allowing the Fregean to treat the *de se* in parallel with the *de re*, as had been suggested by Stanley (2011). Appendix A shows how the account can also be extended to an array of further challenging data.

The second half of the paper considered how Fregeans might give an account of the "third readings" discovered by Fodor (1970). It is fairly clear that they can give an account of such data using sophisticated theories of movement (such accounts are developed in detail in Appendix B). But if these approaches were the only ones available to Fregeans, Fregeanism would depend on controversial claims about syntactic movement, and the theory would thus be "hostage to fortune" in an arguably problematic way. In Sect. 3.2, I showed that Fregeanism is not hostage to fortune in this way, by developing a Fregean version of the "Standard Solution" from possible-worlds semantics. This theory not only brings Fregeanism and sententialism up to date as competitors to more standard theories, but also has immediate payoffs. Whereas possible-worlds-based accounts must treat certain data about counterfactual attitudes using machinery that goes beyond what is required to handle third readings, Fregeanism offers some hope for handling these two apparently disparate classes of reports by the same mechanism.

Fregean and sententialist approaches to the semantics of attitude reports have largely fallen out of the mainstream in formal semantics in recent years. But these alternatives to mainstream possible-worlds semantics offer a distinctive perspective on data which have proven difficult to accommodate within more mainstream theories. With more systematic investigation, they may offer us not just a fresh perspective, but empirically more adequate and theoretically more satisfying accounts.

Appendix A: Extending the basic system

In this section I present an extension of the basic theory from Sect. 2.

I will motivate this extension by a series of examples, starting from the "Paderewski" example of Kripke (1979) (cf. Schiffer (1979) on "Thelma"). In Kripke's example, Peter knows that Paderewski is a pianist, and that Paderewski is prime minister of Poland, but thinks that Paderewski the pianist is not a politician. Kripke observes that the following can both be used truly to describe Peter:

15. Peter believes that Paderewski is prime minister;
16. Peter does not believe that Paderewski is a prime minister.

The basic theory of Sect. 2 cannot account in a natural way for the fact that both of these sentences have true readings. One could extend the theory to do this if one says that the name "Paderewski" is context-sensitive and associated with different senses in different contexts, or that the name "Paderewski" as used to refer to this particular Paderewski is ambiguous, and that different disambiguations are associated with different senses.[32] Here I will pursue a different approach, on which the verb

[32] Since sententialists take the second coordinate of expressions to be the expressions themselves, they can't say that names are associated with different expressions in different contexts, if they are the very

"believe" is context-sensitive and can (in effect) express different relations on different occasions of use.

To achieve this result, we slightly alter the lexical entry for the verb "believe", in a manner inspired by Carnap (1947), Davidson (1968), Kaplan (1989, §XX) and Richard (1990) among many others. The word "believe" will now denote not the BEL relation itself, but rather the relation which holds between a person and a thought m when the person stands in the BEL relation to *some thought which is relevantly related* to m. Formally, a *transformation* t is a function from D_{m_p} to D_{m_p} such that $\Delta(t(m)) = \Delta(m)$, that is, a function which maps every thought to a thought which determines the same intensional proposition. We assume that context supplies a function f from individuals to sets of transformations which are salient relative to those individuals:

Extended Believe $\pi_1([\![\text{believe}]\!]^{g,h,f}) = \lambda p.\lambda x.$ for some t such that $t \in f(x)$ $x\ BEL\ t(p)$.

(We omit κ here, since the present extension is thought of as directly extending the account of Sect. 2.)

This extension of the basic system allows us to give a straightforward account of the fact that both 15 and 16 have true readings. The first will be true in contexts where some transformation which is salient relative to Peter maps the thought expressed by "Paderewski is a prime minister" to one that Peter is belief-related to, and the second will be true in contexts where all of the salient transformations for Peter map this thought to one that Peter is not belief-related to. This result depends crucially on the fact that the lexical entry introduces existential quantification only over the transformations which are salient in context, and that which transformations are salient can differ in different contexts.[33]

This flexibility raises some questions about the predictiveness of the approach, which I'll return to at the end of this appendix. But before I do that, I want to consider how our new lexical entry handles three other kinds of examples which pose problems for the basic account presented in the main text. The first is from Quine (1956) and has come to be known as an example of "double vision". Here is a version of Quine's case:

Context Ralph sees Ortcutt by the docks. Ralph concludes on the basis of what he sees that Ortcutt is a spy. Later, Ralph watches Ortcutt's mayoral inauguration address

Footnote 32 continued
same name. So a sententialist version of this approach would be to say that there are two different names "Paderewski" (see, e.g. Fiengo and May (2006)).

[33] In the main text, I have presented the formalism on the assumption that h will change from context to context. But the use of transformations opens up a different possibility. A Fregean could hold instead that for each individual, there is a distinguished "*de re*" sense, and that context always supplies this sense as the sense of the relevant variable, so that h would not change from context to context. There are various ways one could spell out this view further, by adding hypotheses about the metaphysics of mind. For instance, one might hold that people who stand in the belief-relation to particular thoughts featuring particular "acquaintance-based" senses, automatically stand in the belief-relation to relevant thoughts composed of the *de re* sense. (For discussion of related ideas, with citations, see Yalcin (2015, §7).) On a natural version of this view, if the sense of "Venus" is an "acquaintance-based" sense, anyone who stands in the belief-relation to the thought expressed by "Venus is bright" thereby automatically stands in the belief-relation to the thought composed of the *de re* sense of Venus and the sense of "is bright".

on TV. Ralph thinks that no mayor could possibly be a spy; the background checks are simply too rigorous. So he concludes that Ortcutt the mayor is not a spy.

17. Ralph believes that Ortcutt is a spy.
18. Ralph believes that Ortcutt is not a spy.

On the natural assumption that in this case for any sentence such that Ralph is belief-related to the thought expressed by that sentence, he is not belief-related to the thought expressed by the negation of the sentence, the Fregean theory from Sect. 2 cannot predict that these two sentences are true in the same context. But it might seem that in fact these two sentences can be true in the same context.[34]

Extended Believe allows us to give a true reading of these two sentences in a single context, while upholding the assumption that for any sentence such that Ralph is belief-related to the thought expressed by that sentence, he is not belief-related to the thought expressed by the negation of the sentence. The sentences can be true in a context where a transformation which maps the thought expressed by "Ortcutt is a spy" to something like that expressed by "The actual person Ralph saw by the docks is a spy" is salient relative to Ralph, and in addition a transformation which maps the thought expressed by "Ortcutt is not a spy" to something like that expressed by "The actual mayor of Ralph's town is not a spy" (which may or may not be the same transformation) is also salient relative to Ralph.

The second kind of example, from Stephen Schiffer, gives rise to what is sometimes called the "Madonna problem":

> Consider
>
> 19. Everyone who has ever known her has believed that Madonna was musical.
>
> According to the Fregean proposal, there is a particular mode of presentation m of Madonna and a particular mode of presentation m' of the property of being musical such that the foregoing utterance of 19 is true only if everyone who has ever known Madonna has believed the proposition $\langle m, m' \rangle$. Yet this is surely too strong a requirement on the truth of 19. It requires that everyone who has ever known Madonna shared a single way of thinking of her and a single way of thinking of the property of being musical, and this is most unlikely given that there may have been people who knew her as a child and then died and that someone like Helen Keller may have been among them." (Schiffer, 1992, pp. 507–508; numbering of the example altered to fit the present paper)

If cognitively diverse people cannot stand in the belief-relation to the thought expressed by "Madonna is musical", then the Fregean theory of Sect. 2 cannot predict the true reading of this sentence.

But Extended Believe allows us to handle this report, while preserving the assumption that very cognitively diverse people cannot stand in the belief-relation to the thought expressed by "Madonna is musical". We simply assume that the various transformations made salient by the story above map the thought expressed by "Madonna is

[34] One might argue that they are not true in the same context, for instance by suggesting they are "revisionist reports" Blumberg and Lederman (2020). I discuss this style of response and more refined arguments against it in detail in Lederman (2021, §5.1 and 7.1).

musical" to a wide array of different Fregean thoughts. If each relevant person stands in the belief-relation to one of these thoughts, Extended Believe can allow that the the sentence is true.

A third kind of example was to my knowledge originally developed by Soames (1990, p. 198f.) (cf. Higginbotham (1991, p. 362 ex. 42) and, more extensively, Soames (1994)), but has recently been brought back into the spotlight by Sharvit (2010) and Charlow and Sharvit (2014):

Context John knows that Jupiter is bigger than Mars, and that Mars orbits the sun faster than Jupiter. He believes no planet is bigger than Jupiter, and no two planets are exactly the same size. He thinks that Hesperus is Jupiter and thinks that Phosphorus is Mars.

20. There's a planet which John thinks is as big as Jupiter and orbits the sun as fast as Mars.

Intuitively, this sentence has a true reading. But this true reading is hard to account for on the theory of Sect. 2: since the trace which is the subject of "is as big as Jupiter" and the one which is the subject of "orbits the sun as fast as Mars" are coreferential, in the system of that section they will be forced to be associated with the same sense.

Extended Believe, however, once again allows a straightforward way of handling this datum. In contexts where a transformation which is salient relative to John maps the thought expressed by the complement clause on an assignment to the thought expressed by "Hesperus is as big as Jupiter and Phosphorus orbits the sun as fast as Mars", the sentence will have a true reading.

As I mentioned earlier, an important objection to this way of extending the account is that it is too unconstrained. Transformations themselves are very flexible, since they can map a thought to any other thought which determines the same intensional proposition.[35] Moreover, I have not said anything about how a given backstory or environment might make some transformations salient as opposed to others. This latter question is especially urgent since, as I emphasized in connection to 15 and 16, the theory requires that different transformations be salient in different contexts.

One way of strengthening the account would be to impose a further constraint on which transformations can be supplied by context. Say that a transformation t is a *translation* if and only if (i) for any elementary sense (i.e. sense of an elementary expression) m, $\Delta(t(m)) = \Delta(m)$ and (ii) $t(f \oplus x) = t(f) \oplus t(x)$. The idea would then be that context supplies a more restrictive relation between people and *translations* in context, not merely between people and transformations.[36]

Constraints like this one are worth exploring further as a way of winnowing down the readings our Fregean allows. But I am not sure that this constraint is the correct

[35] This objection may be particularly concerning for Fregeans who take the first-coordinate of the semantic values of sentences to be truth-values (as opposed to intensional propositions), since in that case the fact that the relevant mappings must be transformation functions does very little to constrain which readings are available.

[36] This requirement is, in effect imposed by Richard (1990). Richard does not give an explicit treatment of abstraction, making it hard to compare his theory with mine directly. But in addition to the fact that Richard imposes this extra requirement on transformations, there is one important difference between our two theories, namely, that Richard's lexical entry has "believe" operate (essentially) on *pairs* of first-coordinates and second-coordinates.

one. One reason for doubt is that it is incompatible with our treatment of 20 above. The transformation I appealed to there must not be a translation; it must map the two "occurrences" of the sense associated with the planet Venus to different senses (in one case, a sense of Mars, and in the other a sense of Jupiter). There are other tools which might be used to handle that example, but in my view the most promising way of doing so introduces other dimensions of flexibility (and hence lack of predictive power) which are as worrying as the kind of flexibility this new constraint was designed to avoid.[37]

Other than to say that this constraint does not seem to me the right one, I do not have a general answer to the question of how Fregeans should make their theory more predictive. I do, however, believe that the fact that we do not yet have a predictive account of which transformations are salient in context should not bar us from exploring this theory further. On the contrary, investigating how Fregeans might further constrain the theory presented here seems to me an important line of future research.

Appendix B: Sophisticated movement

The goal of this appendix is to develop Fregean accounts of third readings which rely on sophisticated forms of movement.[38]

On the simplest of these theories—which will be the main focus of the appendix—the predicate "winner" is assumed to be able to move out of the scope of "John believes", and the system is altered in such a way that if this expression moves it denotes not the property of being a winner, but the rigid property of belonging to the set or plurality of the actual winners (or, at least, the winners in the relevant world of evaluation). Given these assumptions, our sentence would have the (rough) paraphrase "*being an actual winner* is an F such that John believes two Fs lost".

Footnote 36 continued

Chalmers (2011, p. 618) similarly gives a related account of simple attitude ascriptions. But again, when it comes to quantifying in (§8) he gives some sentence-level paraphrases without a systematic treatment, making it hard to situate his account with respect to how it handles abstraction.

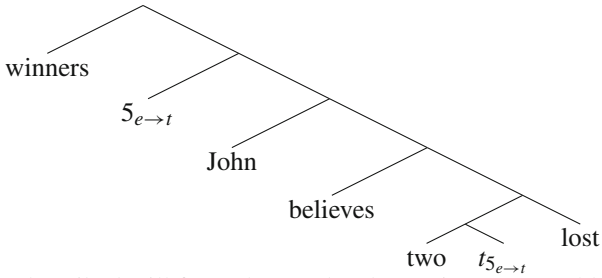
[37] Here are two alternative treatments. First, one could hold that there *are* senses s such that John stands in the think-relation to the thought produced by substituting s in for the senses of "Hesperus" and "Phosphorus" in the thought expressed by "Hesperus is as big as Jupiter and Phosphorus orbits the sun as fast as Mars", maybe the distinctive *de re* sense described in n. 33. Second, following Santorio (2014) and Barker (2016), one could postulate that variables have not one but two indices. We let h be not just a function from individuals to senses, but a function from individuals to functions from indices to senses, so that for all traces $t_{i,j}$, and all g, h, we assume that $[\![t_{i,j}]\!]^{g,h} = \langle g(i), h(g(i))(j)\rangle$ and for all i and j, $\Delta(h(g(i)(j))) = g(i)$. We then say that the two occurrences of traces bound under "thinks" in 20 have distinct second indices. This move allows us to say that, even though they are assigned to the same individual (because their first index is the same, and this is all that g is sensitive to), h can still assign them distinct senses. This approach would require some extra assumptions to ensure that changes in the labeling of bound variables would not make a difference to the semantic value of the sentence.

[38] A broadly "sophisticated movement approach" postulates "split intensionality" (Keshet, 2008, 2010, 2011). On this approach, a third scope is postulated for attitude verbs such that if a determiner moves to this scope, its quantificational force will be interpreted under the attitude verb, while its attitudinal status is fixed by the world of evaluation of the attitude verb. It seems to me an interesting question whether Fregeans can develop a version of this approach as well, but I will not undertake to answer this question here.

Formally, we develop this idea as follows. We use a slightly different semantics than the one introduced in Sect. 2; in particular we now use a world-indexed denotation function $[\![\cdot]\!]^{w,g,h}$. Many expressions (e.g. names, e-type variables) have the same world-indexed denotation as the non-world indexed denotation they had in the old system. But expressions whose semantic values had world-sensitive first coordinates, for instance, predicates, are now handled somewhat differently. Predicates are still thought of as assigned functions from individuals to worlds to truth-values by the lexicon (think of this as the first coordinate of a "pre-semantic value"), but for a predicate α which is assigned such a function f in the lexicon, $\pi_1([\![\alpha]\!]^{w,g,h}) = \lambda x.f(x)(w)$ for every g, h, and w. In words, the first coordinate of the world-indexed semantic value of such an expression is given by saturating the second argument of the function assigned it by the lexicon with its world-index, yielding instead a function from individuals to truth-values. Higher-type expressions (e.g. determiners) have lexical entries altered in the obvious way to reflect the fact that their arguments are now extensions, as opposed to intensional properties. When the meaning of an expression (such as a modal operator, or an attitude verb) requires an intensional property as its argument, the type-mismatch between it and its argument is repaired by the introduction of an abstraction over the world index where it is needed (as in the "Intensional Function Application" of Heim and Kratzer (1998, p. 308)). To produce an intensional proposition as the asserted content of an utterance of a sentence, we also require that abstraction over the world-index occurs at the head of the sentence.

We still assume that Δ maps the senses of predicates not to extensions, but to intensional properties. In particular, we assume that the sense of a predicate is mapped by Δ to the intensional property which it is assigned by the lexicon (what I earlier called its "pre-semantic value"). In general this means that while Intensional Function Application is needed to repair type mismatches in the first coordinates of the semantic values of attitude reports, no corresponding special rule is needed for second coordinates; we assume that the abstraction over world-indices we have been discussing has no effect on the second coordinates of sentences.

Our target syntax for the intended reading of 3 will be:



The rules just described will force there to be abstraction over worlds both immediately beneath "believes", and at the head of the sentence, but these are not marked syntactically. The abstraction below "winners" (marked by $5_{e\to t}$) has type $e \to t$—where t is the type of truth values—as opposed to $e \to p$, because in our new system the first coordinates of the semantic values of predicates are extensions, not intensional properties. This fact will be crucial to achieving our desired result: if the first coordinate of the semantic value of "winner" were an intensional property (as in the

original system), and if the binder were over a trace of type $e \rightarrow p$, then the first coordinate of the semantic value of the syntax above would be the same as that of the syntax where no movement had occurred at all. In the system of this appendix, by contrast, "winners" supplies an *extension* for the trace inside the scope of "believes", and this extension is not shifted by the world-abstraction beneath "believes", allowing a different interpretation of the sentence on the syntax where the predicate has moved.

One might have hoped that we could simply re-use the clause for abstraction over individuals here for abstraction over extensions. But since we want h to produce not a sense of an extension, but a sense of an intensional property for the trace $5_{e\rightarrow t}$, we need to modify the clause for the first coordinate slightly:

Property Predicate Abstraction If α is a branching node with daughters β and γ, where β dominates only a numerical index $i_{e\rightarrow t}$ of type $e \rightarrow t$, then for any assignment functions g and h,

(a) $\pi_1([\![\alpha]\!]^{w,g,h}) = \lambda F.\pi_1([\![\gamma]\!]^{w,g[F/i],h[h(\lambda w.\lambda x.F(x)(w))/F]})$;
(b) $\pi_2([\![\alpha]\!]^{w,g,h}) = \text{LAMBDA- I}_{e\rightarrow p} \oplus \pi_2([\![\gamma]\!]^{g,h[\text{I}_{e\rightarrow p}/g(i_{e\rightarrow t})]})$.

The key new features are in clause (a). We simply set the value of h on the relevant extension to its value on the rigid property which returns that extension at every world. In the second clause, we assume that the sense LAMBDA- I and the corresponding trace-sense I have type $e \rightarrow p$ because, as discussed above, we still assume that senses have the types used for them in Sect. 2. This clause produces the truth-conditions described informally in the main text.

Von Fintel and Heim (2002, 8.3.2 Way 1) also suggest a different sophisticated movement approach, which is in some ways more attractive than the one just described. On this approach, the whole determiner phrase "two winners" is assumed to move, as it does in the simple movement strategy. But unlike in the simple movement strategy, it is assumed that the phrase can move leaving behind a trace of its own type, i.e. $((e \rightarrow t) \rightarrow t)$, not of type e. As in the predicate movement approach just described, this approach is implemented in a setting with a world-indexed denotation function, designed so that the first coordinate of the semantic value of the higher-order trace is not the world-variable denotation of the original determiner phrase, but instead a rigid "extension" of the type of the determiner phrase. The resulting syntax is thus interpreted as roughly "*Two actual winners* is a Q such that John believes Q lost." This "paraphrase" is hard to interpret directly, but the effect is to closely approximate the truth-conditions of the predicate movement approach. Since the whole Q is reinterpreted underneath "John believes", the truth-conditions do not require that there be two particular winners which John believes lost; as we might put it, the quantificational force of "two winners" is interpreted below "believes". But, since the expression has undergone movement, the predicate "winners" is interpreted outside the scope of "believes", and thus interpreted from the speaker's perspective, not from the attitude holder's.

We can implement this strategy straightforwardly by modifying the "Predicate Abstraction" rule just given to apply to the type of determiner-phrases.

References

Abreu Zavaleta, M. (2019). Weak speech reports. *Philosophical Studies, 176*(8), 2139–2166. https://doi.org/10.1007/s11098-018-1119-2.

Anand, P. (2006). *De De Se*. Ph.D. thesis, MIT.

Barker, C. (2016). The linguistics of coordinated thought. Ms., New York University.

Baron, C. (2015). Generalized concept generators. *Proceedings of NELS*, *46*, 59–68.

Bäuerle, R. (1983). Pragmatisch-Semantische Aspekte der NP-Interpretation. In M. Faust, R. Harweg, W. Lehfeldt, & G. Wienold (Eds.), *Allgemeine Sprachwissenschaft, Sprachtypologie und Textlinguistik* (pp. 121–131). Tübingen: Narr.

Bigelow, J. C. (1978). Believing in semantics. *Linguistics and Philosophy, 2*(1), 101–144. https://doi.org/10.1007/BF00365131.

Blumberg, K. (2018). Counterfactual attitudes and the relational analysis. *Mind, 127*(506), 521–546. https://doi.org/10.1093/mind/fzx007.

Blumberg, K., & Lederman, H. (2020). Revisionist reporting. *Philosophical Studies*. https://doi.org/10.1007/s11098-020-01457-4.

Carnap, R. (1947). *Meaning and necessity: A study in semantics and modal logic*. Chicago, IL: University of Chicago Press.

Chalmers, D. J. (2011). Propositions and attitude ascriptions: A Fregean account. *Noûs, 45*(4), 595–639.

Charlow, S., & Sharvit, Y. (2014). Bound 'de re' pronouns and the LFs of attitude reports. *Semantics and Pragmatics, 7*(3), 1–43. https://doi.org/10.3765/sp.7.3.

Church, A. (1946). A formulation of the logic of sense and denotation (abstract). *Journal of Symbolic Logic, 11*, 31.

Church, A. (1951). A formulation of the logic of sense and denotation. In P. Henle, H. Kallen, & S. Langer (Eds.), *Structure, method, and meaning: Essays in honor of Henry M. Sheffer* (pp. 3–24). New York: Liberal Arts Press.

Church, A. (1973). Outline of a revised formulation of the logic of sense and denotation (part i). *Noûs, 7*(1), 24–33.

Church, A. (1974). Outline of a revised formulation of the logic of sense and denotation (part ii). *Noûs, 8*(2), 135–156.

Church, A. (1993). A revised formulation of the logic of sense and denotation. Alternative (1). *Noûs, 27*(2), 141–157.

Cresswell, M. J. (1975). Hyperintensional logic. *Studia Logica, 34*(1), 25–38. https://doi.org/10.1007/BF02314421.

Cresswell, M. J. (1985). *Structured meanings*. Cambridge, MA: MIT Press.

Cresswell, M. J., & von Stechow, A. (1982). De re belief generalized. *Linguistics and Philosophy, 5*(4), 503–535. https://doi.org/10.1007/BF00355585.

Crimmins, M., & Perry, J. (1989). The prince and the phone booth: Reporting puzzling beliefs. *The Journal of Philosophy, 86*(12), 685–711. https://doi.org/10.2307/2027014.

Davidson, D. (1968). On saying that. *Synthese, 19*(1/2), 130–146.

Dorr, C. (2014). Transparency and the context-sensitivity of attitude reports. In M. Garcia-Carpintero & G. Martí (Eds.), *Empty representations: Reference and non-existence* (pp. 25–66). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199647057.003.0002.

Dummett, M. (1981). *Frege: Philosophy of language*. Cambridge, MA: Harvard University Press.

Field, H. (1978). Mental representation. *Erkenntnis, 13*(July), 9–61. https://doi.org/10.1007/BF00160888.

Field, H. (2017). Egocentric content. *Noûs, 51*(3), 521–546. https://doi.org/10.1111/nous.12141.

Fiengo, R., & May, R. (2006). *De Lingua Belief*. Cambridge, MA: Bradford Book/MIT Press.

von Fintel, K., & Heim, I. (2002). Lecture notes on intentional semantics. Ms., MIT.

Fodor, J. D. (1970). *The linguistic description of opaque contents.* Ph.D. thesis, MIT.

Frege, G. (1948). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100, 25–50 (1892). https://doi.org/10.2307/2181485. Translated as "On Sense and Reference" by Max Black in *Philosophical Review*, 57(3), 209–230.

Goodman, J., & Lederman, H. (2021). Perspectivism. *Noûs*, *55*(3), 623–648.

Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Oxford: Blackwell.

Higginbotham, J. (1991). Belief and logical form. *Mind and Language, 6*(4), 344–369. https://doi.org/10.1111/j.1468-0017.1991.tb00261.x.

Hintikka, J. (1962). *Knowledge and Belief.* Ithaca, NY: Cornell University Press.

Janssen, T. M. V., & Partee, B. H. (1997). Compositionality. In J. van Bentham & A. ter Meulen (Eds.), *Handbook of logic and language* (pp. 417–473). Amsterdam: Elsevier.
Kaplan, D. (1968). *Quantifying in. Synthese, 19*, 178–214. https://doi.org/10.1007/BF00568057.
Kaplan, D. (1986). Opacity. In L. E. Hahn & P. A. Schilpp (Eds.), *The philosophy of W.V. Quine* (pp. 229–289). La Salle, IL: Open Court.
Kaplan, D. (1989). Demonstratives. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 481–563). Oxford: Oxford University Press.
Keshet, E. (2008). *Good intensions: Paving two roads to a theory of the de re/de dicto distinction*. Ph.D. thesis, MIT.
Keshet, E. (2010). Possible worlds and wide scope indefinites: A reply to Bäuerle 1983. *Linguistic Inquiry, 41*(4), 692–701. https://doi.org/10.1162/LING_a_00019.
Keshet, E. (2011). Split intensionality: A new scope theory of de re and de dicto. *Linguistics and Philosophy, 33*, 251–283. https://doi.org/10.1007/s10988-011-9081-x.
Keshet, E., & Schwarz, F. (2019). De re/de dicto. In J. Gundel & B. Abbott (Eds.), *The Oxford handbook of reference* (pp. 167–202). Oxford: Oxford University Press.
King, J. C. (2007). *The nature and structure of content*. Oxford: Oxford University Press.
Klement, K. C. (2001). Russell's paradox in Appendix B of 'The priciples of mathematics': Was Frege's response adequate? *History and Philosophy of Logic, 22*(1), 13–28.
Klement, K. C. (2002). *Frege and the logic of sense and reference*. New York: Routledge.
Kripke, S. A. (1979). A puzzle about belief. In A. Margalit (Ed.), *Meaning and use* (pp. 239–283). Dordrecht: Reidel. https://doi.org/10.1007/978-1-4020-4104-4_20.
Kusliy, P., & Vostrikova, E. (2019). Hard cases of third readings in terms of the standard solution. *Proceedings of Sinn und Bedeutung, 22*(2), 37–52. https://doi.org/10.18148/sub/2018.v22i2.69.
Lederman, H. (2021). Fine-grained semantics for attitude reports. *Semantics and Pragmatics, 14*(1), 1–33. https://doi.org/10.3765/sp.14.1.
Lewis, D. K. (1970). General semantics. *Synthese, 22*(1–2), 18–67. https://doi.org/10.1007/BF00413598.
Maier, E. (2015). Parasitic attitudes. *Linguistics and Philosophy, 38*(3), 205–236. https://doi.org/10.1007/s10988-015-9174-z.
Mates, B. (1952). Synonymity. In L. Linsky (Ed.), *Semantics and the philosophy of language* (pp. 111–136). Urbana, IL: University of Illinois Press.
Ninan, D. (2008). *Imagination, content, and the self*. Ph.D. thesis, MIT.
Ninan, D. (2012). Counterfactual attitudes and multi-centered worlds. *Semantics and Pragmatics, 5*(5), 1–57. https://doi.org/10.3765/sp.5.5.
Partee, B. H. (1973). The semantics of belief-sentences. In K. J. J. Hintikka, J. M. E. Moravcsik, P. Suppes (Eds.), *Approaches to natural language: Proceedings of the 1970 Stanford workshop on grammar and semantics* (pp. 309–336). Dordrecht: Springer. https://doi.org/10.1007/978-94-010-2506-5_18.
Pavese, C. (2017). Know-how and gradability. *Philosophical Review, 126*(3), 345–383.
Percus, O. (2000). Constraints on some other variables in syntax. *Natural Language Semantics, 8*(3), 173–229. https://doi.org/10.1023/A:1011298526791.
Percus, O., & Sauerland, U. (2003). On the LFs of attitude reports. *Proceedings of Sinn und Bedeutung,* 7, 228–242. https://doi.org/10.18148/sub/2003.v7i0.805.
Pickel, B. (2017). Structured propositions in a generative grammar. *Mind, 128*(510), 329–366. https://doi.org/10.1093/mind/fzw074.
Pickel, B. (2021). The functional composition of sense. *Synthese, 199*, 6917–6942.
Quine, W. V. (1940/1981). *Mathematical logic*. Cambridge, MA: Harvard University Press.
Quine, W. V. (1956). Quantifiers and propositional attitudes. *The Journal of Philosophy, 53*(5), 177–187. https://doi.org/10.2307/2022451.
Rabern, B. (2012). Against the identification of assertoric content with compositional value. *Synthese, 189*(1), 75–96. https://doi.org/10.1007/s11229-012-0096-9.
Richard, M. (1990). *Propositional attitudes: An essay on thoughts and how we ascribe them*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511625282.
Rieppel, M. (2017). Names, masks, and double vision. *Ergo, 4*(8), 229–257.
Russell, B. (1903). *The Principles of mathematics*. Cambridge: Cambridge University Press.
Russell, B. (1905). On denoting. *Mind, 14*, 479–493. https://doi.org/10.1093/mind/XIV.4.479.
Santorio, P. (2014). On the plurality of indices. Ms., University of Leeds.

Schiffer, S. (1979). Naming and knowing. In H. K. Wettstein, A. French Peter, & T. E. Uehling (Eds.), *Midwest Studies in Philosophy II: Contemporary Perspectives in the Philosophy of Language* (pp. 28–41). Minneapolis, MN: University of Minesota Press.

Schiffer, S. (1992). Belief ascription. *The Journal of Philosophy, 89*(10), 499–521. https://doi.org/10.2307/2941005.

Schwager, M. (2009). Speaking of qualities. *Semantics and Linguistic Theory, 19*, 395–412.

Schwarz, F. (2012). Situation pronouns in determiner phrases. *Natural Language Semantics, 20*(4), 431–475. https://doi.org/10.1007/s11050-012-9086-1.

Sharvit, Y. (2010). Covaluation and unexpected bt effects. *Journal of Semantics, 28*(1), 55–106. https://doi.org/10.1093/jos/ffq012.

Soames, S. (1990). Pronouns and propositional attitudes. *Proceedings of the Aristotelian Society, New Series, 90*, 191–212. https://doi.org/10.1093/aristotelian/90.1.191.

Soames, S. (1994). Attitudes and anaphora. *Philosophical Perspectives, 8*, 251–272. https://doi.org/10.2307/2214173.

Soria Ruiz, A. (forthcoming). On mates' puzzle. *Mind & Language*.

Stanley, J. (2011). *Know how*. Oxford: Oxford University Press.

Tancredi, C., & Sharvit, Y. (2020). Qualities and translations. *Linguistics and Philosophy, 43*(3), 303–343. https://doi.org/10.1007/s10988-019-09277-4.

Von Stechow, A. (1985). Review of Cresswell, 'Structured meanings'. *Journal of Semantics, 4*(2), 165–191. https://doi.org/10.1093/jos/4.2.165.

Yalcin, S. (2015). Quantifying in from a Fregean perspective. *Philosophical Review, 124*(2), 207–253. https://doi.org/10.1215/00318108-2842186.

Yanovich, I. (2011). The problem of counterfactual de re attitudes. *Semantics and Linguistic Theory, 21*, 56–75.

Yli-Vakkuri, J. (2013). Propositions and compositionality. *Philosophical Perspectives, 27*(1), 526–563. https://doi.org/10.1111/phpe.12025.

Zimmermann, T. E. (2018). Fregean compositionality. In D. Ball & B. Rabern (Eds.), *The science of meaning* (pp. 276–305). Oxford: Oxford University Press.