Synthese (2022) 200:433
https://doi.org/10.1007/s11229-022-03915-4

ORIGINAL RESEARCH

Mates and the hierarchy

Gurpreet Rattan[1] · **Marion Durand**[2]

Received: 11 March 2022 / Accepted: 27 September 2022 / Published online: 20 October 2022

Abstract

Mates's Puzzle has flown below many philosophers' radar, despite its relations to both Frege's Puzzle and the Paradox of Analysis. We explain the relations amongst these puzzles on the way to arguing that Mates's Puzzle suggests a generalization of Frege's Puzzle, and of the sense-reference distinction itself, in the form of hierarchy of senses. We explain how Mates's Puzzle and the hierarchy, to different degrees, illuminate each other, and how their connection is missed in the literature. However, we argue that the potential of Mates's Puzzle to illuminate the hierarchy is yet to be fully actualised. We suggest that in order to do better, we need to formulate Mates's Puzzle as a puzzle in the philosophy of thought and not as puzzle in the philosophy of language. What is needed, to which the present paper is a precursor, is an account of *the cognitive significance of higher-order senses*.

Keywords Mates's Puzzle · Frege's Puzzle · The Paradox of Analysis · Fregean hierarchy · Cognitive significance

Every philosopher knows about Frege's Puzzle. Lots of philosophers know about the Paradox of Analysis. Mates's Puzzle, by contrast, flies below many philosophers' radar. This is so despite the fact that the three puzzles are all members of the same family of puzzles about foundational matters in the semantics of thought and language. Mates's Puzzle is also the least well understood of these puzzles. This is a little surprising because although Mates's Puzzle is not well-known, it attracted the attention of some prominent twentieth century philosophers of language, including

✉ Marion Durand
marion.durand@philosophy.ox.ac.uk

Gurpreet Rattan
gurpreet.rattan@utoronto.ca

1 University of Toronto, Toronto, Canada

2 Corpus Christi College, University of Oxford, Oxford, UK

Alonzo Church (1954), Hilary Putnam (1954), Israel Scheffler (1955), Wilfrid Sellars (1955), and, later, Tyler Burge (1978) and Saul Kripke (1979),[1] after appearing as a passing remark in Benson Mates's 'Synonymity'(1950). Its relatively neglected status is no doubt partly due to the formulation of the puzzle suggested by Mates and subsequently substantively filled out by others. The formulation is difficult to keep in mind, employing as it does logically complex sentences occurring in multiply embedded propositional attitude attributions.[2] Nevertheless, it was recognized early on that Mates's Puzzle potentially contains deep challenges for the most foundational issues in the philosophy of language, including issues about synonymy, compositionality, and direct versus indirect discourse. Although we think that the lessons for these foundational issues in the philosophy of language are well worth sorting out, and although we do engage with them to some degree here, our larger and more ultimate interest, which extends past the current paper, lies elsewhere: in the philosophy of thought, and in particular as it derives from Frege and concerns the notion of *sense*.

Despite this impressive history, and despite some more recent attention,[3] very little work has been done to understand the philosophical significance and bearing of what one might call the *prehistory* of Mates's Puzzle. It is well-known that Mates's Puzzle was directed against Carnap's analysis of belief sentences given in terms of intensional isomorphism; and also that intensional isomorphism itself was introduced by Carnap as a modification of his broadly Fregean theory of the semantics of belief attributions in response to problems posed by the Paradox of Analysis (Carnap, 1947, pp. 14–15). What is missing, though, is an understanding of the connections between Mates's Puzzle and the Paradox of Analysis and, rather surprisingly, between Mates's Puzzle and Frege's Puzzle.

We will be arguing that there are strong connections here. Notably, Mates's Puzzle can be understood in terms of a generalisation of Frege's Puzzle and in particular a generalisation of the sense-reference distinction, in the form of an application of at least a fragment of a robust—what we'll call, following Terence Parsons (1981), a *libertine*—hierarchy of senses. This strong connection between Mates's Puzzle and the hierarchy of senses is the main original result of this paper, one that, as far as we know, has hardly been recognized and nowhere discussed in detail.[4]

[1] We consider more recent literature concerning the puzzle in detail below.

[2] Indeed, it is mainly unhelpful to present the puzzle outside of its philosophical context—which is why we postpone its presentation to later in the paper, once the context is established.

[3] See fn5 and 6, below.

[4] Tyler Burge, in fn 12, of his 'Postscript to "Frege and the Hierarchy"'(2005b) discusses the hierarchy in connection with Mates's Puzzle, and even considers and explains how Mates's Puzzle could support not only the hierarchy, but also the possibility of a robust or libertine hierarchy (which is *not* the view that Burge is defending in the main text to which the footnote appended). However, Burge's remarks are speculative and he nowhere as far as we know elaborates on the footnote. Kit Fine (2009, pp. 129–131) considers but argues against the idea that Fregean higher-order senses can help with Mates's Puzzle. Kripke (2008) defends a view about the hierarchy according to which it is not libertine, and expresses misgivings about the view (fn. 92) which may be based on the Paradox of Analysis or Mates's Puzzle; unfortunately, he does not provide details. Mark Sainsbury and Michael Tye (2012, pp. 4.6–4.7) perhaps come closest to recognizing the view that we argue for. They argue that Mates's Puzzle compromises what they call "two-level" (without the hierarchy) Fregeanism but that "multi-level" (with the hierarchy) Fregeanism is poised to do much better because "it gives an explanation of why deeper embeddings may increase the persuasiveness

Central to our overall conclusion is a distinction between the philosophy of language and the philosophy of thought. We assume that the Fregean notions of *sense* and *Thought*, and so of the hierarchy of senses more generally, are in the first instance notions in the philosophy of thought. Frege's idea is that thoughts are what they are because of their *cognitive significance* for thinkers. Although we do not complete the task in this paper, we begin to connect this way of thinking about Thoughts to Mates's Puzzle. Our approach, and, as a result, our rendering of the Puzzle, differs from a more narrowly language-focused approach that puts weight on linguistic intuition and issues about intersubstitutivity and semantic compositionality. By contrast, on a thought-focused approach, the main issues are the extent and structure of epistemic possibility, the rationality of judgement and inference, and the identity or distinctness of concepts in the constituent structures of Thoughts as determined, ultimately, by cognitive significance, of higher-order senses in particular.[5]

Initially, we discuss Mates's Puzzle using the language-focused approach, because much of the literature with which we will be engaging itself approaches the issues in a language-focused way. However, we bring to the fore the distinction between language- and thought-focused approaches at a number of points throughout the paper, and especially in closing, where we argue for the limitations of the language-focused approach to Mates's Puzzle for making fully intelligible the ramifications of Mates's Puzzle, and in particular with respect to the hierarchy and the nature of higher-order senses.

Our conclusion is thus two-sided. On the first side, we argue that the Fregean hierarchy has been overlooked as a natural and valuable means of understanding Mates's Puzzle. We show how the hierarchy serves to illuminate Mates's Puzzle and we correct what we argue are errors in some recent literature that addresses it. On the other side, we argue that the language based-formulation of Mates's Puzzle fails to actualise the potential of the Puzzle to illuminate the nature of the hierarchy, and the rich philosophical questions it raises, and we suggest that transitioning from a language-focused to a thought-focused approach serves to highlight the question: what is the *cognitive significance* of higher-order senses? We do not answer this important question here, leaving it for a sequel to the current paper.

Footnote 4 continued
of Mates cases" (80). They however reject this solution for broadly Davidsonian learnability considerations (Davidson, 1965). Their discussion assumes what we identify as a language-focused approach to Mates and Frege Puzzles, the limitations of which we discuss below.

[5] Recent discussions of the paradox have tended to be firmly in the language-focused paradigm, as e.g., Soria Ruiz forthcoming, and the works we engage with in more depth in Sect. 3. For other relevant, but more distantly related, literature, also in the language-focused paradigm, see Baron (2015), Caie et al. (2019), Charlow and Sharvit (2014), Goodman and Lederman (2020, 2021), Lederman (2021), Tancredi and Sharvit (2020) and Schwager (2011). For discussion congenial to our approach, see Heck (2014) Sect. 2, where Heck makes a distinction between Frege's Puzzle type arguments that rely on intuition and those that rely on issues of cognitive explanation and significance. See also McCullagh (forthcoming) for a discussion of Mates's puzzle that argues for a nuanced view, "interpretative modesty", that deals directly with Mates's Puzzle and draws conclusions about concept possession and propositional attitude attribution.

1 The prehistory of Mates's Puzzle

We will be describing Mates's Puzzle in detail shortly but suffice it to say for now that in its language-focused conception, the puzzle posed is one about how the substitution of (even) synonyms fails to preserve truth-value in certain (non-quotational) contexts. The focus on *substitution-failure* brings Mates's Puzzle into close relation with Frege's Puzzle (also in its language-focused conception), but unlike Frege's Puzzle, the semantic property shared by expressions that fail substitution is not co-reference, but synonymy. The focus on substitution failures for *synonyms* brings Mates's Puzzle into the close relation with the Paradox of Analysis, which also involves substitution-failure for synonyms. In what follows, we try to relate these puzzles and paradoxes from a thought-focused perspective, tracing the historical debates and philosophical context which led to the genesis of Mates's Puzzle, and which are, we believe, necessary to fully understanding and appreciating it. We begin with Frege's Puzzle and the notion of cognitive significance.

1.1 Frege's Puzzle

Frege's Puzzle is a puzzle about propositional contents ('Thoughts', in Frege's terminology (1956)). On our take, the fundamental Fregean idea is that the Thoughts are to be understood in terms of cognitive significance. Thoughts are what they are because of their cognitive significance for thinkers. More specifically, Thoughts are individuated in terms of their cognitive significance. Thoughts are the same when their cognitive significance for thinkers is the same, and they are different when their cognitive significance for thinkers is different. That is,

Thoughts Are As Cognized
The Thought that p = the Thought that q iff p $\equiv_{CE}$ q

where '≡CE' denotes the relation of cognitive equivalence, a relation defined on contentful cognitive attitudinal states or events which obtains when these attitudinal states or events have the same content. To avoid trivialising the principle, *having the same content* should not be understood as standing in a cognitive relation to the same Thought.[6]

This still leaves open, though, a characterization of what cognitive significance and cognitive equivalence are. The relation of cognitive equivalence can be explained in terms of the possibility or not of conflict in mental attitudes, in terms of a principle along the following lines:

[6] Having made this point, we proceed below to speak in terms of a cognitive relation to propositions for ease of exposition and to connect with previous work on the issue. Note also that Thoughts Are As Cognized takes the form of an abstraction principle, hinting at the possibility that Thoughts are abstractions from an equivalence relation on contentful cognitive states or events. The idea that thoughts are objects of abstraction is underexplored. For some discussion, see Field (2017), Grzankowski and Buchanan (2019), Schiffer (2003) and Wrigley (2006).

Cognitive Equivalence
p $\equiv_{CE}$ q iff it is not possible for a thinker to take conflicting attitudes to p and q without irrationality or partial understanding.

Together, *Thoughts Are as Cognized* and *Cognitive Equivalence* entail:

Criterion of Difference
The Thought that p = the Thought that q iff it is not possible for a thinker to take conflicting attitudes to p and q without irrationality or partial understanding.[7]

It will also be useful to have a version of the *Criterion of Difference* that applies to the sense *c*onstituents of *T*houghts:

Criterion of Difference$_{cT}$
The sense s_1 = the sense s_2 iff it is not possible for a thinker to take conflicting attitudes to p = ...s_1... containing s_1 as a constituent and q = ...s_2... differing at most from p in containing s_2 as constituent without irrationality or partial understanding.

These theoretical principles about Thoughts and cognitive significance are illustrated by way of familiar and intuitive examples. For example, the Thought that Hesperus is Hesperus is distinct from the Thought that Hesperus is Phosphorus because they differ in cognitive significance. This is familiar, but an important question with which we will be closing this paper is the question of how to generalise this kind of intuitive example to illustrate the principles about Thoughts and cognitive significance for higher-order senses.

It is worth noting that there are two kinds of cases in which these principles accommodate the possibility of thinkers taking conflicting cognitive attitudes to cognitive equivalents: if the thinker (or the relevant thinking) is irrational, or if the thinker incompletely or partially understands the relevant Thought.[8] Recognizing this is important because it leaves open the possibility of an explanation of the existence of conflicting attitudes that does not proceed in terms of a difference in Thoughts. Put another way, the Criterion of Difference is not to be applied mechanically, with a conflict in cognitive attitudes automatically entailing a difference in Thoughts. It does not entail

[7] The Criterion of Difference formulated here follows the tradition of neo-Fregeanism in which the individuation of senses or sense expressed is understood in terms of conflicting attitudes and rational coherence. Here is, for one example, Gareth Evans (1982):

> The thought associated with one sentence S as its sense must be different from the thought associated with another sense S' as its sense, if it is possible for someone to understand both sentences at a given time while coherently taking different attitudes towards them, i.e., accepting (rejecting) one while rejecting (accepting), or being agnostic about, the other. (1982: pp. 18–19, his italics)

Here Evans takes rejecting and accepting to be (what we call) *conflicting attitudes* because accepting and rejecting sentences that one understands that have the same sense associated with them is not rationally coherent. Evans takes agnosticism to conflict with accepting, and presumably to conflict with rejecting as well. For more examples from across the decades see Longworth (2013, pp. 58-61), Goldberg (2008, p. 165) and Peacocke (1992, p. 2).

[8] We take the notion of incomplete or partial understanding from the externalist and anti-individualist writings of Hilary Putnam (1975) and especially Tyler Burge (1979, 1986). For discussion of whether incomplete or partial understanding is consistent with the idea of individuating thoughts in terms of cognitive significance, see Rattan (2009) and Rattan and Wikforss (2017).

such a difference when the conflict in cognitive attitudes springs from irrationality or partial understanding. These considerations will be highly relevant for the transition from language-focused to thought-focused formulations of Mates's Puzzle. While intuitions about specific examples might help us identify candidates for the application of the Criterion, they are in themselves insufficient to establish whether Thoughts are in fact the same or different, and require supplementation by consideration of matters of rationality and whether understanding is incomplete.

1.2 The Paradox of Analysis

Although the Paradox of Analysis arguably has roots in Plato's *Meno*, its contemporary formulation is most often credited to G.E. Moore (1903, chap. 1), with the term first introduced by C.H. Langford.[9] The backdrop for the paradox is a conception of philosophical inquiry as methodologically comprising in large part the analysis of concepts. The paradoxical result of the reasoning in the paradox is that philosophical inquiry so understood is impossible. Either philosophical inquiry can be informative, but in that case it cannot be analytic; or it can be analytic, but in that case it cannot be informative.

Most relevant here is that the reasoning makes use of the Criterion of Difference. Consider the putative analysis that all brothers are male siblings, and what follows from supposing (i) that it is informative, and (ii) that it is analytic.

(i) Suppose it is informative that all brothers are male siblings. Then it is possible not to believe that all brothers are male siblings. But, it is assumed, it is not possible for those who possess the (complete) concept of a brother not to (rationally) believe that all brothers are brothers. This means, according to the Criterion of Difference, that the Thought that all brothers are brothers has a different cognitive significance, and hence is distinct from the Thought that all brothers are male siblings. But if these are different Thoughts, then in what sense is one the analysis of the other? So, if the analysis is informative it looks not to be analytic.
(ii) Suppose now that it is analytic that all brothers are male siblings. If it is an analysis, the Thought that all brothers are male siblings is the same as the Thought that all brothers are brothers. But in that case, according to the Criterion of Difference, it is not possible to believe that all brothers are brothers without believing that all brothers are male siblings. But this means, according to the Criterion of Difference, that the Thought that all brothers are male siblings is not informative, at least to anyone who believes that all brothers are brothers, which we assume is everyone who possesses the concept of a brother. So, if the analysis is analytic, it cannot be informative.

This is the paradox of analysis. The paradox relies on the Criterion of Difference and thus puts stress on it. But the Criterion of Difference is the implementation of the principle that Thoughts Are As Cognized. So the more general threat is either to trivialise the notion of sense or to undermine its foundational principle. The notion of sense is trivialised if we stick to the Criterion of Difference, in which case *even*

[9] As noted by Beaney (1996). Cf. Church (1946) and Langford (1942).

the concept of a brother and the concept of a male sibling need to be distinguished. This suggests that sense is so fine-grained that no ordinary analysis will be sense-preserving. We can prevent this kind of trivialisation by allowing that sometimes there can be differences in cognitive significance that do not result in differences in sense. But in that case, we have compromised not only the Criterion of Difference but also the foundational principle that Thoughts Are As Cognized.[10]

The problem that the Paradox of Analysis poses for Frege's theory of sense is addressed by Carnap's theory of intensions and intensional isomorphism. Carnap's account is developed as part of a larger semantic account, what he calls "the method of intension and extension", a method which agrees "in many cases" with Frege's (1947, p. 50).[11] Carnap recognizes that his intensions will have, like contemporary accounts of intensions, problems with the semantics of belief attributions, because belief attributions form not intensional but hyperintensional contexts where the substitution of co-intensional expressions does not generally preserve truth-value.[12] What is required, then, is a criterion of intersubstitutability ("interchangeability" in Carnap's language) that is more fine-grained than sameness of intension. This criterion is intensional isomorphism.

On Carnap's view, intensions are unstructured, and intensional isomorphism is a relation between sentences.[13] It is the relation sentences bear to each other when they express the same intension in the same structured way, with corresponding parts of the sentences expressing the same intensions.[14] Intensional isomorphism is obviously more fine-grained than sameness of intension and so intensional isomorphism can serve as a stricter condition for interchangeability than co-intensionality. Intensional isomorphism can then explain why some substitutions of co-intensional expressions in belief attributions do not preserve truth-value (when the attributing sentences that express the truths differ in structure) and explain why analysans cannot be substituted for analysandum in belief contexts (because the attributing sentences expressing the analysans and the analysandum differ in structure). The overall strategy is to argue that both intensional isomorphism and sameness of intension are theoretically important, and when this is recognized the intuitions about sameness and difference that underpin the paradox can be simultaneously and non-paradoxically explained.

10 Cf. Nelson (2008, p. 175).

11 For further discussion see Chalmers (2012, p. 250) and Stanley (2014, p. 1).

12 This is obviously *not* true of the Fregean account of Thoughts, when Thoughts are individuated in accordance with the Criterion of Difference.

13 See Carnap (1947, pp. 26–7), around proposition 6.2, where he identifies intensions with propositions and individuates propositions in terms of his notion of "L-equivalence". Although Carnap's view is about sentences, it can be transposed to the level of thought to fit with our framework by allowing for structured intensions.

14 Here is Carnap's informal definition: "If two sentences are built in the same way out of designators (or designator matrices) such that any two corresponding designators are L-equivalent, then we say that the two sentences are intentionally isomorphic or that they have the same intensional structure" (1947, p. 56).

1.3 Mates's Puzzle

In response to Carnap, whose view he finds attractive, Mates raises the following worry:

> Let "D" and "D'" be abbreviations for two intensionally isomorphic sentences. Then the following sentences are also intensionally isomorphic:
>
> Whoever believes that D, believes that D.
>
> Whoever believes that D, believes that D'.
>
> But nobody doubts that whoever believes that D believes that D. Therefore, nobody doubts that whoever believes that D believes that D'. This seems to suggest that, for any pair of intensionally isomorphic sentences-let them be abbreviated by "D" and "D'",-if anybody even doubts that whoever believes that D believes that D', then Carnap's explication is incorrect. (1950, p. 215)

The exposition is swift, but the threat is powerful. It rests on the idea that no one is likely to doubt what is expressed by any sentence of the form "whoever believes that D believes that D", but we can readily imagine that someone might doubt what is expressed by a sentence of the form "whoever believes that D believes that D'", even where "D" and "D'" are intensionally isomorphic. Although intensional isomorphism was supposed to account for subtitutivity problems for intensions, it has substitutivity problems of its own. The conclusion, as Mates articulates it, is more general:

> …any adequate explication of synonymity will have this result, for the validity of the argument is not affected if we replace the words "intensionally isomorphic" by the word "synonymous" throughout. (1950, p. 215)

According to Mates, there is a problem with accounting for substitution failures by an appeal to the notion of synonymy, one which intensional isomorphism as an "approximate explicatum" (Mates, 1950, p. 215) for synonymy cannot solve.

It is easier to work through these ideas with an example. While Mates does not provide one, Putnam (1954, p. 117), gives an example using the expressions "Greeks" and "Hellenes" for "D" and "D'". Putnam stipulates their intensional isomorphism and synonymy "for the sake of illustration" (117).[15] Indeed, for the example to yield the kind of puzzle Mates had in mind, one in which schematics "D" and "D'" are intensionally isomorphic and synonymous, we must take Putnam's "Greeks" and "Hellenes" to be intensionally isomorphic, and, more generally, synonymous, or, in Fregean terms, to have the same sense. In other words, Putnam's example assumes and requires us to grant that the Thought constituent or sense expressed by "Greeks", which we write as "<Greeks>", is the same as the Thought constituent or sense expressed by "Hellenes", which we write as "<Hellenes>". That is, Putnam's example takes for granted that

[15] Putnam does not himself commit to the synonymy of "Greeks" and "Hellenes"—and neither are we inclined to—but insofar as he appeals to common usage of the two terms by newspapers, there is no doubt that he intends them to be widely thought to be synonymous. Intuitions are relevant here but the problem as conceived by Mates and indeed by Putnam also importantly depends on a theoretical surround and involves theoretical choices in formulating the example.

<Greeks> = <Hellenes>.[16]

Moreover, since the other constituents are trivially identical, and are composed in the same way, it follows that

<all Greeks are Greeks> = <all Greeks are Hellenes>

Now, using "Greeks" and "Hellenes" in the Mates formulation yields the following pair of sentences:

a) Whoever believes that all Greeks are Greeks believes that all Greeks are Greeks.
b) Whoever believes that all Greeks are Greeks believes that all Greeks are Hellenes.

Introducing the doubt yields the following iterations:

c) Nobody doubts that whoever believes that all Greeks are Greeks believes that all Greeks are Greeks.
d) Nobody doubts that whoever believes that all Greeks are Greeks believes that all Greeks are Hellenes.

The Puzzle is then usually taken to run as follows: (a) and (b) are both true and preserve truth-value upon the substitution of synonyms. In addition, (c) is generally agreed to be true. However, (d) is intuitively false. That is, although (b) is true, we seem to have an intuition that someone might doubt that it is. So the truth-values of (c) and (d) differ, despite the synonymy of "Greeks" and "Hellenes".

Putnam summarises the significance of this observation in one particular and influential direction:

> In short, on any theory of synonymity, the synonymity of [doubly embedded attributions] must follow if the synonymity of 'Greek' and 'Hellene' is assumed. If we take this seriously, there is but one conclusion to which we can come: 'Greek' and 'Hellene' are not synonyms, and by the same argument, neither are any two different terms. (1954, p. 117)

This is a classic statement of the idea that Mates's Puzzle shows the notion of synonymy to be trivialised. Others, however, doubt that the substitution is successful even in (a) and (b) and are inclined to believe that while (a) is true, (b) is false, so that the second level of iteration in (c) and (d) is unnecessary.[17] This is an immediate peril of the language-focused approach and its central reliance on linguistic intuition. The result is to dismiss an important aspect of the Puzzle as Mates and Putnam formulate it.

However, our thought-focused approach allows us to bypass these intuitions and focus on the consequences that attend the puzzle when we assume that <Greeks> = <Hellenes>. On this assumption, (a) and (b) cannot differ in truth-value, since they

[16] We take up in more detail below an issue about how the hierarchy articulates and interprets the notion of synonymy itself, allowing synonymy relations of different strengths. See Sect. 2b below. The notion we use here is the weakest notion. See also Rattan (2019).

[17] See, for example, Burge (1978, p. 126), which we discuss at greater length below. This has been a common reaction from audiences as well.

have the same constituents and thus are the same Thought. This makes the second-order formulation an essential feature of the puzzle, just as Mates and Putnam presupposed. This also allows one side of our two-sided conclusion to come to the fore: that Mates's Puzzle finds an intuitive and valuable resource in Fregean semantics (including the sense hierarchy) which set off the debates from which it arose. Exploring this connection will allow us to consider Mates' Puzzle in a different light, one obscured by the language-focused approach.

2 The Fregean Hierarchy and Mates's Puzzle

Our idea is to explain the failure of substitution which is traditionally identified in (c) and (d) by appealing to the Fregean hierarchy and differences in higher-order senses.[18] The hierarchy has been discussed at length.[19] We present here only a brief summary of the theory in order to show its relation to Mates's Puzzle.

2.1 The Fregean hierarchy

The Fregean hierarchy flows from basic Fregean semantics. As is well known, Frege introduced the distinction between sense and reference, arguing that expressions refer to their referents and express their senses. For example, in

(1) Superman is strong

"Superman" refers to Superman and expresses a sense,<Superman>. In addition, on Frege's view, belief contexts induce a shift in reference. Expressions within the scope of propositional attitude verbs refer not to their customary referent but to their customary sense, or mode of presentation, and express an indirect sense. Thus, in

(2) Lois Lane believes that Superman is strong

"Superman is strong" refers to the sense, or Thought,<Superman is strong>and expresses a higher-order sense, which we write<<Superman is strong>>, "Superman" expresses a higher-order sense <<Superman>>and refers to the first-order sense or mode of presentation <Superman>, and "is strong" expresses a higher-order sense <<is strong>> and refers to the first order sense or mode of presentation<is strong>. Assuming that Thoughts are structured out of component senses, we can then analyse (2) in the following way, adopting Christopher Peacocke's (2008) notation[20]:

(2') BEL (Lois, <Superman>ˆ<is strong>)

[18] Whether Frege himself held the view as we outline it here is debated, though it is suggested in a comment in a letter to Russell dated 28.12.1902 (reprint. in McGuinness, 1980), in which Frege talks of "indirect senses of the second-degree" (154).

[19] In addition to work we discuss below, see for example Boisvert and Lubbers (2003), Burge (2005a), Davidson (1965), Kripke (2008), Künne (2010), Parsons (1981, 2009), Salmon (1986), Simchen (2018) and Skiba (2015).

[20] The angle bracket notation pre-dates Peacocke, but different authors differ on other details about how to use this notation, and on these differences, we follow Peacocke (with some reservations indicated in the next note).

Following Peacocke, we use “^” as a symbol for combination of senses, and BEL as a two-place relation between an agent and a Thought. (2), then, states that Lois stands in relation to a Thought, namely the Thought <Superman is strong> , which is built from the senses <Superman> and <is strong> , modes of presentations of Superman and the property of being strong, respectively.[21] The doctrine of reference shift in belief contexts gives rise to a distinction between first-order senses, which are expressed by unembedded expressions and referred to by expressions within the scope of a propositional attitude verb, and second-order or indirect senses, which are expressed by expressions within the scope of a propositional attitude verb. The hierarchy is generated by iterating embedding, which generates higher-order indirect senses.

Expressions within the scope of an embedded propositional attitude verb refer to an indirect sense and express a higher-order indirect sense. For example, consider:

(3) Amy believes that Lois Lane believes that Superman is strong

We know from our analysis of (2) that “Lois Lane believes that Superman is strong” refers to a first-order sense, or Thought, <Lois Lane believes that Superman is strong> . Because “Superman is strong” is now doubly embedded, its reference is shifted, so that it refers to a second-order or indirect sense, <<Superman is strong>> , which is a mode of presentation of the Thought <Superman is strong>, and expresses, a higher-order indirect sense, <<<Superman is strong>>>. The truth condition for (3) can be rendered as follows:

(3’) BEL (Amy, <BEL>^<Lois>^<<Superman>>^<<is strong>>)

One sees, in comparison with (2’), the ascent in levels. Further iteration yields a hierarchy of senses.

2.2 Connecting the Fregean hierarchy and Mates’s Puzzle

We now begin connecting the Fregean hierarchy to Mates’s Puzzle. In subsequent sections, we analyze this connection in greater depth, and in particular try to discern what can be learnt about each from the other. This section lays out the general ideas and draws some basic lessons.

Let us begin with the first pair of sentences in Putnam’s example formulation:

(a)Whoever believes that all Greeks are Greeks believes that all Greeks are Greeks.

[21] We treat ‘<Superman> ^ <is strong>’ as coreferential with ‘<Superman is strong>’. This simplification is fine for unembedded sentences, but obscures information about differences in levels of hierarchy for different expressions in embedded sentences (e.g., ‘Lois believes that Superman is strong’ strictly speaking expresses the Thought<BEL>^<Lois>^<<Superman>>^<<is strong>>, the hierarchy levels of which are obscured in <Lois believes that Superman is strong>). More generally, it is worth noting that Peacocke’s notation does not distinguish between the combination of <<Superman>> and <<is strong>> and the combination of <Lois> and <<Superman>>, although these are arguably different. An ideal notation would stipulate an order of operations, as it were, making it clear that the combinations of <<Superman>> and <<is strong>> on the one hand and of <Lois> and <BEL> on the other are prior to that of <Lois> and <<Superman>>. We use the notation despite this limitation, and the simplification for readability where it does not affect the discussion.

(b) Whoever believes that all Greeks are Greeks believes that all Greeks are Hellenes.

Recall that, on the thought-focused account, the assumption that “Greeks” and “Hellenes” are synonymous expressions is understood as the assumption that.

<Greeks> = <Hellenes>

and therefore that

<All Greeks are Greeks> = <All Greeks are Hellenes>

“All Greeks are Greeks” and “All Greeks are Hellenes” refer, in (a) and (b), to this Thought. The pair can thus be rendered in the following way:

(a’) ∀x(BEL(x, <All Greeks are Greeks>)→BEL(x,<All Greeks are Greeks>))
(b’) ∀x(BEL(x, <All Greeks are Greeks>)→BEL(x,<All Greeks are Hellenes>))

Since <All Greeks are Greeks> = <All Greeks are Hellenes>, it should now be clear that (a) and (b) cannot differ in truth-value. This isn’t to say that the truth of (b) is self-evident such that no one could doubt it: it may be true that a thinker believes that not all Greeks are Hellenes, but she also believes that all Greeks are Hellenes, since <All Greeks are Greeks> = <All Greeks are Hellenes>, and so believing the former just is believing the latter. An inclination to doubt (b) or an intuition that someone might doubt (b) stems not from there being two distinct Thoughts being referred to in (b), but rather from irrationality or from partial understanding of the Thought, and this is something for which the Criterion of Difference makes provision. We see here our thought-focused approach at work, in contrast with a language-focused approach, which risks putting too much weight on intuitions about (b) and thereby obscuring the need for and interest in the next level of the puzzle—a problem to which we return below.

Let us now turn to the iterated sentences in Putnam’s example formulation:

(c) Nobody doubts that whoever believes that all Greeks are Greeks believes that all Greeks are Greeks.
(d) Nobody doubts that whoever believes that all Greeks are Greeks believes that all Greeks are Hellenes.

The inclination to doubt (b) or the intuition that someone might doubt (b) is at play here: the truth-value of (c) and (d) are different because we can imagine someone, say Bates, who might falsify (d). Bates might falsify (d) because she herself doubts (b) or because she believes that someone, say Cates, doubts (b). On our approach, and in accordance with the Criterion of Difference, either scenario can be explained in terms of partial understanding. Bates might doubt (b) as a result of her own partial understanding of the concept <Greeks> (= <Hellenes>) or she might doubt that Cates believes (b) because she attributes partial understanding (or irrationality) to Cates.

Although the thought-focused approach allows us to appeal to partial understanding in this way, it is important to note that an appeal to partial understanding is not required in order to explain how (c) and (d) can have different truth-values. It need not be the case that anyone actually doubts (b) in order for (d) to be false. We only need to assume that someone, say Bates, might believe that someone else, say Cates, doubts (b), even though Bates herself does not doubt (b), and regardless of whether or not Cates does (or whether Cates even exists). Bates can herself have full understanding of <Greeks> (= <Hellenes>) so that she knows (b), and can be wrong in ascribing doubt to Cates who also has full understanding of <Greeks>and knows (b). It only matters that Bates believes it possible for Cates, or someone, to doubt (b). In other words, the puzzle can formulated even in a world in which no one in fact doubts (b)—that is, in a world in which everyone knows that <Greeks> = <Hellenes>, and so in a world in which everyone has full understanding of the concept <Greeks>.[22] If this is right, then any view that takes the multiple embedding in Mates's Puzzle to be inessential, and that thereby collapses Mates's Puzzle into the Paradox of Analysis, automatically mistakes the nature of Mates's Puzzle, since the Paradox of Analysis has to involve incomplete understanding, but Mates's Puzzle does not.

The difference in truth-value between (c) and (d) is a result of the multiple embedding in Mates's Puzzle. Because of the double embedding, "all Greeks are Greeks" in (c) and (d) does not refer to the first-order sense, <All Greeks are Greeks>, like it did in (a) but to the second-order sense, <<all Greeks are Greeks>> . Similarly, "all Greeks are Hellenes" in (d) refers to the second-order sense <<all Greeks are Hellenes>>. (c) and (d) can have different truth-values because

<<All Greeks are Greeks>> ≠ <<All Greeks are Hellenes>>

On this account, there are multiple higher-order senses (<<All Greeks are Greeks>> and <<All Greeks are Hellenes>>) that present one and the same first-order sense (<All Greeks are Greeks> = <All Greeks are Hellenes>).

The hierarchy allows us to respect the need for synonymy in the Puzzle, but it does so with a more articulated understanding of synonymy. What synonymy as sameness of sense amounts to, once the hierarchy is in place, depends on sameness *of which* sense(s). Sameness of customary or first-order sense constitutes *a*—the weakest—relation of synonymy. Sameness of sense all the way up the hierarchy constitutes the strongest relation of synonymy. Intermediate strengths can also be defined as sameness of sense up to some n-th level with a difference in n + 1 level senses. The Puzzle as actually presented by Mates and Putnam and interpreted through the hierarchy requires synonyms according to the weakest relation of synonymy. There is room in logical space for versions of the Puzzle different from the ones Mates and Putnam actually present that are nevertheless genuine versions of the Puzzle, using synonyms in accordance with a weak, but not necessarily the weakest, relation of synonymy. Such versions of the Puzzle would be generated by further embeddings. However, no

[22] This point is similar to one made in passing by Kripke (1979, p. 282 n46):

> Mates's problem would not arise in a world where no one ever was under a linguistic or a conceptual confusion, no one ever thought anyone else was under such a confusion, no one ever thought anyone ever thought anyone was under such a confusion, and so on.

Kripke's point is that Mates's Puzzle requires someone to be in the grips of linguistic or conceptual confusion, or someone to think that someone is in the grips of linguistic or conceptual confusion, etc. Our point, which follows from Kripke's, is that, in his terms, no speaker or thinker needs to be in the grips of linguistic or conceptual confusion in order to be able to formulate the puzzle. McCullagh forthcoming concurs with the spirit of Kripke's remark, although not with the letter.

version of the Puzzle can be formulated using synonyms according to the strongest relation of synonymy. That is, the Puzzle requires a difference in some higher-order senses. Synonyms in the strongest sense, with sameness of sense all the way up the hierarchy, could not, under any level of embedding, yield sentences with different truth-values. The hierarchy not only interprets the Puzzle, it interprets its key notion, synonymy.

We will return below to what Mates's Puzzle can—and, importantly, cannot—tell us about the hierarchy (Sect. 4), but for now it is worth noting that this application of the hierarchy to Mates's Puzzle challenges some common views about the hierarchy in the literature. First, it poses a problem for views that deny the existence of the hierarchy. For example, in an Appendix to his 'Quantifying in from a Fregean Perspective' (2015 Appendix A), Seth Yalcin expresses suspicion about the hierarchy. However, in his formal semantics, the hierarchy is ruled out by stipulation:

> We stipulate that there are only two values that the r-parameter can take, given any language L. One value is r_s, the reference function determined by the sense function s given by the model for the language. (We assume that every sense determines a reference, so that any *s* induces a corresponding r_s.) (....) The second value the reference parameter can take is *s* itself. In that case, the expression is mapped directly to its sense. (244)

The "r-parameter" is a parameter to which the semantic value of an expression is relativized; intuitively, it is the parameter relative to which the shift in semantic value that Fregean semantics posits in belief reports is to be understood. Yalcin gives no justification for his stipulation, the result of which is to eliminate the hierarchy. But, we think, eliminating the hierarchy leaves Yalcin open to Mates's Puzzle worries, since the hierarchy is required to generate expressions which share the same first-order sense but differ in second-order sense.

Second, the application of the hierarchy to Mates's Puzzle poses problems for views that constrain the kind of relation that can hold between higher-order senses and the senses that they are senses of. This is the case with views that limit the hierarchy to *rigid hierarchies.* Proponents of a rigid hierarchy claim that although there can be multiple first-order or customary senses that present one and the same customary referent, there cannot be multiple higher-order indirect senses presenting one and the same lower-order sense. That is, beyond the level of customary reference, lower-order senses determine higher-order senses, so that the relation between higher-order senses and the senses that they are senses of is one-to-one.

Proponents include Christopher Peacocke (2008) and Terence Parsons (1981, 2009).[23] Peacocke notes in response to the question "Does a hierarchy of canonical concepts collapse into, or at least make available, a single-level account of sense and concepts?":

[23] One advantage of a rigid hierarchy is that it admits a response to Donald Davidson's learnability objection (Davidson, 1965). It is, as far as we know, an open question whether and how a libertine hierarchy can be squared with learnability considerations. We do not take up the issue here but hope that the discussion here will encourage attention to it.

> In his early paper 'Frege's Hierarchies of Indirect Sense and the Paradox of Analysis', Terence Parsons argues that the hierarchy does so collapse. He classifies as 'rigid' any theory that holds that 'the customary sense of an expression uniquely determines its indirect sense' (p. 44). The theory I have offered is certainly rigid in this sense. (313)

Parsons in fact takes this further with an argument—which Peacocke resists—that the rigid nature of the hierarchy allows the higher-order senses to collapse into the first-order senses. But rigid hierarchies of the kind Peacocke and Parsons defend do not explain Mates's cases. To account for Mates's cases, there must be the possibility of multiple indirect senses presenting one and the same lower-order sense. Seen in this light, Mates's Puzzle shows both that the hierarchy must exist, and that it cannot be rigid.

3 Mates's Puzzle in light of the hierarchy

Understanding the connection to the hierarchy in turn sheds light on our understanding of Mates's Puzzle. In this section, we discuss two examples from the literature where there is an instructive failure to grasp this connection.

3.1 Moffett, Mates, and level of embedding

Marc Moffett (2002) presents an argument for the conclusion that any instance of Mates's Puzzle is equivalent to some instance of Frege's Puzzle. Moffett's conclusion is roughly that Mates's Puzzle collapses into Frege's Puzzle. Our view by contrast is that Mates's Puzzle suggests a generalisation of the standard Frege Puzzle, and of the distinction between sense and reference in the form of the hierarchy. On our view, Moffett's argument fails to appreciate the relationship between Mates's Puzzle and the hierarchy, and the significance of the doubly embedded formulation.

Moffet specifically sets out to show that there is no "strictly lexical solution" to Frege's Puzzle. According to him, Mates's Puzzle trivialises the notions of lexical synonymy, making it useless for solving Frege's Puzzle. Using "f is computable" and "f is recursive" as examples of synonymous expressions, he justifies this trivialisation result as follows:

> If somebody, say x, doubts that whoever believes that f is recursive, believes that f is computable, they must on reflection doubt that the proposition that f is recursive is the same proposition as the proposition that f is computable. For this person doubts that the proposition that f is recursive has the property of being believed by y (for some y) while the proposition that f is computable lacks this property. (163)

The claim about trivialisation is supposed to follow from the claim that x "doubts that the proposition that f is recursive is the same proposition as the proposition that f is computable". The idea is that if *even* this identity of Thoughts can be doubted, there cannot be any non-trivial Thought identities. Further, the claim that it is dubitable that the proposition that *f* is recursive is the same proposition as the proposition that *f* is computable is supposed to follow from the claim that x "doubts that the proposition that *f* is recursive has the property of being believed by y (for some y) while the proposition that *f* is computable lacks this property."

However, noting the embedded formulation and the interaction with the hierarchy shows that Moffett actually makes two mistakes here.

First, supposing that x does doubt that the proposition that *f* is recursive is the same proposition as the proposition that *f* is computable (that is, x doubts that < *f* is computable > = <*f* is recursive>), it does not follow that the two propositions are not the same—that is, that < *f* is computable > ≠ < *f* is recursive>). It follows only that certain modes of presentation of those propositions are not identical—that is, that <<*f* is computable>> ≠ <<*f* is recursive >>. But trivialisation of the sort Moffett is trying to argue for requires that < *f* is computable > ≠ < *f* is recursive>.

Second, there is in fact no reason to think that x doubts the identity. From x's doubt that whoever believes that*f* is recursive believes that x is computable, it does not follow that x doubts that < *f* is computable > = < *f* is recursive>. What does follow is that there are distinct modes of presentation of the proposition <*f* is computable>(= < *f* is recursive>), namely << *f* is computable >> and << *f* is recursive>>, such that x has distinct attitudes to two distinct propositions which contain these respective modes of presentation, namely:

the proposition that y believes that if*f* is computable, then*f* is recursive
=<y believes that if *f* is computable, then *f* is recursive>
=<BEL>^<y>^<<if *f* is computable, then *f* is recursive>>
=<BEL>^<y>^<<IF >>^<< *f* is computable>>^<< *f* is recursive>>

which contains both << *f* is computable >> and << *f* is recursive >> , and is such that it has the property of being doubted by x; and

the proposition that y believes that if*f* is recursive, then *f* is recursive
= <y believes that if *f* is recursive, then *f* is recursive>
=<BEL>^<y>^<<if *f* is recursive, then *f* is recursive>>
=<BEL>^<y>^<<IF>>^<<*f* is recursive>>^<<*f* is recursive>>

which contains <<*f* is recursive>> but not <<*f* is computable>> and is such that it does not have the property of being doubted by x. So x's attitudinal states of affairs are:

DBT (x,<BEL>^<y>^<<IF>>^<<*f* is computable >>^<<*f* is recursive>>)
BEL (x,<BEL>^<y>^<<IF>>^<<*f* is recursive>>^<<*f* is recursive> >)

From this it does not follow that x doubts that <*f* is computable> = <*f* is recursive>. As we noted earlier in connection with our original examples, Bates needn't herself doubt that all Greeks are Hellenes in order to doubt that Cates believes that all Greeks are Hellenes. This is, indeed, a key difference between Mates' Puzzle and the Paradox of Analysis. Similarly, x does not have to doubt that if *f* is computable, then *f* is recursive in order to doubt that y believes that if *f* is computable, then *f* is recursive.

So, we conclude, appreciating the relationship of Mates's Puzzle to the hierarchy and the importance and significance of the iterated formulation shows that it is incorrect that every instance of Mates's Puzzle is an instance of (a standard) Frege's Puzzle. The relation between Mates's Puzzle and Frege's Puzzle is different. Mates's Puzzle does not collapse into Frege's Puzzle but rather suggests a generalisation of Frege's Puzzle in the form of a libertine hierarchy, where Frege-style Puzzles can arise at any level of the hierarchy. Our conclusion helps separate out the respective philosophical contributions of Mates's Puzzle, Frege's Puzzle, and the Paradox of Analysis.

3.2 Sorensen, mates, and meaningless beliefs

Roy Sorensen (2002) presents a "proof" of the possibility of meaningless beliefs using Mates's Puzzle. The proof begins by formulating Mates sentences that include the meaningless expressions "reywals" and "yenrottas" ("lawyer" and "attorney" spelt backwards and pluralised), yielding:

(a') Whoever believes reywals are wealthy believes that reywals are wealthy.
(b') Whoever believes reywals are wealthy believes that yenrottas are wealthy.
(c') Nobody doubts that whoever believes reywals are wealthy believes that reywals are wealthy.
(d') Nobody doubts that whoever believes reywals are wealthy believes that yenrottas are wealthy.

Just as with the original puzzle, we can imagine someone, say Cates, who believes (a') but doubts (b'), making (c') true but (d') false. It is worth noting again that Cates needn't really hold those beliefs, it only needs to be possible for us to imagine that she does. As Sorensen emphasises, "Cates does [her] damage even if [s]he is a figment of Professor Mates's imagination".

On the basis of this, Sorensen presents his proof for the possibility of meaningless beliefs:

[1] I believe that someone can believe that reywals are wealthy.
[2] If I believe that someone believes something, then I attribute an object of belief to him.
[3] If I believe that someone believes that reywals are wealthy, then if I attribute any object of belief to him, it is that reywals are wealthy.
[4] Any object of belief can be believed. (172)

The conclusion is that that reywals are wealthy is something that can be believed. Sorensen takes this to be tantamount to the idea that there can be meaningless beliefs.

However, reading Mates's Puzzle in light of the Fregean hierarchy shows that premise [2] is problematic. [2] states "if I believe that someone believes something, then I attribute an object of belief to him". Since Sorensen attributes a belief to Cates, namely.

(4) Cates believes that reywals are wealthy.

there should be an object of belief that he attributes to Cates. Now if "reywals are wealthy" expressed a Thought, "that reywals are wealthy" would refer to it and would express a second-order sense << reywals are wealthy >> that presents it.

However, it is commonly accepted that there can be, on a Fregean view, sense without reference.[24] There are empty names, for example "Vulcan", which have no referent, but nonetheless express a sense. This can happen at higher levels of the hierarchy also. That is, there can be empty names of Thoughts, which express higher-order indirect senses, but have no referent. When we attribute a belief, we name the Thought which we attribute. When we attribute a meaningless object of belief to someone, we use an empty name. "that reywals are wealthy", as it appears in (4), can thus be analyzed as an empty name: although it expresses a higher-order sense <<reywals are wealthy>> , there is no Thought to which it refers.[25]

It is therefore not the case that from the attribution of the belief to Cates in (4) we can infer that there is in fact an object of belief. (4) does not require the existence of an object of belief; it requires only the existence of a name or singular concept, which may be empty. Understanding the connection between Mates's Puzzle and the hierarchy therefore undermines Sorensen's proof for the possibility of meaningless beliefs, and shows an application for sense without reference for names of Thoughts.

4 The hierarchy in the light of Mates's Puzzle

We have been looking at the connection between Mates's Puzzle and the hierarchy. We hope to have shown that appreciating this connection allows for a firmer grasp on Mates's Puzzle as formulated by Mates and Putnam than is currently found in the literature. We close this paper by considering what these connections can and can't teach us about the hierarchy and higher-order senses themselves. In so doing, we highlight the limits of the language-focused treatment of Mates's Puzzle. We argue that Mates's Puzzle has the potential to illuminate the hierarchy, much as the hierarchy illuminates it, but that, in its language-focused formulation, Mates's Puzzle forms an inadequate basis for understanding the hierarchy and especially the nature of higher-order senses. We suggest that we need to replace the language-focused formulation with a thought-theoretic formulation that attends to the question of what *the cognitive significance of higher-order senses* consists in—a question the answer to which will have to await another occasion.

Let's begin by turning back to our theoretical principles governing Thoughts. Although they remain largely intact once we recognize the hierarchy, we can now see that our initial formulation of the *Criterion of Difference*$_{cT}$—the criterion of difference for constituents of Thoughts—requires revision. Indeed, higher-order senses are not Thoughts themselves, but constituents of Thoughts, and they are individuated by their occurrences and behaviour in higher-order embeddings. That is, higher-order senses are not by their nature constituents of first-order thinking, but of second- and higher-order thinking. But level of embedding is not mentioned in the *Criterion of Difference*$_{cT}$. This omission needs rectifying.

Recall the *Criterion of Difference*$_{cT}$:

***Criterion of Difference*$_{cT}$**

[24] This is of course contentious – see Evans (1982, chap. 1) and McDowell (1984, 1986).

[25] One consequence of this is that (4) will lack a truth-value.

> The sense s_1 = the sense s_2 iff it is not possible for a thinker to take conflicting attitudes to p = ...s_1... containing s_1 as a constituent and q = ...s_2... differing at most from p in containing s_2 as constituent without irrationality or partial understanding.

Now, these constituents, s_1 and s_2, must both be of the same order. For example, it cannot be the case that one is a higher-order sense and the other a customary sense. If that were the case, one or the other would be an ill-constituted Thought, or nonsense, and our thinker would necessarily be taking an attitude towards it *without understanding*, which is the limit case of partial understanding. Thus, the *Criterion of Difference*$_{cT}$ as formulated above leads to the trivialising result that any customary sense is identical to any higher-order sense.

Two principles therefore need to be incorporated into the *Criterion of Difference*$_{cT}$ in order to accommodate higher-order senses. The first constrains the original principle as it applies to higher-order senses, and says that no sense of any order *n* is the same sense as any sense of order *m*, if $n \neq m$. The second modifies the original principle to incorporate the relevance of order for higher-order senses.

Let s^0 be some customary sense and let $s^n{}_i$ be one of the *n*-th order senses (the *i*-th, on some arbitrary ordering) of s^0. For example, $s^0{}_1$ is one of the indirect senses (the first, on some arbitrary ordering) of s^0. Further, let 'about$_n$' denote level of embedding, where 'about$_0$' denotes the relation that holds between Thoughts and the world, 'about$_1$' denotes the relation between Thoughts and the Thoughts that they are thoughts about, etc. We can now formulate *the generalised Criterion of Difference*, the Criterion of Difference for Constituents of Thoughts, *including* higher-order senses:

***Criterion of Difference*$_{generalised}$**

1. For all *i, j, m, n* such that $n \neq m$, $s^n{}_j \neq s^m{}_i$
2. The sense $s^n{}_i$ = the sense $s^n{}_j$ iff it is not possible for a thinker to take conflicting attitudes to Thought about$_n$ Thoughts …$s^n{}_i$… containing $s^n{}_i$ as a constituent and Thought about$_n$ Thoughts …$s^n{}_j$… differing at most from Thought about$_n$ Thoughts …$s^n{}_i$… in containing $s^n{}_j$ as constituent, without irrationality or partial understanding.

This modification of the *Criterion of Difference*$_{cT}$ is the result of introducing higher-order senses and the hierarchy into our understanding of Thoughts.

As far as we know, a criterion of difference of this sort that applies to higher-order senses has not been formulated in the literature. Note that this criterion serves to individuate higher-order senses by their canonical roles in higher-order attitude attributional Thoughts, but this does not imply that higher-order senses *cannot* occur in ordinary first-order (or other order) thinking. For example, I can believe that the higher-order sense <<<<brother>>>> is identical to the higher-order sense <<<<male sibling>>>>>. In order to think about some n-th level sense, one does not need to think a thought that includes n + 1-orders of embedding; one can in principle just think about that sense directly in first-order thinking.

However, although higher-order senses can be thought about in first-order thinking, their role in such thinking is not part of a constitutive explanation of higher-order senses. We can make higher-order senses the subject matter of thinking, but we do not really know what, for example, the higher-order sense <<<<brother>>>> *is* such that it could be presented to even higher-order thinking. Although we will not be able to show this in detail here (because it requires addressing the question of the cognitive significance of higher-order senses), we think that higher-order senses are much more illuminatingly explained in their occurrence and behaviour in multiply embedded attributional Thoughts, as exemplified in Mates's Puzzle.

Yet, we do not think that Mates's Puzzle, in its standard language-focused formulation, despite its use of multiple embedding, can carry the burden of such an explanation. Although Mates's Puzzle as standardly formulated is illuminating and suggestive about Fregean senses and especially the hierarchy, we do not think that it provides an adequate foundation for understanding the hierarchy and the nature of higher-order senses. More specifically, in its language-focused formulation, Mates's Puzzle fails to justify (a) clearly moving *even one rung up* the hierarchy, and, once at the first-rung, (b) moving *any rungs beyond* the first rung of the hierarchy. Let us take each in turn.

(a) As we have emphasised, a key feature of Mates's puzzle is the second-order formulation. Yet, some have raised questions over this formulation. Tyler Burge, for example, at one time cast doubt on the idea that the second-order formulation is essential, suggesting that the puzzle could be formulated using just a single-embedded formulation. Burge (1978) writes:

> If the examples are taken at face value, theories that allow substitutions of synonyms in belief contexts *on logical grounds* are wrong even for unembedded contexts. In what follows, I shall be assuming that it takes an *argument* to show that the examples should not be taken at face value… (126)

Burge's point relies on the idea, supported by linguistic intuition, that substitution into even singly-embedded attributions could result in difference in truth-value. This would suggest (in Putnam's example) that ‘Greeks’ and ‘Hellenes’ are after all not synonyms or at least not co-sensical, and this would in effect collapse Mates's Puzzle into a version of the Paradox of Analysis.

A thought-focused conception of Mates's Puzzle can, however, hold more tightly to co-sensicality, in the form of the identity <Greeks> = <Hellenes>. In particular, the thought-focused formulation can appeal to considerations about rationality and partial understanding to explain away apparent differences in truth-value. As applied to the Paradox of Analysis, the thought-focused formulation allows an interpretation on which analysans and analysandum express identical senses, but of which speakers quite ubiquitously display a partial or incomplete understanding. As we have shown, a thought-focused analysis of Mates' Puzzle avoids assigning different truth-values to the two first-order attributions, since they are, as a matter of their constituent structure, the same Thought. By moving away from linguistic intuitions, the thought-theoretic approach recognizes the possibility of a puzzle—Mates's—that does not rest on or require actual partial or incomplete understanding and avoids confusing it with a paradox that does—the Paradox of Analysis.

(b) Even if the language-focused account were given a pass on the problem it faces with respect to the essentially second-order formulation, it encounters further problems in justifying the next move up a robust hierarchy of senses, one that both goes beyond the second-order and allows the many-to-one relation of sense to reference to continue (as it does at the second-order). The problem runs as follows. To move up the first step in the hierarchy, the language-focused formulation considers synonyms (‘Greeks’ and ‘Hellenes’ in Putnam’s example), and shows a failure of substitution without a change in truth-value in double-embedded attitude attributions. How is the process to be iterated to take us further up the hierarchy? What is required are expressions that are (i) coreferential, (ii) *first-level* synonyms (as we assumed ‘Greeks’ and ‘Hellenes’ to be in Putnam’s example), (iii) *second-level* synonyms (unlike ‘Greeks’ and ‘Hellenes’), but, (iv) not *third-level* synonyms. And so on at higher levels.

Our claim is that a language-focused extension to levels beyond the doubly-embedded as envisioned here is dubious. The language-focused account does not explain what these different levels of synonymy are or how they could be instantiated in our propositional attributional practices. There is no account of the nature of such synonymy relations, or putting it in Fregean terms, of the sameness and difference of higher-order senses. We have the individuating principle, but we do not really understand the *linguistic semantic* situation that underpins it and for which the individuating principle is meant to provide senses. The language-focused account of Mates’s Puzzle is explanatorily out of gas by the second embedding, and does not give us insight into higher-order senses in general and the hierarchy as a totality in all its infinite and libertine glory.

We suggest that these are questions worth pursuing, and that what is needed is to turn to the theory of thought, searching for the structures not in *linguistic semantics* but in *epistemological semantics*[26] to elucidate the true nature of higher-order senses and the hierarchy, and arrive at an account of the *cognitive significance of higher-order senses*. Mates’s Puzzle provides an opening clue for understanding the explanatory basis of the hierarchy by contributing an insight about multiple embedding and ascending up a hierarchical order of propositional attitude attributions. Nonetheless, its standard language-focused formulation fails to actualise this potential to render the connection fully intelligible. What could render the hierarchy fully intelligible, we think, is an extension of a thought-focused interpretation of Frege’s Puzzle. Just what such an extension amounts to is beyond the scope of this paper, but a few words are no doubt in order.

On the thought-focused interpretation, Frege’s Puzzle illustrates how issues about cognitive significance can be understood in terms of the rationality or not of conflicting attitudes to explain how Thoughts are individuated. An extension of this idea would also foreground considerations about cognitive significance and would seek to understand issues about cognitive significance in terms of the rationality or not of conflicting attitudes, specifically, to Thoughts about Thoughts, and to explain the individuation, specifically, of higher-order senses. The extension would pick up on and generalise the role of conflicting attitudes. Call the conflicts in ordinary Frege’s Puzzle cases *intrasubjective factual conflicts*, since they are conflicts within a single thinker about the world. The extension would generalise this, first, to *intersubjective factual conflict*, where thinker*s* take conflicting attitudes to Thoughts about the world, and then second, to *intersubjective interpretive conflicts*, where thinkers take conflicting

[26] Cf. Chalmers (2012, chap. 5) and Rattan (2014).

attitudes to Thoughts about Thoughts, in the form of conflicts about Thoughts about the attribution of propositional attitudes to others. We believe that it is in this nexus of intersubjective conflict, at arbitrary levels in a hierarchical order of propositional attitude attributions, that an account of the nature of higher-order senses is to be found. Mates's Puzzle—but not in its language-focused formulation—is naturally suited to help us articulate this cognitive structure.[27]

Funding No funding was received to assist with the preparation of this manuscript.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Baron, C. (2015). Generalized concept generators. In C. Hammerly, & B. Prickett (eds.), *Proceedings of the 46th annual meeting of the north east Linguistics Society*. GLSA: Amherst

Beaney, M. (1996). Conceptions of analysis in early analytic philosophy. *Acta Analytica, 15*, 97–115.

Boisvert, D. R., & Lubbers, C. M. (2003). Frege's commitment to an infinite hierarchy of senses. *Philosophical Papers, 32*(1), 31–64.

Burge, T. (1978). Belief and synonymy. *The Journal of Philosophy, 75*(3), 119–138.

Burge, T. (1986). Intellectual norms and foundations of mind. *The Journal of Philosophy, 83*(12), 697.

Burge, T. (2005a). Postscript to 'Frege and the hierarchy.' In T. Burge (Ed.), *Truth, thought, reason: Essays on Frege* (pp. 167–210). Oxford University Press.

Burge, T. (2005b). *Truth, thought, reason: Essays on Frege*. Oxford University Press.

Burge, T., & Center, P. D. (1979). Individualism and the Mental. *Midwest Studies in Philosophy, 4*, 73–121.

Caie, M., Goodman, H., & Lederman, J. (2019). Classical opacity. *Philosophy and Phenomenological Research, 48*(4), 631–648.

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

Chalmers, D. (2012). *Constructing the world*. Oxford University Press.

Charlow, S., & Sharvit, Y. (2014). Bound 'de re' pronouns and the LFs of attitude reports. *Semantics and Pragmatics, 7*(3), 1–43.

[27] Many thanks to audiences at the Canadian Philosophical Association Meeting in Montreal in May 2018, the Tilburg Conference on the History of Analytic Philosophy in July 2019, and the Language and Metaphysics Workshop in Banff in February 2020, where versions of this paper were presented. For comments and conversation about this material, we would also like to thank Dominic Alford-Duguid, Josh Armstrong, Herman Cappelen, Josh Dever, David Hunter, Jim Hutchinson, Bernard Katz, David Liebesman, Genoveva Martí, Mohan Matthen, Mark McCullagh, Bernhard Nickel, Roy Sorensen, David Sosa, Nathan Wildman, and Byeong-uk Yi. Thanks as well to anonymous referees for helpful feedback.

Church, A. (1946). Morton G. White. A note on the 'paradox of analysis.' Mind, n.s. vol. 54 (1945), pp. 71–72. - Max Black. The 'paradox of analysis' again: a reply. Mind, n.s. vol. 54 (1945), pp. 272–273. - Morton G. White. Analysis and identity: a rejoinder. Mind, n.s. vol. 54 (1945), pp. 357–361. - Max Black. How can analysis be informative? Philosophy and phenomenological research, vol. 6 no. 4 (1946), pp. 628–631. *Journal of Symbolic Logic, 11*(4), 132–133.

Church, A. (1954). Intensional isomorphism and identity of belief. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, 5*(5), 65–73.

Davidson, D. (1965). Theories of meaning and learnable languages. In Y. Bar-Hillel (Ed.), *Proceedings of the international congress for logic, methodology, and philosophy of science* (pp. 3–17). Amsterdam.

Evans, G. (1982). *The varieties of reference*. Oxford University Press.

Field, H. (2017). Egocentric content. *Noûs, 51*(3), 521–546.

Fine, K. (2009). *Semantic relationism*. Wiley.

Frege, G. (1956). The thought: A logical inquiry. *Mind, 65*(259), 289–311.

Goldberg, S. (2008). Must differences in cognitive value be transparent? *Erkenntnis, 69*(2), 165–187.

Goodman, J., & Lederman, H. (2020). Sense, reference and substitution. *Philosophical Studies, 177*(4), 947–952.

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

Grzankowski, A., & Buchanan, R. (2019). Propositions on the cheap. *Philosophical Studies, 176*(12), 3159–3178.

Heck, R. G. (2014). Intuition and the substitution argument. *Analytic Philosophy, 55*(1), 1–30.

Kripke, S. A. (1979). A puzzle about belief. In A. Margalit (Ed.), *Meaning and use* (pp. 239–283). Springer.

Kripke, S. A. (2008). Frege's theory of sense and reference: Some exegetical notes. *Theoria, 74*(3), 181–218.

Künne, W. (2010). Sense, reference and hybridity. *Dialectica, 64*(4), 529–551.

Langford, C. H. (1942). The notion of analysis in Moore's philosophy. In P. A. Schilpp (Ed.), *The philosophy of G. E. Moore* (pp. 319–342). Northwestern University.

Lederman, H. (2021). Fine-grained semantics for attitude reports. *Semantics and Pragmatics, 14*(1), 1–33.

Longworth, G. (2013). Sharing thoughts about oneself. *Proceedings of the Aristotelian Society, 113*, 57–81.

Mates, B. (1950). Synonymity. In G. Stahl (Ed.), *Meaning and interpretation* (pp. 201–226). University of California Publications in Philosophy.

McCullagh, M. (forthcoming). Interpretative Modesty. *Journal of Philosophy*.

McDowell, J. (1984). De Re Senses. *The Philosophical Quarterly (1950-), 34*(136), 283–294.

McDowell, J. (1986). Singular thought and the extent of 'inner space.' In J. McDowell & P. Pettit (Eds.), *Subject, thought, and context.* Clarendon Press.

McGuinness, B. (1980). *Philosophical and mathematical correspondence of Gottlob Frege*.

Moffett, M. A. (2002). A note on the relationship between Mates' Puzzle and Frege's Puzzle. *Journal of Semantics, 19*(2), 159–166.

Moore, G. E. (1903). *Principia ethica*. Cambridge University Press.

Nelson, M. (2008). Frege and the paradox of analysis. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, 137*(2), 159–181.

Parsons, T. D. (1981). Frege's hierarchies of indirect senses and the paradox of analysis. *Midwest Studies in Philosophy, 6*(1), 37–58.

Parsons, T. (2009). Higher-order senses. In J. Almog & P. Leonardi (Eds.), *The philosophy of David Kaplan* (p. 45). Oxford University Press.

Peacocke, C. (1992). Sense and justification. *Mind, 101*(404), 793–816.

Peacocke, C. (2008). Representing thoughts. In C. Peacocke (Ed.), *Truly understood* (pp. 285–318). Oxford University Press.

Putnam, H. (1954). Synonymity, and the analysis of belief sentences. *Analysis, 14*(5), 114–122.

Putnam, H. (1975). The meaning of 'meaning.' *Language, Mind, and Knowledge Minnesota Studies in the Philosophy of Science, 7*, 131–193.

Rattan, G. (2009). Intellect and concept. *The Baltic International Yearbook of Cognition, Logic and Communication 5*.

Rattan, G. (2014). Epistemological semantics beyond irrationality and conceptual change. *Journal of Philosophy, 111*(12), 667–688.

Rattan, G. (2019). Understanding semantic coordination in cognition. *Dialectica, 73*(3), 289–313.

Rattan, G., & Wikforss, Å. (2017). Is understanding epistemic in nature? *Pacific Philosophical Quarterly, 98*(2), 271.

Sainsbury, R. M., & Tye, M. (2012). *Seven puzzles of thought: And how to solve them: an originalist theory of concepts*. Oxford University Press.
Salmon, N. U. (1986). *Frege's Puzzle*.
Scheffler, I. (1955). On synonymy and indirect discourse. *Philosophy of Science, 22*(1), 39–44.
Schiffer, S. (2003). *The things we mean*. Oxford University Press.
Schwager, M. (2011). Speaking of qualities. *Proceedings of SALT, 19*, 395–412.
Sellars, W. (1955). Putnam on synonymity and belief. *Analysis, xv 5*, 4.
Simchen, O. (2018). The hierarchy of Fregean senses. *Thought: A Journal of Philosophy, 7*(4), 255–261.
Skiba, L. (2015). On indirect sense and reference. *Theoria, 81*(1), 48–81.
Sorensen, R. A. (2002). Meaningless beliefs and Mates's problem. *American Philosophical Quarterly, 39*(2), 169–182.
Soria-Ruiz, A. (forthcoming). On Mates's puzzle. *Mind and Language*.
Stanley, J. (2014). Constructing meanings. *Analysis, 74*(4), 662–676.
Tancredi, C., & Sharvit, Y. (2020). Qualities and translations. *Linguistics and Philosophy, 43*, 303–343.
Wrigley, A. (2006). Abstracting propositions. *Synthese, 151*(2), 157–176.
Yalcin, S. (2015). Quantifying in from a Fregean perspective. *Philosophical Review, 124*(2), 207–253.