

# The Composition of Thoughts

*Richard G Heck Jr and Robert May*

**Brown University and the University of California, Davis**

Among the many innovations that mark Frege's *Begriffsschrift* as a revolutionary work, perhaps the most important is its presentation of the first formal system of logic. Frege believed that the introduction of a new notation, especially for the expression of generality, was necessary if the logical relationships between contents were to be made apparent. This new notation made it possible, at least in an important range of cases, to establish that a given content could be inferred from certain other contents simply by examining the structure of the sentences that expressed those contents.<sup>1</sup> In this regard, Frege's great advance was that, in his system, the logical rules could be stated in purely formal terms, without any reference to the contents expressed by the sentences of his formal language, *Begriffsschrift*, the conceptual notation. But it is important to recognize that logic, according to Frege, is not 'formal' in any sense that would oppose form to content: The sentences of *Begriffsschrift* are not mere forms. Frege's goal was "not... to present an abstract logic in formulas, but to express a content through written symbols in a more precise and perspicuous way than is possible with words" (AimCN, pp. 90–1). Frege's formal system is intended to be one we can actually use in reasoning, that is, in inferring truths from other truths, which is to say that we can prove theorems in this system, where theorems are true contents.<sup>2</sup> If so, the sentences of *Begriffsschrift* must express those contents. The point of presenting proofs in a formal system is thus not to empty mathematics of content (FTA, esp. opp. 97ff.),<sup>3</sup>

---

<sup>1</sup> Much of the discussion in this paper will concern a period during which Frege's views and terminology were in transition, but we shall generally use the term 'content' to refer to the *intuitive* notion that Frege's *technical* notions of 'conceptual content', 'sense', and 'reference' were supposed to capture.

<sup>2</sup> This point is found throughout Frege's writings, both early, when he is comparing his logic to Boole's (BLC, pp. 12–13), and later, in his debate with Hilbert (Geo1; Geo2). It is rightly echoed in van Heijenoort's (1967) insistence that, for Frege, logic is a language and not just a calculus.

<sup>3</sup> Page references marked by "op." and "opp." are to the pages in the original publication. These

nor even to make us more certain of what we are proving (Gl, §2), but rather to force every assumption on which a given proof depends to be made explicit, so that the grounds on which the result rests can be reliably ascertained (Bg, preface; Gg, v. I, p. vii; PCN, pp. 362–3). This is what Frege intends by his insistence that proofs must be rigorous and ‘gap-free’. Only then can we be certain that we have shown how a truth follows from other truths.

Given Frege’s perspective, what makes formalization possible is thus the fact that, in *Begriffsschrift*, the logical properties of contents are made visible in the signs that express those contents. And that, Frege seems to have understood from the start, is because these signs—the sentences of *Begriffsschrift*—are composed of parts that have their own contents, contents that together determine the content of the sentence. The expression ‘ $2^2 = 4$ ’, for example, occurs in the expression ‘ $2^2 = 4 \rightarrow 2^4 = 16$ ’, and, as a part of the latter sentence, it has the very same content it has when it occurs on its own.<sup>4</sup> Of course, Frege was hardly the first to gesture in the direction of the principle of compositionality. But what distinguished Frege from his predecessors—and what made it possible for Frege to develop the first compositional semantics for a language of reasonable expressive power—is that he understood what the really hard question was: How *exactly* are the contents of the parts supposed to combine to yield the content of the whole?<sup>5</sup> This paper tells the story of Frege’s struggle with that question.

There are two familiar models for understanding this sort of composition. The first is the composition of a sentence from its parts. A sentence consists of its parts in a straightforward sense: The concatenation of a noun phrase and a verb phrase simply *is* a sentence, just as a top set upon four legs simply is a table; the table is not something other than the top set upon the legs, and the sentence is not something other than the noun phrase concatenated with the verb phrase.<sup>6</sup> The second model is the application of a function to an argument. The result of applying a function to an argument is not (normally) a complex consisting of the function and the argument. It is, rather, the value of the function for that argument.

---

are indicated in the margins of the *Collected Papers* (Frege, 1984) and of the *Frege Reader* (Frege, 1997).

<sup>4</sup> This is Geach’s (1965) famous ‘Frege point’.

<sup>5</sup> Frege’s appreciation of this fact was also a consequence of his new understanding of generality: Despite the surface similarities, the way the parts are combined in a sentence like ‘Socrates is mortal’ is very different from how they are combined in a sentence like ‘All men are mortal’. (See PMC, pp. 100–1, for example.)

<sup>6</sup> The question what exactly the relationship is between a table and its parts, being a philosophical question, is of course controversial. But a sentence wholly consists of its parts in whatever sense a table wholly consists of its parts.

This model of composition is the one Frege ultimately applies in the realm of reference: A concept is a function from objects to truth-values, and the reference of a (non-embedded) sentence is its truth-value; “Frege had a beard” refers to the truth-value that is the result of applying the ‘concept-function’ denoted by the predicate to the object denoted by the subject. But it is obvious that, for the mature Frege, this cannot be a complete answer to the question how the content of a sentence is determined from the contents of its parts, since we thereby explain only how the reference of the whole is determined by the references of the parts. A complete answer must also explain how the *sense* of the whole is determined by the senses of the parts. And Frege is much less explicit about how thoughts are composed from senses, even though the answer is essential to his enterprise: Without it, we will not understand how formal relations among expressions might reflect logical relations among their contents. Logical relations do not hold among contents simply in virtue of facts about reference. This, as we shall see, is one of the many lessons of Frege’s famous puzzle about substitution.

We shall be arguing that Frege’s account of how language expresses content can fulfill the explanatory ambitions he has for it only if thoughts are composed in a manner akin to the first model: A thought is constituted from its constituent senses much in the way a sentence is constituted from its constituent phrases. As Frege puts it, “To the structure of the thought there corresponds the compounding of words into sentences” (Neg, p. 378, op. 148) A thought in this regard is a complex. The central questions we want to address here are why Frege is committed to this view and what its implications are.

The question whether thoughts are compositionally complex has been much discussed,<sup>7</sup> but it remains controversial, although the textual case for an affirmative answer is *prima facie* quite strong. In *Grundgesetze*, for example, Frege expresses the view that thoughts have parts quite directly: “If a name is part of a name of a truth-value [that is, a sentence], then the sense of the former name is part of the thought expressed by the latter name” (Gg, v. I, §32). That said, however, Frege suggests in “On Sense and Reference” that we can also say that the reference of a part of a complex expression is part of the reference of the whole expression (SM, p. 165, opp. 35–6), but this is an absurd suggestion given Frege’s other commitments. The reference of ‘Annette’ is not part of the reference of ‘the father of Annette’. One might therefore be tempted to take neither suggestion seriously. But Frege qualifies the suggestion that references are compositionally

---

<sup>7</sup> The contributors include Hans Sluga (1980), Steven Wagner (1983), David Bell (1987), Michael Dummett (1991b), and James Levine (2002), among many others.

complex when he makes it—he says that he is using “the word ‘part’ in a special sense” and remarks that, really, “[a] special term would need to be invented” (SM, p. 165, opp. 35–6)—and he later disowns this suggestion even while reaffirming his commitment to the view that thoughts are compositionally complex:

We can regard a sentence as a mapping of a thought: corresponding to the whole–part relation of a thought and its parts we have, by and large, the same relation for the sentence and its parts. Things are different in the domain of reference. We cannot say that Sweden is a part of the capital of Sweden. (Darm, p. 255)

And, we find the following remarks in Carnap’s notes on Frege’s lectures of 1913:<sup>8</sup>

The *reference* of the parts of a sentence are not parts of the reference of the sentence.

However: The sense of a part of the sentence is part of the sense of the sentence. (Reck and Awodey, 2004, p. 87)

Such passages are undeniably suggestive, but we cannot expect to resolve the question whether Frege regarded thoughts as compositionally complex simply by exhibiting proof-texts. Unless the claim that thoughts are compositionally complex can be shown to play a significant role in Frege’s considered views, remarks like the one just quoted can be dismissed as mere slips. So we shall need to consider what role this doctrine might play.

We will proceed as follows. We shall begin, in section 1, by discussing Frege’s earliest views on these issues, arguing that, although Frege did regard conceptual contents as inherently unstructured in *Begriffsschrift*, he began almost immediately thereafter to move away from that view and towards one on which conceptual contents are composed of parts corresponding to ‘function’ and ‘argument’. This change was connected with a broader change in Frege’s view of language, which then took on a strictly representational role: Linguistic terms designate elements of content, what Frege calls “objective ideas” in *Die Grundlagen*. But Frege’s initial way of understanding the relation between linguistic items and their contents turns out to be unstable. This is the most immediate lesson of Frege’s famous puzzle about the substitution of co-referential expressions, as we will bring out in section 2. We argue there that this puzzle forced Frege to abandon his earlier conception of how the conceptual content of a sentence was determined by

---

<sup>8</sup> We have changed the translation of ‘*Bedeutung*’ from ‘meaning’ to ‘reference’.

the contents of its parts and, in doing so, posed in an especially sharp form the question why “The evening star is a planet” and “The morning star is a planet” have different contents. We consider Frege’s response to this question in section 3, arguing that it depends essentially upon the claim that the senses of the proper names that occur in the two sentences just mentioned are parts of the thoughts those sentences express. However natural this view may seem, however, a proper development of it requires an answer to the question how thoughts are composed from senses, that is, to the question what in the realm of sense corresponds to the syntactic operation of concatenation. We turn to this question in section 4, arguing that it should be understood not as a metaphysical question but as a semantic one: The question concerns not the intrinsic nature of residents of the ‘third realm’ but how thoughts are expressed by sentences and how thoughts determine truth-values.

## 1 Concepts and Conceptual Articulation

Throughout Frege’s thinking about content, there is an underlying theme that remains invariant: that we have no direct cognitive access to content.<sup>9</sup> The reason for this, however, has nothing to do with whether we can apprehend objects, *qua* elements of content. Rather, the problem is that mere apprehension of entities is not tantamount to having knowledge of them; for instance, we have, according to Frege, intuitive access to geometrical entities, but that, in and of itself, is not sufficient for any substantial knowledge about space. What makes knowledge possible is our ability to apprehend something else about content, its logical structure; only through our apprehension of this structure can we know of the properties and relations that hold of any objects that we may apprehend, and hence what is true of them. But we have no direct cognitive grip on this structure and hence no direct grip on content itself. Rather, there must be some mediating system, to which we can be cognitively related, which provides, in some relevant sense, the requisite access to structure. Thus, at base, the issue for Frege begins, and remains throughout, the nature of the mediating system through which we can grasp the structure of content. What varies in Frege’s thought is his conception of the nature of the system that displays the structure of content for us; the shift we encounter is from a system that is itself external to content, through which structure is “viewed” from without, to one that is internal to content, structuring it from within.

As noted, Frege was interested in the correlations his logic revealed between

---

<sup>9</sup> With one exception. We do have direct cognitive access to senses; we grasp them.

formal relations and logical relations. There is nothing terribly original about the view that logic is formal in some such sense. What was original was Frege's conception of how sentences and the contents they express are structured. In logic as it existed prior to Frege, sentences, or judgements,<sup>10</sup> were constructed from predicates using a small number of operators corresponding to traditional forms of judgment, such as universal affirmative judgments of the form 'All *F*s are *G*'. Proper names, such as 'Socrates', were regarded as predicates, that is, as being of the same logical type as expressions like 'is mortal'. Frege came to believe, however, that this traditional identification of names as predicates is incompatible with the proper representation of generality—in particular, with the representation of scope—and what Frege himself advertises as the cornerstone of his advance in *Begriffsschrift* is his insistence that sentences must be regarded as constructed from *arguments* and *functions*. For example, in the sentence 'Socrates is mortal', if we take the argument to be 'Socrates', then the function is "the part that remains invariant in the expression" when we replace 'Socrates' by other names, such as 'Plato' or 'Thales' (Bg, §9).

There is a lineage, an important one, between the function–argument distinction as Frege initially draws it and his later distinction between concept and object. But there are also important discontinuities, centered on Frege's insistence in *Begriffsschrift* that the distinction between function and argument "has nothing to do with the conceptual content, but only with our way of viewing it" (BgBynum, §9).<sup>11</sup> What are we to make of this remark? First off, it appears that Frege is being explicit that the notions of function and argument are not to be taken as internal aspects of content itself, but rather as external to it. But Frege is also saying that function and argument are the tools for analyzing content; it is how we "view" content. Thus, an analysis in terms of function and argument, while not part of content, is nonetheless essential if content is to be represented in such a way that it can be comprehended as having logical or, as Frege puts it in *Begriffsschrift*, conceptual structure. Accordingly, it is natural for Frege to say, in describing the proper application of the terms 'function' and 'argument' that:

---

<sup>10</sup> Prior to Frege, use and mention were not carefully distinguished; indeed, Frege can be accused of being none too observant himself, especially in his earlier writings. But see the discussion at the end of this section.

<sup>11</sup> That is Bynum's translation. van Heijenoort reads it: "... has nothing to do with the conceptual content; it comes about only because we view the expression in a particular way". That just seems wrong, because what we are viewing is not the expression but the content. The German is: "Diese Unterscheidung hat mit dem begrifflichen Inhalte nichts zu thun, sondern ist allein Sache der Auffassung."

If, in an expression (whose content need not be assertable), a simple or a complex symbol occurs in one or more places and we imagine it as replaceable by another [symbol] (but the same one each time) at all or some of these places, then we call the part of the expression that shows itself invariant [under such replacement] a function and the replaceable part its argument. (Bg, §9)

Frege's point here is that functions and arguments are to be understood as *expressions* of the conceptual notation, that is, as linguistic symbols. The function of Frege's 'content stroke', he tells us, is to "tie[] the symbols [for functions and arguments] which follow it into a whole". If we suppose that what an expression so composed represents is a particular formal analysis of content—function and argument being the analytic terms provided by Begriffsschrift for a structural analysis of judgeable content—then content can be depicted as structurally divided between function and argument when 'viewed' through the lens of language, a language, that is, with the expressive resources of the conceptual notation.<sup>12</sup> Thus, for Frege, the issue is not whether the content *Socrates is mortal* divides itself into function and argument in any determinate way. Rather, it is that one could regard 'Socrates' as the argument and 'is mortal' as the function, but one could equally well regard 'is mortal' as the argument and 'Socrates' as the function.<sup>13,14</sup> In *Begriffsschrift*, the issue is not whether content has an inherent or intrinsic structure

---

<sup>12</sup> So the shortcoming of natural languages, for Frege, is that they do not permit content to be logically viewed or regarded, at least not to the standard of rigor and precision that Frege requires for the business at hand.

<sup>13</sup> This is so, according to Frege, as long as the terms both have 'determinate' meanings, in the sense that the meaning of one does not depend upon the meaning of the other. This is true of both 'Socrates' and 'is mortal' but not of 'all men'. Rather, terms like these are indeterminate in and of themselves; their meanings depend upon that of other (determinate) terms. Accordingly, 'All men are mortal' will inherently have a conceptual organization that reflects this relation, so that "the whole splits up into *function* and *argument* according to its own content, and not just according to our way of looking at it" (BgBynum, §9). This seems much like what Frege would later say about 'Socrates is mortal': 'Socrates' is a proper name, and refers to an object, 'ξ is mortal' is a predicate, and refers to a function (a concept), and how we regard them is irrelevant to this classification. See below for further discussion.

<sup>14</sup> Afficionados will note that we did *not* just say that one can regard 'ξ is mortal' as the argument and the second-level concept 'Φ(Socrates)' as the function, which is how Frege would have seen the matter in his later work. The reason is that there is no notion of 'second-level concept' to be found in *Begriffsschrift*. And similarly, there is no clear distinction between first- and second-order quantification. There is but one axiom of universal instantiation, proposition 58, and it is used indiscriminately to justify both what we would regard as first-order inferences and what we would regard as second-order inferences.

and, if not, what might be its source. Rather what is at issue is the analytic priority of language for logically comprehending that structure, regardless of whether it is exposed or imposed by language; it is only through our grasp of *Begriffsschrift* that we have access to the logical structure of content. Language, then, has a very special role in Frege's earliest work: It is the locus of the structural relation of function and argument, the structure we must view content as having if we are to have the logical wherewithal to show how true contents follow from other true contents.

Thus, for Frege at the time of *Begriffsschrift*, the prime role of language was to reveal the conceptual organization of content. Although external to content, language is nevertheless the locus of compositional complexity vis-à-vis content. But whatever merits Frege may have perceived in seeing things this way, there was a problem that could not have escaped his notice for long. If function and argument are linguistic notions, then the complex that results from their composition—what Frege designates by the addition of the content stroke—is also linguistic; what we have is a linguistic complex whose constituent parts are linguistic expressions, including a 'function' and an 'argument'. What the combination of function and argument yields is therefore just a sentence. But if so, then it is hard to see how function–argument analysis helps us understand how the formal properties of sentences of *Begriffsschrift* might reflect the logical properties of their contents; no connection has yet been established between the function–argument analysis of a given sentence and any features at all of its content. But if so, then there is a large semantic lacuna in the exposition of *Begriffsschrift*: We have been given no particular reason to suppose that the formal manipulations perform as advertised, that is, not only that they are rigorous and gap-free, but that they take us from *truths* to *truths*. This lacuna is one that Frege almost immediately sets out to fill.

The initial step in Frege's re-orientation occurred shortly after the publication of *Begriffsschrift*. Almost everyone who reviewed the book criticized Frege for his not having discussed Boolean logic,<sup>15</sup> and Frege therefore wrote a paper in 1881, "Boole's Logical Calculus and the *Begriffsschrift*", that was intended to explain the motivation behind his new approach.<sup>16</sup> Frege there makes the following

---

<sup>15</sup> Hardest on Frege in this regard was John Venn. Ernst Schroeder, the leading German Boolean logician, was rather more sympathetic, if still uncomprehending of Frege's accomplishments. It appears that they were right to take Frege to task, as apparently at the time of writing *Begriffsschrift*, Frege was not familiar with Boole. The reviews are collected in the Bynum edition of *Begriffsschrift* (BgBynum). Risto Vilkkio (1998) has an insightful discussion of their content.

<sup>16</sup> Although this paper was not published during Frege's lifetime, he submitted it for publication on three separate occasions—see the editors' notes—so we may regard it as representing Frege's



remarks:

If... you imagine the 2 in the content of possible judgement  $2^4 = 16$  to be replaceable by something else, by  $-2$  or by 3 say, which may be indicated by putting an  $x$  in place of the 2:

$$x^4 = 16,$$

the content of possible judgement is thus split into a constant and a variable part. The former, regarded in its own right but holding a place open for the latter, gives the concept ‘4<sup>th</sup> root of 16’. ... And so instead of putting a judgement together out of an individual as subject and an already formed concept as predicate, we do the opposite and arrive at a concept by splitting up the content of possible judgement. (BLC, pp. 16–17)

When Frege remarks that a concept is to be “regarded in its own right but *holding a place open for*” a suitable argument, this is highly suggestive of a view very different from that of *Begriffsschrift*. While Frege does not use the words “incomplete” and “unsaturated” until the following year (PMC, p. 101), saying that a concept has a place in it that is held open for an object is obviously very much in the same spirit. And if it is to make any sense at all to speak of replacing the *object* 2 with other *objects* in the content of possible judgement  $2^4 = 16$ , then the object 2 must itself occur in that content—it must somehow be a part of it—as must what remains constant when this part is varied; that is to say, the content itself must be articulated as concept and object.<sup>17</sup> If so, then there has been a significant change in Frege’s view, for now there is a structure inherent to content, that of concept and object. The very distinction between these introduces the connection to truth: “In the case of a concept, it is always possible to ask whether something, and if so what, falls under it, questions which are senseless in the case of an individual” (BLC, p. 18). It is but a quick step from “falling under” to truth.

One might be tempted to object at this point that there is really no shift in Frege’s view here, but that he is simply confusing use and mention. In *Begriffsschrift*, so the objection would go, Frege speaks as if functions and arguments are

---

considered view at the time.

<sup>17</sup> Frege does not commit himself here on the question whether there is a *unique best* articulation of the content. For discussion, see the exchange between Bell (1987) and Dummett (1991b) and the more recent contributions by Rumfitt (1994), Levine (2002) and Textor (2009), and see note 71 for some remarks on the issue.

(parts of) expressions: When we replace the name ‘Socrates’ with other names in the expression ‘Socrates is mortal’, the function we discover is “the part that remains invariant in the *expression*” (Bg, §9, our emphasis). But, the objection might continue, we need not conclude from this that Frege’s considered view was that functions were linguistic expressions: It would be enough to suppose that, in 1879, he was no clearer about the distinction between use and mention than were his contemporaries. And if so, Frege’s remarks in “Boole’s Logical Calculus” about replacing the object 2 with other objects need not be taken to indicate any change in his views about conceptual content. While Frege does, that is to say, speak of replacing the *object* 2 with other *objects* in the *content of possible judgement*  $2^4 = 16$ , perhaps this is just loose talk; he really means to be speaking, as he did in *Begriffsschrift*, of replacing the *expression* ‘2’ by other *expressions* in the *sentence* ‘ $2^4 = 16$ ’.

It is no doubt true that, by contemporary standards, Frege does tend to run use together with mention, but he is hardly unaware of the distinction. His emphatic claim that the function–argument distinction is *not* a distinction in content shows his awareness of it. His view that the function–argument distinction is one drawn within language and not within content makes sense only if use is being distinguished from mention—if content is being distinguished from language. Indeed, during the period we are discussing, it becomes increasingly important to Frege clearly to distinguish between content and language, between content structured in its own right, and how language represents that structure. We can see this emerge in the continuation of the remarks quoted above from “Boole’s Logical Calculus”:

... [I]nstead of putting a judgement together out of an individual as subject and an already previously formed concept as predicate, we do the opposite and arrive at a concept by splitting up the content of possible judgement. Of course, if the expression of the content of possible judgement is to be analysable in this way, it must already be itself articulated. We may infer from this that at least the properties and relations which are not further analysable must have their own simple designations. But it doesn’t follow from this that the ideas of these properties and relations are formed apart from objects: on the contrary, they arise simultaneously with the first judgement in which they are ascribed to things. Hence, in *Begriffsschrift*, their designations never occur on their own, but always in combinations that express possible contents of judgement. ... A sign for a property never appears without a thing to which it might belong being at least in-

licated, a designation of a relation never without indication of the things which might stand in it. (BLC, p. 17)

Here Frege is quite carefully distinguishing between properties and relations, on the one hand, and our representations of them, on the other. Indeed, Frege seems to be arguing here that ‘signs for properties’ and ‘designations of relations’ must always occur in complexes, along with appropriate expressions in their argument-places, *because* our ideas of what they designate are not “formed apart from objects” but only through their combination, in judgment.<sup>18</sup>

Beginning about 1881, then, and essentially continuing through *Grundlagen*, Frege takes on a conception of the relation between language and content that is opposed to his initial position in *Begriffsschrift*. As the view that the distinction between function and argument has “nothing to do with the conceptual content” fades away, we find a new view emerging, according to which language represents a structure already inherent in content, a structure characterized by the distinction between concept and object. And this change, we have argued, is well-motivated: It is needed if Frege is to align the formal notion of structure with the logical one, which he must, since only then can the structural articulation of sentences serve to reveal the logical properties of their contents. As we shall now see, however, this new view is incompatible with other assumptions central to Frege's thought.

## 2 Frege's Puzzle About Substitution

Readers no doubt will have noticed a terminological change during the course of the previous section. Whereas we began by considering functions, we ended by discussing concepts. To a large extent, this terminological shift highlights a more substantive one due to Frege's sharpening his conception of the relation between language and content in the two or three years following the publication of *Begriffsschrift*. As we have seen, concepts, on Frege's emended view, are ‘unsaturated’ parts of content of a sort fundamentally different from objects. But they are also what function–argument analysis reveals to us; the analysis of a content into function and argument is now clearly intended to reflect structural features inherent within content itself. Frege's identification of concepts as functions is simply the result of this shift, and it is supposed to reveal something fundamental about the role concepts play in logic, that is, in the logical articulation of content. It is from

---

<sup>18</sup> As we shall see, the order of explanation is reversed in Frege's mature philosophy.

the logical perspective that concepts are to be characterized as functions; logically speaking, concepts *are* functions.

But the identification raises an odd problem. Not all functions are concepts. For example, the addition function is not a concept.<sup>19</sup> So there must be something that distinguishes concept-functions, as a class, from ordinary functions. The natural suggestion is that this difference is to be found in their arguments and values, in what constitutes the domain and range of these functions. We need not dwell on the question what the arguments of concept-functions are supposed to be: They are objects, the contents of names.<sup>20</sup> It is not so obvious, however, what the *values* of concept-functions were supposed to be. So consider a simple sentence of the form  $Fa$ . One would have supposed that applying the concept-function that is the content of  $F$  to the object that is the content of  $a$  should yield the content of the complex expression that they together constitute. Otherwise, it is hard to see how the analysis of that content into function and argument is supposed to help us understand how the content is structured. So Frege's view at this time must have been the following: A proper name has as its content an object; a predicate has as its content a function from objects to conceptual contents; and the conceptual content of a sentence of the form  $Fa$ , which is a content of possible judgement, is the result of applying the function that is the content of  $F$  to the object that is the content of  $a$ .<sup>21</sup>

But this position is inconsistent with other of Frege's commitments. If the content of a sentence  $Fa$  is the result of applying the function that is the content of  $F$  to the object that is the content of  $a$ , so that the content of:

(E)           The Evening Star is a planet

is the result of applying the function that is the content of 'ξ is a planet' to the object that is the content of 'the Evening Star', then its content must be the same as that of:

(M)           The Morning Star is a planet.

<sup>19</sup> We ignore the concept *horse* problem here and limit our attention to what Frege would later call first-level concepts.

<sup>20</sup> Frege held already in *Begriffsschrift* that the content of a proper name is the object it denotes (Bg, §8).

<sup>21</sup> Our sense is that most people who have considered this question have come to the same conclusion we have. See, for example, the recent discussion by Michael Beaney (2007). We shall see some additional evidence in favor of this interpretation below. See note 40.

And this conclusion would have been unacceptable to Frege.<sup>22</sup>

To see why, consider the following remarks from *Begriffsschrift*:

... [T]he contents of two judgements may differ in two ways: either the consequences derivable from the first, when it is combined with certain other judgements, always follow from the second, when it is combined with these same judgements, or this is not the case. The two propositions, "The Greeks defeated the Persians at Plataea" and "The Persians were defeated by the Greeks at Plataea" differ in the first way. ... Now I call that part of the content that is the *same* in both the *conceptual content*. ... Everything necessary for a correct inference is expressed in full, but what is not necessary is generally not indicated; *nothing is left to guesswork*. (Bg, §3, emphasis in original)

Some commentators have found in this passage the claim that, if *A* and *B* have the same consequences, then they must also have the same content; it is a short step from this to an inferentialist conception of content. We don't read the passage that way, but we need not enter this controversy here. A weaker thesis is clearly enough present: If two judgements have *different* consequences, then they have *different* conceptual contents. But it is obvious that (E) and (M) have different logical consequences, since it is certainly no truth of *logic* that the Evening Star is the Morning Star.<sup>23</sup> Hence, (E) and (M) must have different conceptual contents.<sup>24</sup>

How, then, can Frege avoid the conclusion that (E) and (M) must have the same content? He cannot give up the thesis that the content of a proper name is its bearer; that is central to his understanding of identity as objectual.<sup>25</sup> He cannot

<sup>22</sup> Thomas Ricketts (2005, pp. 23–4) reads this part of Frege's argument in much the same way we do.

<sup>23</sup> The contrary thesis has been held (Marcus, 1995). But that was before Kripke (1980) showed us the differences between logical, metaphysical, and epistemic necessity.

<sup>24</sup> The point Frege is making in the quoted passage seems badly expressed. The underlying idea is that conceptual content is what matters to logic: logical differences between judgements imply differences in conceptual content. But even if  $a = b$  is a truth of logic, it is still logically different from  $a = a$ . Perhaps the right thing to do here would be to introduce some notion of 'immediate' consequence and say that judgements have the same conceptual content if they have the same 'immediate' consequences. But Frege never pursues the matter, in large part, we suspect, because his notion of content soon changes so dramatically that the question is no longer important. That said, Bob Hale (2001a; 2001b) has worked hard to reconstruct some such notion. Unfortunately, the attempt has serious problems (Potter and Smiley, 2001, 2002).

<sup>25</sup> Contrary to what often seems to be assumed, the transition from the earlier view that identity is a relation between names to the later view that it is a relation between objects occurs well before

give up the thesis that concepts are functions;<sup>26</sup> that is key to his understanding of logical generality. Nor can he give up the principle that the content of a simple sentence like (E) or (M) is fully determined by the contents of its parts, since that is just an aspect of his treatment of concepts as functions: If predicates denote functions, then the value of any such function will be determined by the value of its argument, just as the value of the function  $\xi^2$  is determined by the value of its argument. But if all of these are non-negotiable, then there is an antinomy within Frege's developing philosophy of logic. Something else must go, and there is only one thing that can go: The values of concept-functions *cannot* be conceptual contents. More precisely, the conceptual content of  $Fa$  cannot be the result of applying the concept-function that is the content of  $F$  to the object that is the content of  $a$ .

The immediate question that faces Frege, then, is what the values of concept-functions are, if they are not conceptual contents. The key insight leading to the answer is already in place by 1882, when, in the remark cited above, Frege insists that what distinguishes concepts from objects is that a concept is the kind of thing under which an object does or does not fall (PMC, p. 101). If this is what is distinctive of concepts, then this is what concept-functions must capture if they are to be worthy of the name. What Frege eventually realized was that it is therefore sufficient to take the values of concept-functions to be *truth-values*, truth and falsity corresponding to whether an object falls under a concept or not. This of course was Frege's mature view. As he puts it in *Grundgesetze*, "...it seems appropriate straightforwardly to call a *concept* a function whose value is always a truth-value" (Gg, v. I, §3). But the change occurred considerably earlier.

This new treatment of concepts as characteristic functions was deeply embedded in Frege's thinking no later than 1884.<sup>27</sup> The evidence is as follows. In "Boole's Logical Calculus", which dates from 1881, Frege explicitly rejects the view that concepts are extensional. Indeed, he does not just insist that concepts are intensional: He uses the claim that there can be more than one concept true

---

the publication of "On Sense and Reference" in 1891. As May (2001) notes, the change has clearly occurred already by *Die Grundlagen*, and there is strong evidence, in fact, that it is already in place by 1881 (Heck, 2013; Heck and May, 2006).

<sup>26</sup> Note that the argument requires only the claim that the meaning of the predicate *determines* a function from objects to conceptual contents; it does not require in fact the claim that the meaning of a predicate is such a function. See also note 34.

<sup>27</sup> We have discussed the reasons Frege came to this view elsewhere (Heck and May, 2006, §3). They are connected with Frege's discovery of truth-functions, which probably emerged as a result of his reading of Boole. This connection reinforces the argument we are about to give.

just of the Earth as a premise in an argument for the important thesis that concepts must be distinguished from objects, in particular, that we must distinguish concepts true of just the Earth from the Earth itself (BLC, p. 18). Three years later, however, in *Die Grundlagen*, Frege says, in an infamous footnote, that he is “convinced” that the view that concepts are extensional can be defended (Gl, §68, fn). He does not there indicate *how* he thinks the extensional view can be defended, but surely Frege would not have abandoned the much more natural view that concepts are *intensional* unless he had acquired theoretical commitments that conflicted with it. We know of no other view of Frege’s that conflicts with it than the view that concepts are characteristic functions. So Frege must have held, already in 1884, that concepts are functions from objects to truth-values.

One might object that there is another explanation for Frege’s change of view, namely, that he should have adopted an extensional view of *functions* and that this is what led him to an extensional view of *concepts*. And, indeed, it is clear that the view that concepts are characteristic functions does not by itself entail that concepts are extensional, since functions themselves might be intensional. But the view that functions are extensional does not entail that concepts are extensional, either. For suppose that the conceptual contents *Fred has a heart* and *Fred has a kidney* are distinct, as one would naturally suppose. Then the concept-functions *has a heart* and *has a kidney* map Fred to different conceptual contents, and so these functions are distinct, even if they are extensional. The right conclusion to draw, then, is that Frege’s abandonment of the intensional view of concepts makes sense only if he has committed himself *both* to an extensional view of functions *and* to the treatment of concepts as characteristic functions. It is the latter commitment that matters here, though we are happy to agree that these same considerations show that the view that functions are extensional was also in place by 1884. Whether Frege had always regarded functions as extensional, or whether this was another shift in his views, is an issue we need not decide here.<sup>28,29</sup>

---

<sup>28</sup> For what it is worth, we think the story is actually pretty straightforward: So long as Frege was confusing functions with the expressions that denote them, an intensional treatment would have been natural; once he had clearly distinguished these, the extensional conception, which is the one that is true to how mathematicians speak of functions, could emerge.

<sup>29</sup> Here is another objection. On the basis of evidence from the Frege *Nachlass*, Sundholm argues that “Frege did not have the doctrine of (objectual) truth-values” in 1889 (Sundholm, 2001, p. 62). And yet we are claiming that Frege held already by 1884 that truth-values were the values of concept-functions. But we are *not* claiming that Frege already held in 1884 that truth-values are *objects*. As Dummett notes (1991a, ch. 12), the ontological thesis that truth-values are objects is bound up with the syntactic thesis that sentences are proper names. So it may be that Frege had not arrived at this view by 1884, either. And there is reason to suppose he hadn’t. There are

It is a major virtue of Frege's new view that concepts are characteristic functions that he can now at least start to make good on the critical claim left undefended in *Begriffsschrift*: that logic, *his* logic, allows for the derivation of truths from other truths, not just of formal symbols from other formal symbols. Truth, that is to say, does at least enter the picture now. But the new view has problems of its own; it seems not so much to resolve the problem posed by the substitution puzzle as to make matters worse. On Frege's earlier view, the content of a name is an object; the content of a predicate is a function; and the content of a sentence is the result of applying the function to the object. The new view no longer treats concept-functions as mapping objects to conceptual contents; instead, concept-functions map objects to truth-values. But the general architecture of the view is unchanged, so much so that it might seem as if what has really changed is Frege's conception of what the contents of sentences are. It might seem, in particular, that his new view is simply that the content of a sentence is its *truth-value*. But then it would not just be pairs like "The Evening Star is a planet" and "The Morning Star is a plant" that had the same content. *All* true sentences would have the same content.<sup>30</sup>

That would be a misunderstanding, to be sure, but we should not dismiss it too quickly.<sup>31</sup> It is, in fact, a misunderstanding Frege discusses explicitly in an important passage from *Function and Concept*, a passage in which he is introducing his view that the 'meaning' of a sentence is its truth-value:

. . . What ' $2^2 = 4$ ' means is the True just as, say, ' $2^2$ ' means 4. And ' $2^2 = 1$ ' means the False. Accordingly, ' $2^2 = 4$ ', ' $2 > 1$ ', and ' $2^4 = 4^2$ ' all mean the same thing, viz. the True. . . . The objection here suggests itself that ' $2^2 = 4$ ' and ' $2 > 1$ ' nevertheless tell us quite different things, express quite different thoughts. (FC, op. 13)

significant differences between his statements of the context principle in *Die Grundlagen* (Gl, p. x, §62) and in *Grundgesetze* (Gg, v. I, §29), differences that seem to reflect a change from a view on which sentences have a privileged position to one on which they are but a kind of proper name. This has been the subject of much controversy, so we would not want to rest too much upon it. But it is reasonable to suppose that it took Frege some time fully to adjust his conception of logic to his new discovery that truth-values are the values of concept-functions. It seems, in fact, to have taken several years, as a satisfying new philosophy of logic does not seem to have emerged until sometime in 1889 or 1890, as Sundholm (2001) makes clear. That, we would suggest, is what accounts for Frege's long silence between 1885 and 1891.

<sup>30</sup> And all false sentences, too, though one different from that of the true sentences.

<sup>31</sup> Something like this misunderstanding is what underlies Thau and Caplan's (2001) argument that Frege never abandoned the view that identity is a relation between names. We have discussed their arguments elsewhere (Heck, 2003; May, 2001).



We are here translating Frege's German term '*bedeuten*' (and its cognates) using the ordinary English equivalent, 'mean', as one simply cannot appreciate the force of the objection Frege is considering if one translates this difficult term any other way.<sup>32</sup> But when we do so translate it, we can see that, in saying that sentences 'mean' their truth-values, Frege seems simply to be inviting the misunderstanding we just mentioned. Sentences cannot 'mean' their truth-values, he has his interlocutor say, because otherwise all true sentences would have to "tell us" the same thing, which they plainly do not.<sup>33</sup> So much the worse, then, for the view that the 'meaning' of a sentence is its truth-value! But Frege draws a different conclusion:

... [L]ikewise ' $2^4 = 4^2$ ' and ' $4 \times 4 = 4^2$ ' express different thoughts; and yet we can replace ' $2^4$ ' by ' $4 \times 4$ ', since both signs have the same meaning. Consequently, ' $2^4 = 4^2$ ' and ' $4 \times 4 = 4^2$ ' likewise have the same meaning. We see from this that from sameness of meaning there does not follow sameness of the thought expressed. If we say 'the Evening Star is a planet with a shorter period of revolution than the Earth', the thought we express is other than in the sentence 'the Morning Star is a planet with a shorter period of revolution than the Earth'; for somebody who does not know that the Morning Star is the Evening Star might regard the one as true and the other as false. And yet both sentences must have the same meaning; for it is just a matter of interchange of the words 'the Evening Star' and 'the Morning Star', which mean the same thing, i.e., are proper names of the same heavenly body. (FC, opp. 13–4)

What Frege is doing here is simply applying the lesson he had learned from the substitution puzzle: We cannot treat concept-functions as functions from objects to the contents of sentences. The argument is straightforward: If we assume that 'meaning' is compositional, so that the 'meaning' of a complex expression is determined by the 'meanings' of its parts, and if we assume further that the 'meaning' of a proper name is its bearer and that the 'meaning' of a sentence  $Fa$  is a function<sup>34</sup> of the bearer of  $a$ , then it follows immediately that ' $2^4 = 4^2$ ' and

<sup>32</sup> In our own text, we shall use scare-quotes around the term "meaning" so long as we are using it in the sense of *Bedeutung*.

<sup>33</sup> It is here that the translation in terms of 'reference' fails to convey the force of the objection. The fact that ' $2^2 = 4$ ' and ' $2 > 1$ ' refer to the same thing does not naturally suggest the objection that they "nevertheless tell us quite different things".

<sup>34</sup> Frege of course held that the 'meaning' of  $F$  was such a function, and so that the 'meaning' of  $Fa$  was the result of applying that function to the bearer of  $a$ . But the puzzle about substitution

' $4 \times 4 = 4^2$ ' must have the same 'meaning'. But this shows no more, according to Frege, than that sameness of 'meaning' does not imply sameness of content; sentences can have the same 'meaning' even if they have different contents.<sup>35</sup> And this, in turn, straightforwardly implies that the content of a sentence  $Fa$  cannot be the result of applying the concept-function that is the content of  $F$  to the object that is the content of  $a$ . The values of concept-functions, if they are to be the 'meanings' of sentences, *cannot* also be their contents.

Why does this matter? It matters because function–argument analysis was supposed to be our window onto the structural features of contents that are responsible for their logical properties. There were supposed to be structural relationships between the contents  $\forall x(x^2 > x)$  and  $1^2 > 1$ , for example, relationships that were supposed to have been illuminated by the function–argument analysis: In particular, the function  $\xi^2 > \xi$  was supposed to occur in both of these contents. And the *structural* relationship between the contents was supposed to illuminate an important *logical* relationship between them. Moreover, the fact that structural relations between contents can be elaborated by formal features of a 'conceptual notation' was supposed to be what explained why formal argument need not be merely the transposition of forms but can instead constitute logical inference, the derivation of truths from truths. So, in abandoning the view that concept-functions are functions from objects to conceptual contents, Frege is also abandoning the only understanding he had of the relations between formal features of expressions and logical features of their associated contents.

To put the point only slightly differently, when Frege abandons the view that concept-functions map objects to conceptual contents, he is simultaneously abandoning his prior conception of how the content of a sentence is determined by the contents of its parts. The central lesson of the puzzle about substitution, as we saw above, was that his early conception of composition was inadequate to explain why (E) and (M) have the (different) contents they do. What Frege needed, then, was a new conception of how contents compose, and the chief requirement on the new conception is that it must do what the old one could not: It must yield an explanation of why (E) and (M), though composed in the same way of parts that Frege had originally regarded as having the same contents, do not themselves have the same content.

Contemporary presentations typically treat the substitution puzzle as raising

---

does not depend upon the view that this function *is* the 'meaning' of  $F$ , only upon the assumption that there is such a function and that it is determined by  $F$ .

<sup>35</sup> This point was also noted, in a similar connection, by Hugly and Sayward (2000).

the question how it is possible for (E) and (M) to express different contents.<sup>36</sup> This is undoubtedly an important question, but—and we cannot emphasize this point strongly enough—it was *not* the question that bothered Frege. Frege's question is more nearly why it is *necessary* that (E) and (M) should express different contents. Of course, (E) and (M), *qua* formal strings, could have the same content, and they would have the same content if 'the Evening Star' and 'the Morning Star' had the same content. But if 'the Evening Star' and 'the Morning Star' do not have the same content, then (E) and (M) *cannot* have the same content, and it is this conditional that wants explaining.<sup>37</sup> Frege's old view could not explain it. What Frege needs, then, is a view that will.

In the discussion that follows, we shall several times reject candidate interpretations of Frege's views about how thoughts compose on the ground that, were they correct, no explanation of why (E) and (M) have different contents would be forthcoming. It is therefore essential that this explanatory question should be one Frege himself takes to be posed by the puzzle about substitution rather than one we have read into it. That it is Frege's own concern is confirmed by how Frege characterizes his solution to the puzzle. Frege's solution, of course, was first to insist upon the distinction between thought and truth-value and then to generalize this distinction, which applies only to sentences, to one that might apply to expressions of any syntactic category, to names, in particular. The generalization is what we know as the distinction between sense and reference, and it makes its first explicit appearance in *Function and Concept*, immediately following the passages quoted earlier in this section:

'2<sup>4</sup>' and '4 × 4' certainly have the same reference, i.e., are proper names of the same number; but they have not the same sense; *consequently*, '2<sup>4</sup> = 4<sup>2</sup>' and '4 × 4 = 4<sup>2</sup>' refer to the same thing, but have

<sup>36</sup> To give just one example, at the opening of his book *Frege's Puzzle*, Nathan Salmon writes that the question Frege poses in the famous passage at the beginning of "On Sense and Reference" is: "How *can* 'a = b', if true, differ in 'cognitive value'... from 'a = a'?" (Salmon, 1986, p. 11, emphasis ours) Salmon reads the passage this way because he, like everyone else, is focused on the question whether we need to distinguish sense from reference. Frege's argument that we should is supposed to be that only then *can* 'a = b' and 'a = a' express different thoughts. The explanatory question we are about to mention never surfaces.

<sup>37</sup> Obviously, the problem does not concern just (E) and (M). What is wanted is an explanation of why, in general, replacing one expression with another that has a different content will change the content of the whole. It is not enough, therefore, to explain why (E) has the content it does and to explain why (M) has the content it does and then to note that these contents are different. (Thanks to Zoltán Gendler Szabó for clarifying this point.)

not the same sense (i.e., in this case: they do not contain the same thought). (FC, op. 14, our emphasis)

Note Frege's use here of the term 'consequently',<sup>38</sup> which indicates that the sentences he mentions express different thoughts *because* their parts have different senses. There is similar language in a letter Frege wrote to Peano in about 1896:

In the proposition 'The evening star is the same as the evening star', we have only [an instance of the principle of identity]; but in the proposition 'The evening star is the morning star' we have something more.

How can the substitution of one proper name for another designating the same object *affect* such changes? One would think that it could *affect* only the form and not the content. . . . I say that the two names have the same reference but not the same sense. . . . *And so it happens* that the thought of our first proposition is different from that of the second; for the thought we express in a proposition is the sense of the proposition. (PMC, pp. 127–8, our emphasis)

Here again, the claim being made is that the two sentences express different thoughts *because* they are composed of expressions that have different senses.<sup>39</sup>

The substitution puzzle shows, then, that there are two theoretical roles that must be distinguished, roles that Frege had originally conflated: that of the content of a sentence and that of the value of a concept-function, both of which had originally been identified with conceptual content. This is why Frege writes in *Grundgesetze* that conceptual content "now divides into what I call thought and what I call truth-value[, as] a consequence of the distinction between the sense and denotation of a sign" (Gg, v. I, p. x).<sup>40</sup> Content henceforth is to be understood as having two aspects, as being a relation, in the general case, between sense and

<sup>38</sup> The German is: und daher haben ' $2^4 = 4^2$ ' and ' $4 \times 4 = 4^2$ ' zwar dieselbe Bedeutung, aber nicht denselben Sinn. . . .

<sup>39</sup> The German for the last sentence is: So kommt es, dass der Gedanke unseres ersten Satzes von dem des zweiten verschieden ist; denn der Gedanke, den wir mit einem Satze ausdrücken, ist der Sinn dieses Satzes. As Kai Wehmeier pointed out to us, the translation "and so it happens" probably doesn't capture the full force of the remark. The phrase "Wie kommt es eigentlich, dass. . ." is a natural way in German to begin a question that calls for an explanation. Frege's phrase, "So kommt es", thus suggests that he is answering just such a question. So it would not be unreasonable to translate the sentence: That is why the thought of our first proposition is different from that of the second. . . .

<sup>40</sup> Frege's language here obviously suggests that the roles of sentential content and the value of

reference. The content of a sentence will consist of a thought and its truth-value: the sense of the sentence and its reference, or, as we put it above, its ‘meaning’.

It is important to keep in mind that Frege does not identify content with either sense or reference alone. In particular, he does not regard thoughts as conceptual contents re-labeled. It is not that Frege could not have had a notion of conceptual content in his new scheme; having truth-values as the values of concept-functions does not prevent Frege from saying, for example, that the conceptual content of a sentence is a complex consisting of a function and its arguments. But he does not go this route, for reasons that shall emerge directly.

### 3 Thoughts As Compositionally Complex

The problem with which the puzzle about substitution leaves Frege is thus an explanatory one: We need an explanation of why (E) and (M) have different contents. With the introduction of the distinction between sense and reference, this amounts to a call for an explanation of why (E) and (M) express different thoughts.<sup>41</sup> Now, Frege is firmly committed to the principle of compositionality:<sup>42</sup> What thoughts (E) and (M) express will be determined by their logico-syntactic structure and relevant facts about their constituents. But, as we have seen, the ‘relevant facts’ here cannot be limited to facts about reference: We cannot explain why (E) and (M) express different thoughts simply in terms of facts about the ‘meanings’, or references, of their constituents. So there must be some other feature of the constituents that contributes to determining what thought is expressed. Let us call this ‘thought determining’ feature of an expression its *sense*. Then the sense of an expression is that feature of it that determines, so far as that expression is concerned, what thoughts are expressed by sentences in which that expression occurs.

---

a concept-function had previously been played by a single notion, and that this notion is what has now split. This, then, is the additional evidence mentioned earlier, in note 21, for our interpretation of Frege’s early view of these matters: that the values of concept-functions are conceptual contents.

<sup>41</sup> There are two ways to think of this. On one, the old question why (E) and (M) have different contents is no longer of interest, and it has simply been replaced by the new question why they express different thoughts. If the old question is supposed still to be of interest, on the other hand, then, since the old notion of content has divided into a new notion of which thought is one component, it is still sufficient for Frege to show that (E) and (M) express different thoughts.

<sup>42</sup> As shall emerge shortly, we think the label ‘compositionality’ is somewhat ill-chosen, if the thesis it labels is one to the effect that relevant facts about wholes are *determined* by relevant facts about their parts. But that is how such principles are known in the literature.

Frege himself gives a similarly abstract characterization of sense in *Grundgesetze*:<sup>43</sup>

The simple or complex names of which a name of a truth-value [that is, a sentence] consists, contribute to the expression of the thought, and this contribution of the individual [name] is its *sense*. (Gg, v. I, §32)

One might object that this characterization—and, indeed, the way Frege formulates the puzzle about substitution in the first place—assumes without justification that thoughts are expressed by *sentences*, that is, that the notion of sense is a linguistic one rather than a merely psychological one. There is, we think, much to be said about this question, but we cannot pursue it here.<sup>44</sup> For present purposes, it will suffice to note that Frege has a considered view about how languages ought to be individuated for the purposes of logic, namely: If we are to isolate those aspects of language that are required if language is to be used in reasoning, then we must individuate languages not just by their syntax, nor even by syntax plus reference, but by syntax plus sense. That is, languages are to be individuated with respect to the relation of *expression*.

In this respect, then, it is a *linguistic* fact that (M) expresses a different thought than does (E). If so, we may indeed characterize the senses of sub-sentential expressions as Frege does in *Grundgesetze*: The sense of the name ‘the Morning Star’, for example, is what it “contribute[s] to the expression of the thought” by such sentences as (M) (Gg, v. I, §32). To be sure, the notion of sense, so characterized, is almost completely programmatic; indeed, it is Frege’s view that in characterizing sense we can never go any further than elucidating the roles senses play, since for senses (unlike, say, numbers), no explicit definition is possible.<sup>45</sup> But perhaps the notion of sense should simply be left in this programmatic state. Frege’s remarks in *Grundgesetze* are obviously in the spirit of his famous ‘context principle’, which requires us “never to ask for the meaning of a word in isolation, but only in the context of a proposition” (Gl, p. x), and the context principle has often been interpreted as licensing one to reject such questions as what the sense of a name is. We should, the suggestion would be, regard the notion of sense as a theoretical construct, an abstraction from whatever pattern we might observe in

<sup>43</sup> Note that Frege uses the term ‘name’, in *Grundgesetze*, to mean any well-formed expression, so the quoted remark applies not just to proper names but to expressions of all types.

<sup>44</sup> We have discussed it elsewhere, however (Antonelli and May, 2000; Heck, 2002; May, 2006).

<sup>45</sup> In this regard, thoughts are different; they are compositions of senses which refer to truth-values.

the thoughts expressed by all of the many sentences that contain a given expression. In the case of ‘the Morning Star’, one such sentence would be (M), and there would be many other such sentences, too.

There is a serious problem with this sort of holism, however. On this view, it would be enough for ‘the Morning Star’ and ‘the Evening Star’ to have different senses that just *one* appropriate pair of sentences in which they occur should express different thoughts, even if all other such pairs expressed the same thoughts. It is by no means required that *all* such pairs should express different thoughts, so it is not required that (E) and (M) should do so. If the notion of sense were left in its purely programmatic state, then, we would have no general reason to suppose that, if ‘the Morning Star’ and ‘the Evening Star’ have different senses, the particular sentences (E) and (M) must express different thoughts. But then no explanation of that fact is forthcoming. The more important point, however, is that the abstractionist view inverts the appropriate order of explanation.<sup>46</sup> Frege wants an explanation of why (E) and (M) express different thoughts. What he wants to be able to say is that they do so because they contain different parts, that these parts have different senses, and that the thoughts the sentences express are *therefore* different. But if the sense of a name simply encapsulates a ‘pattern’ we find in the thoughts expressed by sentences containing it, then ‘the Morning Star’ and ‘the Evening Star’ only have different senses because these ‘patterns’ are different, that is, because some sentences in which these names occur express different thoughts. It is not, then, that (E) and (M) express different thoughts because ‘the Morning Star’ and ‘the Evening Star’ have different senses; rather, ‘the Morning Star’ and ‘the Evening Star’ have different senses because (E) and (M)—or, at least, some sentences like them—express different thoughts.

To put the point differently: Frege’s so-called ‘criterion of cognitive significance’—viz., *a* and *b* express different senses if someone could rationally believe that  $F(a)$  but not that  $F(b)$ —cannot be used to *individuate* the senses of sub-sentential constituents, even if the criterion is regarded as proffering a necessary as well as a sufficient condition for difference of sense.<sup>47</sup> That (E) and (M) express differ-

<sup>46</sup> Many people have worried that holistic views are ultimately incompatible with any robust account of compositionality, e.g., Dummett (1993) and Fodor and Lepore (1992, esp. Ch. 6). We are expressing a similar concern.

<sup>47</sup> There is some controversy on this point. Most of Frege’s applications of the criterion—and all of his really critical applications of it—depend only upon its being a *sufficient* condition for difference of sense. There are times, however, when he seems to use it as a necessary condition. For example, Frege suggests that ‘*A* and *B*’ and ‘*B* and *A*’ have the same sense, on the ground that no-one who understood both could accept one without accepting the other (CT, op. 39). Our view,

ent thoughts *entails* that ‘the Morning Star’ and ‘the Evening Star’ have different senses, but it cannot *constitute* their having different senses, not if we are to be able to give the explanation Frege wants.

It is thus essential for Frege’s purposes that he should give the notion of sense more substantial content. And, as it happens, Frege had resources antecedently available that would allow him at least to start giving it some. In his discussion of identity-statements in *Begriffsschrift*, Frege famously writes:

... [T]he need for a sign for identity of content rests upon the following consideration: the same content can be completely determined in different ways; but that in a particular case *two ways of determining* it really yield the *same result* is the content of a *judgement*. Before this judgement can be made, two distinct names, corresponding to the two ways of determining the content, must be assigned to what these ways determine. (Bg, §8, Frege’s emphasis)

Or again, Frege writes in *Die Grundlagen*:

Why is it... that we are able to make use of identities with such significant results in such diverse fields? Surely it is... because we are able to recognize something as the same again although it is given in a different way. (Gl, §67)

The distinction between the object to which a name refers and the way in which that object is given to us is thus present throughout Frege’s work, from *Begriffsschrift* to “On Sense and Reference” and beyond. But the way Frege understands this distinction changes. In *Begriffsschrift*, Frege regarded the distinction between an object and how it is given as a distinction between content and something that is not itself content but only presents content. In his mature work, on the other hand, Frege regards this same distinction as one *within* content.<sup>48</sup> What gives the otherwise programmatic notion of sense at least some substance is thus the suggestion that the mode of presentation associated with a name is part of its content (SM, op. 27). Of course, how much heft is thereby added depends upon how much we take ourselves to know about ways in which objects may be presented, and, infamously, Frege says little more directly about the matter. But however desperately one might have wanted him to say more, the crucial question for us is how

---

however, is that this is just a slip. A more important sort of example is provided by such pairs as ‘*a* is parallel to *b*’ and ‘the direction of *a* is the same as that of *b*’. We’ll discuss this sort of case in section 5.

<sup>48</sup> This was first pointed out by Angelelli (1967, pp. 38–40).



the fact that ‘the Morning Star’ and ‘the Evening Star’ are associated with distinct modes of presentation is supposed to help us to explain why (E) and (M) express different thoughts. Remember: It was to answer this question that senses were introduced in the first place.

The temptation here is to invoke a principle of compositionality for senses: The sense of a sentence is determined by the senses of its parts. But this is insufficient. Certainly, if ‘the Morning Star’ and ‘the Evening Star’ have the same sense, then the sentences (E) and (M) must also have the same sense. But compositionality, as just stated, simply does not imply that, if ‘the Morning Star’ and ‘the Evening Star’ have different senses, then (E) and (M) will also have different senses. The *reference* of the whole is also determined by the *references* of the parts, but it does not follow from the fact that the reference of ‘Venus’ differs from that of ‘Mars’ that ‘Venus is a planet’ and ‘Mars is a planet’ must have different truth-values; it follows only that they *might* have different truth-values. Similarly if the sense of a sentence is determined by the senses of its parts, then the fact that (E) and (M) contain parts with different senses makes it *possible* that they should express different thoughts, but it in no way requires that they *must*.

If this point has been obscured, it is because of the difference highlighted earlier between contemporary approaches to the puzzle about substitution and Frege’s own reasons for being concerned with it. As we mentioned, contemporary authors tend to emphasize the question how it is *possible* for (E) and (M) to express different thoughts. And if that is the question, then the principle that the thought expressed by a sentence is determined by the senses of its parts shows the way to an answer: The parts do not have the same senses. But that, we emphasize again, was not the problem the puzzle posed for Frege. Frege’s problem was *to explain why* (E) and (M) express different thoughts. But if *that* is the question, then invoking compositionality in the form of a supervenience thesis hardly helps at all. The claim that the sense of a sentence is determined by the senses of its parts is compatible with the claim that two different compositions of senses should converge on the same thought, just as two different compositions of references may converge on the same truth-value. But then we have no explanation of why (E) and (M) express different thoughts.

It follows that senses cannot just be features of expressions that help to determine the thought expressed by sentences containing them. If that were all they were, then, short of base stipulation, we could not rule out the possibility that different compositions of senses should determine the same thought. So what we need, if we are to explain why (E) and (M) express different thoughts, is to earn a right to that principle: that different compositions of senses *must* determine dif-

ferent thoughts. More precisely, what we need in the present context is just the following special case of that principle: If *a* and *b* are expressions with different senses, then substituting *b* for *a* in any sentence in which *a* occurs will, as Frege puts it in the letter to Peano, “effect” a change in the thought expressed. And we can earn a right to this principle if we regard compositions of senses in a more structural light. So suppose that, rather than merely helping to determine the sense of the whole, the senses of the parts were themselves *parts* of the sense of the whole. A thought, on this view, would be a complex, wholly made up from—that is, composed of—its constituent senses, and so nothing above and beyond their composition. Thoughts, therefore, would be the same only if they were made up of the same parts, and would be different otherwise. As we observed earlier, Frege expresses just this view in *Grundgesetze*: “If a name is part of a name of a truth-value, then the sense of the former is part of the thought expressed by the latter” (Gg, v. I, §32). But Frege’s most vivid expression of this view is in the opening remarks of his last published essay, “Compound Thoughts”:

It is astonishing what language can do. With a few syllables, it can express an incalculable number of thoughts, so that even if a thought has been grasped by an inhabitant of the Earth for the very first time, a form of words can be found in which it will be understood by someone else to whom it is entirely new. This would not be possible if we could not distinguish parts in the thought corresponding to the parts of a sentence, so that the structure of the sentence can serve as a picture of the structure of the thought. (CT, op. 36)

For Frege, the structure of language is thus a mirror of the structure of thought.<sup>49</sup> Given that ‘the Morning Star’ and ‘the Evening Star’ express different senses, it will therefore follow that (E) and (M) express different thoughts. Far from being an affectation, then, Frege’s view that the thought expressed by a sentence is compositionally complex—its constituent parts being the senses expressed by the constituent parts of the sentence—is demanded by the structure of his solution

---

<sup>49</sup> One might suggest that a ‘logically perfect’ language would be one in which distinct expressions can never have the same sense—except, perhaps when they are definitionally related. In that case, we could in principle always detect on purely formal grounds whether two sentences expressed different thoughts, since any divergence would be detectable, given the definitions. But Frege’s point here does not require that language mirror thought to such an extent, and no such claim is needed for our purposes, either. As we mentioned earlier, what needs explaining here is why (E) and (M) must have different senses *if* ‘the Morning Star’ and ‘the Evening Star’ have different senses.

to his famous puzzle about substitution.<sup>50</sup>

## 4 How to Compose a Thought

Unfortunately, it is not at all obvious how to apply mereological notions to thoughts. Frege himself, following up the remark just quoted, warns that “. . . we really talk figuratively when we transfer the relation of whole and part to thoughts. . .” (CT, op. 36). It is all well and good to say that the sense of ‘The Morning Star is a planet’ contains the sense of ‘the Morning Star’ and the sense of ‘is a planet’ as its only parts. But that does not suffice to answer the question how the thought that the Morning Star is a planet is *composed* of these parts, for nothing has yet been said about how the parts are bound together, that is, about how senses cohere to form thoughts. There are some answers that are clear losers. For instance, thoughts cannot be mere agglomerations (or fusions) of their constituent senses. This fails with even the simplest cases, since ‘Alex loves Toni’ and ‘Toni loves Alex’ would then have the same sense. At the very least, it would seem, there has to be some further structure in thoughts if are going to address the issue how the parts cohere.

But the mere addition of structure will not be sufficient either, not if thoughts are to be the sort of things Frege thinks they are. Suppose we were to regard thoughts as some sort of structured sets. The thought expressed by ‘the Morning Star is a planet’ might be something like an ordered pair whose first member was the sense of the name ‘the Morning Star’ and whose second member was the sense of the predicate ‘ $\xi$  is a planet’. The thought that Toni loves Alex would be something like an ordered triple containing, in order, the senses of ‘Toni’, ‘loves’, and ‘Alex’, and so would be different from the thought that Alex loves Toni. No such view can be Frege’s, for no such view can solve the fundamental problem how senses cohere to form a thought. That problem, as Frege aptly puts it in discussing a similar idea, is “shifted, but not avoided” by such a treatment (CO, op. 205). On Frege’s view, thoughts are essentially truth-evaluable, so their truth-evaluability must be a consequence of the correct account of what thoughts are. If thoughts are compositions of senses, therefore, the truth-evaluability of thoughts

---

<sup>50</sup> In a somewhat similar vein, Peter Pagin (2003) has argued that the principle that contents have constituent structure is required if we are to be able to explain how ordinary speakers manage to decide what sentence to use to express a given thought. This sort of principle is also related to what Fodor and Lepore (2002) call ‘reverse compositionality’. There has been a fair bit of discussion of that principle (Johnson, 2006; Patterson, 2005; Robbins, 2005).

must be a consequence of how they are composed from senses. Or, to put the point another way, the question how thoughts are composed of senses will not have been answered unless it has been shown how, through the composition of senses, thoughts determine truth-conditions for the sentences that express them. But this is exactly the issue that the view presently under consideration avoids: For all that has been said, the ordered pair consisting of the sense of ‘the Morning Star’ and the sense of ‘ $\xi$  is a planet’ might be true if the Morning Star is not a planet—or, for that matter, if the Morning Star is the only planet, or if it is one of an odd number of planets. And this issue, one must bear in mind, is not metaphysical. The issue is not whether there can be structured sets of senses. If there are senses at all, then presumably there are sets of them. Rather, the problem is semantic: There is nothing intrinsic to an ordered  $n$ -tuple of senses that tells us under what condition it is true.<sup>51</sup>

Frege was certainly well aware that there is a problem about how thoughts cohere, offering the following insight in “On Concept and Object”:<sup>52</sup>

... [N]ot all parts of a thought can be complete; at least one must be ‘unsaturated’ or predicative; otherwise, they would not hold together. For example, the sense of the phrase ‘the number 2’ does not hold together with that of the expression ‘the concept *prime number*’ without a link. (CO, op. 205)

Frege is here contrasting the noun phrase ‘the concept *prime number*’ with the verb phrase ‘is a prime number’, whose sense he regards as being appropriately predicative. What, however, is it supposed to mean that the sense of this verb

---

<sup>51</sup> A possible response at this point would be to hold that “structured thoughts” are not thoughts at all, but rather *representations* of thoughts. A semantic theory that associates sentences with structured thoughts would then not associate sentences with interpretations of those sentences, but with things that themselves *stand in need* of interpretation. This, however, would be to step back from a central aspect of Frege’s view: that thoughts, *qua* contents, *are* interpretations, not something subject to interpretation (Antonelli and May, 2000). We might try to ameliorate this failing by saying that semantic theory is not so much interpretive, but instead translates sentences of one language into a different language with somewhat peculiar lexical primitives: Rather than words we might write or pronounce, the lexical primitives of this language are senses, objects, properties, or what have you; but it is, for all that, a language that could, in principle, be interpreted in different ways. But this would seem to go too far. Theories that associate structured thoughts with sentences do some semantic work; they associate the lexical primitives of the original language with their interpretations. But such a theory leaves the job of semantics only half done, and it leaves the really hard job—that of explaining the nature of semantic composition—completely *undone*.

<sup>52</sup> Similar remarks can be found in the late essay “Negation” (Neg, op. 155).

phrase is ‘unsaturated’? Perhaps the most obvious way to understand this claim is to take the senses of predicates to be unsaturated in the very same way that their references are: Just as the reference of ‘is a prime number’ is a function, so the sense of ‘is a prime number’ is a function. And once we have made that move, it is very natural to suppose that the composition of senses works like the composition of references: Just as the reference of a sentence is determined by applying the concept-function that is the reference of a predicate to the object that is the reference of the predicate’s argument, so the sense of a sentence will be determined by applying the sense-function that is the sense of the predicate to the sense of the predicate’s argument. The arguments of sense-functions will thus be the senses of the expressions that fill the argument-places of the associated predicates, and the value of the function will be the sense of the complex expression constructed from those parts, that is, the thought the sentence expresses. So the sense of a predicate, on this view, is a function from the senses of names to thoughts: The sense of ‘is a planet’, for example, would be the function that maps the sense of ‘Venus’ to the thought that Venus is a planet; the sense of ‘Sirius’ to the thought that Sirius is a planet; and so forth.<sup>53</sup>

The way this interpretation reduces compositionality for senses to function-application, in tandem with compositionality for references, is undeniably elegant. But it cannot have been Frege’s view either, for we still have no account of the connection between composition and truth. Saying that thoughts are the values of sense-functions in no way explains how the composition of senses determines truth-conditions.<sup>54</sup> Certainly no such account falls out of the conception of how senses compose. All we have been told is that the thought expressed by ‘Venus is a planet’ is some function or other of the sense of ‘Venus’.<sup>55</sup> What has that to do with truth-conditions? But the fatal problem is by now familiar: This account of how senses compose simply does not solve the problem with which Frege was concerned; we still cannot explain why (E) and (M) express different

---

<sup>53</sup> This view is strongly identified with Peter Geach (? , pp. 444–5). Dummett (1981, ch 13) discusses Geach’s view extensively. A more recent proponent is Terence Parsons (1996).

<sup>54</sup> A different sort of problem arises from the fact that, on this view, the sense of ‘ $\xi$  is a planet’ would be a function from the senses of names to complexes of which that very function was itself a part. This is not obviously coherent, though perhaps it could be made so.

<sup>55</sup> It is extremely important to see that we have not actually been told which function the sense of ‘is a planet’ is. We have been told that it maps the sense of an expression  $t$  to the sense expressed by the sentence ‘ $t$  is a planet’. But specifying the function requires saying what its value is for an arbitrary argument—in this case, an arbitrary *sense*—not just saying what its value is when an *expression* denoting a sense is substituted into the argument-place of some *expression* that denotes the function.

thoughts. If the sense of a predicate is just a function from senses to thoughts, there is no obvious reason such a function cannot map distinct senses onto the same thought, just as the reference of a predicate can map distinct objects onto the same truth-value. So while this proposal may have some advantage over the ones previously discussed—it is somewhat more explicit, anyway—there has been no real advance. We could, to be sure, simply stipulate that sense-functions must be one-to-one, so that different arguments map to different values. But what would justify that stipulation?

The view that the senses of predicates are functions from senses to thoughts is an outgrowth, we suggest, of a misreading of Frege’s characterization of sense in *Grundgesetze*. He says there that the sense of an expression is that feature of it that “contribute[s] to the expression of [a] thought” by the various sentences in which it is contained (Gg, v. I, §32). This clearly implies that the senses of a sentence’s parts must determine the sense of the whole sentence. But to say that the sense of a part is what it *contributes* to the expression of a thought is to say a good deal more than simply that it helps to *determine* which thought is expressed. Frege’s language suggests that the thought expressed by a sentence is *constituted* from the contributions made by the sentence’s parts—something that is also suggested by the common meaning of the term ‘composition’. But if that is so, then there is an immediate and important consequence for how we understand the senses of subsentential expressions. In particular, the notion of sense, as it applies to subsentential expressions, must be no less *cognitive* than the notion of thought itself is: If a thought is constituted from what the parts contribute, then the parts must contribute something from which a thought might be constituted.

Frege tells us directly what this contribution is when he speaks of sense as mode of presentation. Why, after all, is it even remotely plausible that the fact that (E) and (M) express different thoughts should follow from the fact that ‘the Morning Star’ and ‘the Evening Star’ are associated with different modes of presentation? Because the notion of a mode of presentation is a cognitive one. To entertain the thought that the Morning Star is a planet is to think of a thing given in *one* way as a planet; to entertain the thought that the Evening Star is a planet is to think of a thing given in a *different* way as a planet. The cognitive difference between the modes of presentation is simply absorbed into the thoughts, where it contributes to—or even constitutes—the cognitive difference between them.<sup>56</sup>

---

<sup>56</sup> On Frege’s view, it is a design feature of Begriffsschrift that formally distinct symbols—*figuren*, as he calls them—represent distinct senses. In this regard, consider Frege’s discussion in “The Thought”, where he suggests distinguishing “Dr. Lauben” from “Gustav Lauben” as a

Conversely, it is difficult to see how we could possibly explain why (E) and (M) express different thoughts if we did not suppose that the notion of sense for sub-sentential expressions was, like the notion of sense for sentences, fundamentally cognitive. Suppose, for example, that we were to take the sense of a primitive expression to be the complete history of its use or, perhaps, the time and place where it was first used.<sup>57</sup> Since no two primitive expressions can first have been used at the same time and place, no two primitive expressions could then have the same sense, and so the sense of a complex expression would trivially be determined by the senses of its primitive parts.<sup>58</sup> Such a theory is utterly unsatisfying, however: That ‘the Morning Star’ and ‘the Evening Star’ have different histories of use does not by itself explain why the sentences (E) and (M) express different thoughts but, yet again, only makes it possible that they should. And the same will be true, we suggest, of any view that identifies the sense of an expression with some non-cognitive feature of it.

The argument here is by no means limited to names; *mutatis mutandis*, a similar point applies to the senses of predicates. In the case of sense-functions, the objection would be this: There is no reason, absent blind stipulation, that different sense-functions could not map the same argument to the same thought—or, for that matter, that different sense-functions should not map *different* arguments to the same thought. There is nothing in this conception of how thoughts are composed from senses that prevents *Fa*, *Fb*, and *Gb* from all expressing the same thought, even if *F* and *G*, and *a* and *b*, have different senses. So if we want an explanation of why *Fb* and *Gb* must express different thoughts if *F* and *G* have different senses, the notion of sense for predicates too must be a cognitive notion. The senses of predicates, that is to say, must contain modes of presentation of their references, and so the sense of a predicate must be a way in which a function from objects to truth-values may be given to us.

What might it mean to say that the sense of a predicate contains a mode of presentation of its referent? Let us first consider ordinary functional expressions,

---

propedeutic for making natural languages more perfect (Tht, p. 359, op. 65).

<sup>57</sup> We are assuming here that primitive expressions, unlike complex ones, cannot be meaningful though utterly unused. Whatever the status of that assumption, however, some such trivial account as the one we are considering ought to be possible. Perhaps most trivially, for example, we could simply take the sense of an expression to be the expression itself, so long as we regarded ambiguous expressions as actually being distinct though homophonic.

<sup>58</sup> It simply won’t be possible to have distinct complexes composed in the same way from parts with the same senses. So even if all complex expressions had the same sense, we’d still have compositionality in the form: If you have different senses for the wholes, you have different senses for the parts. That’s how weak that principle is.

such as ' $\xi^2 - 1$ ' and ' $(\xi + 1)(\xi - 1)$ '. These have the same value for every argument and so have the same reference.<sup>59</sup> That they have different senses, however, is clear from the fact that the statement just made—these functions have the same value for every argument—is potentially informative, the common value the functions have for a given argument being determined in different ways: In the one case, we multiply the argument by itself and then subtract one; in the other, we multiply the result of increasing the argument by one by the result of decreasing it by one. These two ways of describing a single mapping from arguments to values correspond to the senses of the two expressions: The sense of a functional expression must encompass a particular way in which values may be associated with arguments. We might, then, think of the senses of functional expressions as what are sometimes called 'functions-in-intension', that is, as arithmetical functions individuated *intensionally* rather than (as is nowadays common) *extensionally*.

Something like this view seems implicit in one of Dummett's discussions of the senses of incomplete expressions:

The condition for the truth of the thought has to be viewed as the satisfaction by a particular object, given in a certain way, of a certain condition on it. The condition to be satisfied by the object is itself given in a particular manner, corresponding to the sense of the predicate: but it is a condition on the object, that is, on the referent of the proper name. . . . [T]he sense of the proper name determines an object as its referent; and the sense of the predicate determines a mapping from objects to truth-values, that is to say, a concept; the sentence is true or false according as the object does or does not fall under the concept. . . . The mapping of objects on to truth-values is not the *sense* of the predicate, but its referent; the sense is, rather, some particular way, which we can grasp, of determining such a mapping. But the sense of the predicate is to be thought of, not as being given directly in terms of a mapping from the *senses* of proper names on to anything, but, rather, in terms of a mapping of *objects* on to truth-values. (Dummett, 1981, pp. 258–9)

<sup>59</sup> Frege wouldn't have been entirely happy with this way of putting the point, since, according to him, one cannot assert an identity between concepts: identity is a first-level relation between objects. What we can do is assert that  $\forall x[x^2 - 1 = (x + 1)(x - 1)]$ , and Frege regards  $\forall x[\Phi x = \Psi x]$  as the "second-level relation which corresponds to, but should not be confused with, equality. . . between objects" (CSM, p. 121). It is, that is to say, the analogue, for one-place first-level functions, of identity.



So the sense of an incomplete expression is a way of presenting a mapping from objects to truth-values. That, more or less, is what a function-in-intension is supposed to be.

This interpretation retains the idea that the senses of predicates are functions, but it reconceives what the arguments and values of these functions are. As a result, it offers an easy answer to the question what ‘completes’ or ‘saturates’ the sense of a functional expression: What completes the sense of a functional expression is simply an *object*. It might seem as if that is impossible, since we would then be unable to distinguish the sense of, say, ‘4<sup>2</sup>’ from that of ‘(2 × 2)<sup>2</sup>’, since in both cases the sense of ‘ξ<sup>2</sup>’ is completed by the same number. And if the sense of ‘ξ<sup>2</sup>’ were a function from objects to senses—from the referent of a term *t* to the sense of ‘ $t^2$ ’—then of course the objection would be conclusive. But that is not the view. The view, rather, is that the sense of a functional expression is an intensionally individuated function from objects to *objects*, and when an object is provided to this function as argument, it must be given to us in some way or other. What distinguishes the sense of ‘4<sup>2</sup>’ from that of ‘(2 × 2)<sup>2</sup>’ is, quite simply, how the argument is given to us. Suppose, for example, that the sense of ‘4’ presents 4 as the successor of 3. Then the sense expressed by ‘4<sup>2</sup>’ may be characterized as follows: to one who understands the expression ‘4<sup>2</sup>’, this expression presents an object as the number that results if the successor of 3 is multiplied by itself. The expression ‘(2 × 2)<sup>2</sup>’ presents that same object in a different way, namely, as the number that results if the successor of one is multiplied by the successor of one and then this intermediary result is multiplied by itself.

Unfortunately, simply saying that the sense of a functional expression is a function-in-intension doesn’t help us to understand how the sense of a name and the sense of a functional expression cohere.<sup>60</sup> As just emphasized, a function-in-intension is a function from objects to objects: Given an object as argument, it delivers an object as value. How, then, do the function-in-intension and the sense of a name together constitute a new sense?<sup>61</sup> Does the function-in-intension somehow meet the sense of the name at its referent and then, before actually determining a value, wait just long enough that it and the sense of the name can together constitute a compositionally complex sense?<sup>62</sup> It does not seem likely

<sup>60</sup> Thanks to Eli Hirsch for forcing the developments that follow.

<sup>61</sup> Recall Frege’s insistence upon a “link” that holds senses together (CO, op. 205).

<sup>62</sup> One might think it would help to say that the sense of a functional expression is completed by the *sense* of a name so that the sense of a functional expression is a function from the senses of names to objects: The sense of ‘ξ<sup>2</sup>’, for example, would map the sense of ‘4’ to 16; the sense of ‘3’ to 9; and so forth. But even if that did seem helpful, there are serious problems with this

that any satisfying answer will be forthcoming.

The question we should be asking, however, is not how the sense of a name and the sense of a predicate cohere to form a thought, as if what were at issue here were a question about the *metaphysics* of thoughts. What is at issue here is no more a question about the metaphysics of thoughts than it is a question about the psychology of thinking—one that could be solved by appeal to some psychological analogue of predication—and the one temptation is just as dangerous as the other. The question at issue is *semantical*: How does the way a sentence is composed of constituents that express certain senses determine what thought is expressed by the sentence itself? The question, that is to say, concerns how sentences are interpreted. To stay on the straight and narrow, then, we must occupy and steadfastly refuse to abandon the semantic perspective that pervades Frege's mature writings.

We certainly do not wish to deny, of course, that Frege sometimes uses terms like 'unsaturated' and 'incomplete' in a way that suggests a metaphysical interpretation. But when he so speaks, we think Frege is himself straying from the path. (We must do as he says, not as he does.) Consider, for example, Frege's claim that predicates are unsaturated. By way of explanation of this doctrine, one sometimes hears it said that, according to Frege, predicates have a 'gap' in them that awaits completion by an argument. But what, in *metaphysical* terms, could it possibly mean that predicates have a 'gap' in them? We have no idea—and Frege himself warns us that such language is not to be taken literally (CO, op. 205). Its purpose is to help convey to Frege's readers what the fundamental difference between names and predicates is. That is to say, the doctrine that predicates are unsaturated is not a metaphysical one but a *syntactic* one, which is precisely what it ought to be, since it is a claim about what kinds of expressions predicates are: It is part of what it is for an expression to be a predicate, Frege is saying, that, in every one of its occurrences, the expression must occur with an appropriate number of arguments of appropriate types.<sup>63</sup> Perhaps we are so tempted to read

---

suggestion. One is that, if the sense of ' $\xi^2$ ' is a function from senses to objects, then it is not obvious what completes the sense of ' $\xi^2$ ' in a sentence like:  $\exists x(x^2 = x)$ , the problem being that variables do not have senses but only values. Perhaps this technical problem could be addressed, but there is another problem that seems quite fatal: This view leaves it unexplained why the sense of ' $\xi^2$ ' must map the sense of ' $2 \times 2$ ' to the same thing to which it maps the sense of '4' or, more generally, why it must map any two senses with the same reference to the same object. That is to say, this view leaves it unclear why the *referent* of ' $\xi^2$ ' must be a function from objects to objects, which is to say that it fails to represent the sense of ' $\xi^2$ ' as a way in which a function from objects to objects may be given to us.

<sup>63</sup> The first appearance of the claim that predicates are unsaturated—or something in the same

metaphysical content into Frege's pronouncements because this claim is now so familiar to us: Surely Frege is not just making a simple point of syntax when he says that predicates are unsaturated! But the claim was not familiar in Frege's day. It was revolutionary (Heck and May, 2006, §1.2).

Frege's claim that concepts (more generally, functions) are unsaturated should not be understood as a metaphysical thesis either: It is not a claim about the inherent nature of concepts; it is not, in particular, a claim that is independent of concepts' role as the interpretations of predicates. We can see this from the way Frege explains the unsaturatedness of concepts in his mature period.<sup>64</sup> What Frege explains initially is the unsaturatedness of predicates; thus, in opening *Grundgesetze* with an "Introduction to Function, Concept, Relation", he tells us first that "The expression for a *function* is *in need of completion, unsaturated*" (Gg, v. I, §1). Only then does he explain the unsaturatedness of concepts, and he explains concepts' unsaturatedness *in terms of* the unsaturatedness of the expressions that refer to them: Because a sign for an argument fills the argument place of a sign for a function, we can say that "The function is completed by the argument. . ." (Gg, v. I, §1). In "Comments on Sense and Meaning", Frege puts it this way:<sup>65</sup>

. . . [O]ne can always speak of the name of a function as having empty places, since what fills them does not, strictly speaking, belong to it.<sup>66</sup> *Accordingly*, I call the function itself unsaturated, or in need of supplementation, *because* its name has first to be completed with the

ballpark—is in a passage we quoted earlier, from "Boole's Logical Calculus":

A sign for a property never appears without a thing to which it might belong being at least indicated, a designation of a relation never without indication of the things which might stand in it. (BLC, p. 17)

That certainly looks like a syntactic doctrine.

<sup>64</sup> How Frege understands the relation between the unsaturatedness of expressions and the unsaturatedness of concepts changes as his views on concepts develop. In 1881, Frege seems to derive the unsaturatedness of predicates from the unsaturatedness of our ideas of what they designate (BLC, p. 17); in 1882, on the other hand, Frege seems to derive this epistemological thesis from a metaphysical thesis about concepts themselves, and then the unsaturatedness of predicates emerges as a consequence of one or another of these (PMC, p. 101).

<sup>65</sup> There are similar remarks in *Function and Concept* (FC, opp. 6ff), "On Concept and Object" (CO, opp. 194–5), and "What Is a Function?" (WF, op. 665), i.e., in almost all of Frege's mature discussions of the incompleteness of predicates.

<sup>66</sup> Long and White's translation has "them" here rather than "it", as if it were anaphoric on "empty places". It is clear, however, that what Frege means is, as he puts it in *Function and Concept*, that "the argument does not belong with a function" (FC, op. 6).

sign of an argument if we are to obtain a meaning that is complete in itself. (CSM, p. 119, our emphasis)

But how precisely should we understand the relationship between the fact that functional expressions are unsaturated and the fact that functions themselves are? One option would be to regard Frege as simply projecting the unsaturatedness of predicates onto their referents—and here ‘projecting’ would be used in its pejorative sense. Fortunately, there is an alternative. If the unsaturatedness of predicates is to have any consequence whatsoever for the nature of what they denote, that consequence must surely issue from the connection between a predicate and its denotation, that is, from something about the *semantics* of predicates. If so, then just as the claim that predicates are unsaturated is a syntactic thesis, so the claim that concepts are unsaturated is a semantic thesis. In particular, the claim that predicates are unsaturated may be understood as a claim about the nature of semantic composition: It is the claim that predication is ‘internal’ to the semantics of predicates, in the sense that the semantic clause that governs a predicate should of itself account for how the denotation of the predicate composes with the denotations of its arguments (Heck and May, 2006, §2). As Frege puts it in a letter to Anton Marty, written in 1882: “... [T]he relation of subject to predicate is not a third thing added to the two, but it belongs to the content of the predicate” (PMC, p. 101).

Frege’s claim that the senses of predicates are unsaturated can be understood similarly. Just as the unsaturatedness of concepts is parasitic on the unsaturatedness of the expressions that denote them, so the unsaturatedness of the senses of predicates is parasitic on the unsaturatedness of their references. The sense of a predicate is a mode of presentation of a concept: So the sense of a predicate is a mode of presentation of something that is, of necessity, true or false of objects; it is a way in which a function from objects to truth-values may be given to us. It too therefore needs to be completed by an object. The sense of ‘ $\xi$  is a planet’, for example, needs to be completed by an object because it is a way in which a function from objects to truth-values may be given to us, and that function itself needs to be completed by an object.

To illustrate, consider our sentence (M), ‘the Morning Star is a planet’. The sense of the name ‘the Morning Star’ is a mode of presentation of an object. The sense of the predicate ‘ $\xi$  is a planet’ is a way in which a function from objects to truth-values may be given to us.<sup>67</sup> In virtue of how (M) is composed, its truth-value is the result of applying that function to the referent of ‘the Morning Star’.

---

<sup>67</sup> Of course, a finer analysis would distinguish parts within the predicate, too, recognizing at

To entertain the thought expressed by (M) is thus to think of this truth-value as being the True: It is, that is to say, to think of an object that is given in a certain way as being mapped to the True by a function that is given in a certain way. The thought expressed by (M) is thus, very roughly, that the last celestial body visible in the morning is mapped to the True by the function that maps all and only planets to the True. Or much more precisely, albeit much less informatively: that the Morning Star is a planet. To entertain the thought expressed by (E), on the other hand, is to think of an object given as the first celestial body visible in the evening as being mapped to the True by the function that maps all and only planets to the True. The thought expressed is thus: that the Evening Star is a planet. The sentences (E) and (M) therefore express different thoughts, and they do so for just the reason Frege says, namely, that they are composed of parts with different senses.

We can now see why there is no need for any ‘sense-glue’ to bind the parts of a thought together—or, less metaphorically, why we do not need any independent account of how senses compose. On Frege’s view, the parts of a thought are bound together by the interaction of two more fundamental forces: The determination of reference by sense, and the composition of references. The parts of thoughts ‘stick together’ because words that express senses combine, in ways determined by their formal properties, to form sentences that have truth-conditions, as determined by the composition of the references that the senses determine. Thoughts are coherently organized not because there is some organizing principle that binds the senses themselves together, but because senses are related to references that compose *via* function-application. And it is because thoughts are structured in this way that, absent reference-failure,<sup>68</sup> a thought must have one of the two truth-

---

least the tense as another component. So something like a mode of presentation of the present would also enter. How such context-dependence is to be handled in a broadly Fregean framework is a very difficult question: For discussion, see the papers by McDowell (1977), Kaplan (1989), Perry (1993), Burge (2005), Evans (1985), Heck (2002), and May (2006), among many others.

<sup>68</sup> One nice feature of our reconstruction is that it explains why Frege was so ambivalent about sentences containing non-referring expressions. On the one hand, Frege insists in “On Sense and Reference” that such a sentence may express a thought even though it contains a name that does not refer to anything (SM, opp. 32–3). On the other hand, however, as Evans emphasized, Frege elsewhere seems to regard such sentences as not *really* expressing thoughts, or as not expressing *real* thoughts but only “mock thoughts” (Evans, 1985, pp. 297ff). If, as we have suggested, the composition of senses is mediated in part by their relation to their references, then it becomes extremely natural to regard sentences containing parts without reference as exhibiting a grave semantic defect: Even if the parts of the sentence do have senses, the parts of the thought will not hold together in the right way if they do not also have references.

values, and the senses of the parts, through how they determine the references of the parts, will determine under what condition that value will be the True. That is to say, the familiar Fregean doctrines that thoughts are truth-evaluable and that they determine truth-conditions emerge, on our interpretation, as consequences of deep features of Frege's conception of how thoughts cohere, of how they are composed from senses.<sup>69</sup>

The same cannot be said for the view that the senses of predicates are functions from senses to thoughts. Although nothing in that view *precludes* thoughts from being truth-evaluable, it will be impossible to characterize sense-functions absent an antecedent conception of what thoughts are if thoughts are the values of such functions (Dummett, 1981, ch. 13). There will then be no alternative to taking it as axiomatic that thoughts are truth-evaluable and determine truth-conditions, and so these facts will be left unexplained.

## 5 Trouble in Paradise?

The advantages bestowed by Frege's considered conception of content are as obvious as they are profound. But there are also significant drawbacks to the view, and there is one we must mention before closing. In various places, Frege speaks of contents' being 'carved'.<sup>70</sup> For instance, in a famous passage in *Die Grundlagen*, Frege suggests that we can reconceive parallelism as an identity and so regard directions as being what parallel lines have in common. In doing so, "We carve up the content in a way different from the original, and this yields us a new concept" (Gl, §64). This idea is arguably central to Frege's position in *Die Grundlagen*. It emerges in the course of his attempt to use the context principle to justify the introduction of names of directions and, analogously, of numbers.

One advantage of the conception of content with which Frege worked in *Begriffsschrift* is that it made such claims intelligible. Indeed, it is actually possible to assert within the formal language of *Begriffsschrift* that distinct sentences have the very same conceptual content, which they will therefore 'carve' logically in different ways: That  $p$  and  $q$  have the same conceptual content is the official sig-

<sup>69</sup> On this view, the connection between the coherence of thoughts and truth-conditions is intimate; it is, as we have mentioned, a consequence of the account of how senses cohere that thoughts express truth-conditions. This contrasts with much current thinking, which takes coherence and truth-conditions to be distinct matters. The potential implications for contemporary semantics will be clear to initiates, but this matter will have to be deferred to a sequel.

<sup>70</sup> We are not the first to discuss the tension that will shortly emerge. It is also discussed by Dummett (1991b), Hale (2001a), and May (2001), among others.

nificance of ' $p \equiv q$ '. With Frege's mature view of content, however, it is no longer possible for him to make any such claim; the equivalence relations that he now has available are either too strong or too weak. Consider the two statements ' $a \parallel b$ ' and ' $dir(a) = dir(b)$ '. It is certainly true that the two sentences will have the same reference. But saying that these sentences have the same 'content', which they carve differently, cannot be understood in those terms. If that's all it takes, then ' $a \parallel a$ ' and 'Snow is white' have the same content, too, which they carve differently.

The other option would be to say that ' $a \parallel b$ ' and ' $dir(a) = dir(b)$ ' express the same thought yet carve it differently. But these two sentences cannot express the same thought. This is not just because they have different parts. We cannot simply say that sentences that have parts with different senses must express different thoughts and leave it at that. For one thing, that would ignore the fact that the notion of 'same sentence' that is wanted here is not the everyday one but one that is to be reconstructed in light of a sophisticated conception of logical structure. But more fundamentally, it would commit a simple fallacy, comparable to claiming that a single object cannot be composed of different parts: If an object is composed of parts, then it is also composed of the parts of those parts. And so, in this case, one might insist that ' $\xi \parallel \eta$ ' and ' $dir(\xi) = dir(\eta)$ ' have the same sense because they are, ultimately, composed of the same parts.<sup>71</sup> But while this move is, in principle, consistent with Frege's conception of how senses compose, it conflicts with other of his commitments. The only way we can see to implement the proposal is to construe it as a suggestion that parallelism should be analyzed in terms of sameness of direction. But this is something Frege explicitly denies on the ground that "everything geometrical must originally be given in intuition" (Gl, §64). Parallelism is supposed to be the more fundamental notion.

Frege has of course abandoned such 'abstractive' definitions by the time he

<sup>71</sup> This is why the fact that thoughts can be decomposed in different ways does not threaten the thesis that thoughts have a unique best decomposition. Different decompositions of e.g. 'Bill loves Sue' may simply represent different ways of 'grouping' the parts of that thought. The best decomposition is then just the one that articulates the structure of the thought most completely. The textual evidence that Frege would have denied this in his mature period is weak. Textor (2009, p. 113), for example, cites Frege's remark that "The thought itself does not determine what is to be regarded as subject" (CO, op. 199). But we see no reason to suppose that, for Frege, fully articulating a thought's structure would require specifying its 'subject' and so regard Textor's attempt to resolve the resulting tension as unnecessary. That said, it is to Textor's credit that, unlike most who have discussed these issues, he clearly appreciates the difference between cases involving grammatical transformations and the sorts of cases we are discussing in the main text. Our point, in effect, is just that neither the 'different groupings' idea nor anything else one might reasonably say about passivization and the like can help with Law V.

writes *Grundgesetze*. But a similar problem arises nonetheless, and we can still find him making similar remarks, even in his mature writings. When explaining the notion of a value-range in *Function and Concept*, for example, Frege says that ‘ $\acute{\epsilon}(\epsilon^2 - 4\epsilon) = \acute{\alpha}(\alpha(\alpha - 4))$ ’ and ‘ $\forall x[x^2 - 4x = x(x - 4)]$ ’ “express[] the same sense, but in...different way[s]” (FC, op. 11). It is hard to be sure just what Frege has in mind here.<sup>72</sup> Surely “the fundamental law of logic that permits the transformation of an equality holding generally into an equation” between value-ranges (Gg, v. II, §147) is not to be compared to a grammatical transformation, like passivization: ‘ $\acute{\epsilon}(\epsilon^2 - 4\epsilon) = \acute{\alpha}(\alpha(\alpha - 4))$ ’ and ‘ $\forall x[x^2 - 4x = x(x - 4)]$ ’ are not related in the way “John saw Mary” and “Mary was seen by John” are.<sup>73</sup> Another proposal would be analogous to the one about directions and parallelism that we just discussed: Take the second-level functional expressions ‘ $\forall x(\exists x = \Phi x)$ ’ and ‘ $\acute{\epsilon}(\exists \epsilon) = \acute{\epsilon}(\Phi \epsilon)$ ’ to have the same sense. But then, as in the previous case, there seems to be only one way to implement the proposal: to claim that the co-extensiveness of functions should be analyzed in terms of identity of value-ranges. But this conflicts with Frege’s well-known (and distinctive) insistence that concepts are prior to extensions (Gg, v. I, pp. 1–4).

In other places, Frege expresses a somewhat different view. In “Comments on Sense and Meaning”, written but a few years after *Function and Concept*, he insists that “... the reference and not the sense of words [is] the essential thing for logic. ... [T]he laws of logic are first and foremost laws in the realm of reference

<sup>72</sup> As we remarked earlier in note 47, some commentators regard Frege’s ‘criterion of cognitive significance’ as expressing a necessary as well as a sufficient condition for the distinctness of thoughts. If so, then the mentioned formulae express different senses only if one could rationally believe one but not the other, and that, it is said, is impossible. But Frege articulates the criterion as a necessary condition only once in his mature writings (BSLD, p. 197), and that is in an unpublished fragment. Moreover, Frege is clearly aware that the criterion threatens to imply that all of the Basic Laws of the Begriffsschrift express the same thought; if that weren’t bad enough, a Sorities-like argument then threatens to imply that all the theorems do, too. As Rumfitt (1994, pp. 608–9) notes, Frege’s attempt to prevent these consequences is not just *ad hoc* but insufficiently general. We suspect that is why the necessary condition does not appear in the *Logical Investigations*, of which the fragment in question is an early antecedent. Frege is only using it to distinguish sense from tone, or coloring, and there are better ways to do that.

<sup>73</sup> In effect, this is the view of Hans Sluga, who maintains that “a thought concerning a function is the same as one concerning a value-range” (Sluga, 1980, p. 157), so that Basic Law V will be analytic. But this view owes an account of the ‘grammar’ of such statements that would make the structural relation between them clear, as transformational grammars tried to account for passivization. We need to be told *how* “... analysing what is on one side... we obtain what is on the other side...” (Sluga, 1980, p. 156). But then this sort of view has the same obligations as the one we are about to discuss.



and only relate indirectly to sense” (CSM, p. 122). So far as the correctness of proofs is concerned, this is clearly right. If Basic Law V were true, then any thought correctly inferred from it would also be true: It does not matter, so far as the *truth* of the theorems of *Grundgesetze* are concerned, whether the two sides of Basic Law V have the same sense, or are true in the same ‘circumstances’, or what have you. But if our concern is to show that the laws of arithmetic are themselves laws of logic, that they can be proven from logical laws and definitions—and this, of course, was precisely Frege’s concern—then the question whether Basic Law V is itself a law of logic cannot be indefinitely postponed. And then the mere truth of Basic Law V, which is all the material equivalence of its two sides guarantees, is insufficient (Heck, 2007; May, 2001). What Frege seems to need, but does not have, is some intermediate equivalence, one that would follow from identity of sense, entail identity of reference, and be sufficient for analyticity.

The problem, then, is that, by Frege’s lights, thoughts themselves cannot be carved in different ways. Rather, if such talk makes any sense at all, thoughts are what carve content—“content” here being something like the realm of reference, that is, reality—and the problem is precisely that Frege lacks the theoretical machinery even to make such a remark. Lacking such machinery, it is not surprising that Frege should say in *Function and Concept* that the two sides of Law V express the same thought; it’s the best he can do with what he has. But we doubt that this was ever Frege’s considered view, and we know of nowhere besides *Function and Concept* that he makes this claim.<sup>74</sup>

Ultimately, then, Frege has no justification for his claim that Basic Law V expresses a basic law of logic. He can insist that *he* “hold[s] that it purely logical” (Gg, v. I, p. vii), but he has nothing substantial to say to someone who holds otherwise.<sup>75</sup>

<sup>74</sup> And just a few years later, when one might have expected him to repeat it, he doesn’t. In “Comments on Sense and Meaning” (CSM, p. 121), Frege speaks of these two expressions

$$\begin{aligned} \forall x[(x^2 = 1) = ((x + 1)^2 = 2(x + 1))] \\ (\alpha^2 + 1) \overset{\alpha}{\asymp} ((\alpha + 1)^2 = 2(\alpha + 1)) \end{aligned}$$

as having the same sense, where the latter is effectively defined in terms of the former. By way of comparison, he also mentions

$$\hat{\varepsilon}(\varepsilon^2 = 1) = \hat{\alpha}((\alpha + 1)^2 = 2(\alpha + 1)),$$

but he does *not* say or imply that it has the same sense as the other two.

<sup>75</sup> Frege’s main independent criterion for a proposition’s being a primitive law of logic is that is

## 6 Closing

Before arriving at the sense–reference distinction as his stable take on content, Frege held a view that differed in important respects. Frege’s initial position, at the time of *Begriffsschrift*, was that we have a window on logical structure because we grasp a linguistic system—the conceptual notation—that represents just this structure. This approach, however, was seen to be wanting, ultimately because it could not support a notion of truth. And even though this could be ameliorated by an articulation of logical properties as integral to the content represented by *Begriffsschrift*, this too was unstable. Although a substantive notion of truth does become available with the notion of an object’s falling under a concept, the substitution puzzle dooms Frege’s earliest account of how the contents of sentences are determined by the contents of their parts. As a result, Frege developed a view that changed the cognitive locus. Now, it is thoughts that we grasp, and it is through this grasp of thoughts that we have access to logical structure. Content is re-conceptualized as a relation between sense and reference: sense determines reference. Language, in turn, remains representational, but, unlike on the prior view, only indirectly so: Language represents content only by expressing sense, so that the way a sentence is composed of words mirrors the way a thought is composed of senses: “The world of thoughts has a model in the world of sentences, expressions, words, signs” (Neg, op. 148). And with this insight, Frege can now bring into alignment the formal and logical notions of structure, so as to make possible an explanation of how true contents follow from other true contents. Because the sentences of *Begriffsschrift* express thoughts, and because the rules of the *Begriffsschrift* have been so formulated that, “if in accordance with them a sentence is derived from true sentences, the new sentence will also be true” (Gg, II §104), formal derivations in the *Begriffsschrift* are not just of symbols from symbols but of true thoughts from true thoughts.

As we mentioned at the outset, the relations between truth, content, and conceptual structure are critical to Frege’s philosophy: Explicating these relationships is precisely what is required for Frege to reconcile the formal character of his logical system with his insistence that sentences of *Begriffsschrift* are not just empty forms but vehicles of content. And that observation is what allows us finally, we think, to understand why “On Sense and Reference” is no mere diversion

---

should be self-evident. The question what precisely Frege means by self-evidence has been much discussed of late (Jeshion, 2001; ?). The central idea behind this work is that Frege might have thought Law V self-evident even if it is not evident *to us* (Gg, v. I, p. vii; v. II, p. 253). But as Stewart Shapiro (2009) emphasizes, that is not something it is easy to understand.

from Frege's central project.<sup>76</sup> What we have seen is that the distinction between sense and reference emerges *within*, not alongside, the development of the logicist project. It emerges because Frege's famous puzzles about substitution threaten his understanding of the formal character of logic: his understanding of how the logical properties of contents might be mirrored in the signs that express them. That is what forces Frege to revise his early views about how sentences express contents, and it is this revision that leads to the distinction between sense and reference.<sup>77</sup>

## References

- Angelelli, I. (1967). *Studies on Gottlob Frege and Traditional Philosophy*. Dordrecht, Reidel.
- Antonelli, A. and May, R. (2000). 'Frege's new science', *Notre Dame Journal of Formal Logic* 41: 242–70.
- Beaney, M. (2007). 'Frege's use of function-argument analysis and his introduction of truth-values as objects', *Grazer Philosophische Studien* 75: 93–123.
- Bell, D. (1987). 'Thoughts', *Notre Dame Journal of Formal Logic* 28: 36–50.
- Burge, T. (2005). 'Sinning against Frege', in *Truth, Thought, Reason: Essays on Frege*. New York, Oxford University Press, 213–39.
- Dummett, M. (1981). *The Interpretation of Frege's Philosophy*. Cambridge MA, Harvard University Press.
- (1991a). *Frege: Philosophy of Mathematics*. Cambridge MA, Harvard University Press.

<sup>76</sup> Ricketts (2005) aims to address the relation between these. But his estimation of the value of the sense–reference distinction to Frege's project is very different from ours.

<sup>77</sup> This paper was presented as the Gareth Evans Memorial Lecture at Oxford University in May 2006, as well as at Brandeis University, Yale University, and Stanford University, and at the first Semantics and Philosophy in Europe conference. Thanks to all who attended these presentations for their comments and questions. Special thanks are due to George Bealer, Mike Beaney, Sir Michael Dummett, Zoltán Gendler-Szabó, Eli Hirsch, John Hawthorne, Ofra Magidor, Ruth Barcan Marcus, and Andreas Teuber. Thanks to Bill Demopoulos, Michael Glanzberg, Øystein Linnebo, Richard Mendelsohn, Jason Stanley, Kai Wehmeier, and Ed Zalta for helpful comments and discussion, both about this paper and an earlier paper (Heck and May, 2006), in which many of the ideas expressed here were presented in a somewhat different, and less satisfactory, form.

- 
- (1991b). ‘More about thoughts’, in *Frege and Other Philosophers*. Oxford, Clarendon Press, 289–314.
- (1993). ‘What is a theory of meaning? (I)’, in *The Seas of Language*. Oxford, Clarendon Press, 1–33.
- Evans, G. (1985). ‘Understanding demonstratives’, in *Collected Papers*. Oxford, Clarendon Press, 291–321.
- Fodor, J. and Lepore, E. (1992). *Holism: A Shopper’s Guide*. Cambridge MA, Blackwells.
- (2002). ‘Why compositionality won’t go away: Reflections on Horwich’s ‘deflationary’ theory’, in *The Compositionality Papers*. Oxford, Clarendon Press, 43–62.
- Frege, G. (1964). *The Basic Laws of Arithmetic: Exposition of the System*, ed. and tr. by M. Furth. Berkeley CA, University of California Press.
- (1970). *Translations from the Philosophical Writings of Gottlob Frege*, Geach, P. and Black, M., eds. Oxford, Blackwell.
- (1972). *Conceptual Notation and Related Articles*, ed. and tr. by T. W. Bynum. New York, Oxford University Press.
- (1979). *Posthumous Writings*, Hermes, H., Kambartel, F., and Kaulbach, F., eds., tr. by P. Long and R. White. Chicago, University of Chicago Press.
- (1984). *Collected Papers on Mathematics, Logic, and Philosophy*, McGuinness, B., ed. Oxford, Basil Blackwell.
- (1997). *The Frege Reader*, Beaney, M., ed. Oxford, Blackwell.
- (AimCN). ‘On the aim of the conceptual notation’, tr. by T. W. Bynum, in Frege 1972, 90–100.
- (Bg). ‘Begriffsschrift: A formula language modeled upon that of arithmetic, for pure thought’, tr. by J. van Heijenoort. 5–82.
- (BgBynum). *Conceptual Notation*, ed. and tr. by T. W. Bynum. In Frege 1972.

- 
- (BLC). ‘Boole’s logical calculus and the concept-script’, tr. by P. Long and R. White, in Frege 1979, 9–52.
- (BSLD). ‘A brief survey of my logical doctrines’, tr. by P. Long and R. White, in Frege 1979, 197–202.
- (CO). ‘On concept and object’, tr. by P. Geach, in Frege 1984, 182–94. Also in Frege 1997, 181–93.
- (CSM). ‘Comments on sense and meaning’, tr. by P. Long and R. White, in Frege 1979, 118–25. Also in Frege 1997, 172–80.
- (CT). ‘Compound thoughts’, tr. by P. Geach and R. H. Stoothoff, in Frege 1984, 390–406.
- (Darm). ‘Notes for Ludwig Darmstaedter’, tr. by P. Long and R. White, in Frege 1979, 253–257. Also in Frege 1997, pp. 362–7.
- (FC). ‘Function and concept’, tr. by P. Geach, in Frege 1984, 137–56. Also in Frege 1997, 130–48.
- (FTA). ‘Formal theories of arithmetic’, tr. by E.-H. W. Kluge, in Frege 1984, 112–21.
- (Geo1). ‘On the foundations of geometry: First series’, tr. by E.-H. W. Kluge, in Frege 1984, 273–84.
- (Geo2). ‘On the foundations of geometry: Second series’, tr. by E.-H. W. Kluge, in Frege 1984, 293–340.
- (Gg). *Grundgesetze der Arithmetik*. Hildesheim, Georg Olms Verlagsbuchhandlung. Translations of Part I are in Frege 1964, and of portions of Part III, in Frege 1970. Extracts appear in Frege 1997, 194–223.
- (Gl). *The Foundations of Arithmetic*, 2d revised edition, tr. by J. L. Austin. Evanston IL, Northwestern University Press.
- (Neg). ‘Negation’, tr. by P. Geach and R. H. Stoothoff, in Frege 1984, 373–389. Also in Frege 1997, pp. 346–61.
- (PCN). ‘On Mr. Peano’s conceptual notation and my own’, tr. by V. H. Dudman, in Frege 1984, 234–48.

- 
- (PMC). *Philosophical and Mathematical Correspondence*, Gabriel, G., Hermes, H., Kambartel, F., Thiel, C., and Veraart, A., eds., tr. by H. Kaal. Chicago, University of Chicago Press.
- (SM). ‘On sense and meaning’, tr. by M. Black, in Frege 1984, 157–77. Also in Frege 1997, 151–71.
- (Tht). ‘Thoughts’, tr. by P. Geach and R. H. Stoothoff, in Frege 1984, 351–72. Also in Frege 1997, 325–45.
- (WF). ‘What is a function?’, tr. by P. Geach, in Frege 1984, 285–292.
- Geach, P. T. (1965). ‘Assertion’, *Philosophical Review* 74: 449–65.
- Hale, B. (2001a). ‘Grundlagen section 64’, in *The Reason’s Proper Study*. Oxford, Clarendon Press, 91–116.
- (2001b). ‘A response to Potter and Smiley: Abstraction by recarving’, *Proceedings of the Aristotelian Society* 101: 339–58.
- Heck, R. G. (2002). ‘Do demonstratives have senses?’, *Philosophers’ Imprint*, 2. <http://www.philosophersimprint.org/002002/>.
- (2003). ‘Frege on identity and identity-statements: A reply to Thau and Caplan’, *Canadian Journal of Philosophy* 33: 83–102.
- (2007). ‘Frege and semantics’, *Grazer Philosophische Studien* 75: 27–63.
- (2013). ‘Formal arithmetic before *Grundgesetze*’, in P. Ebert and M. Rosberg (eds.), *A Companion to Grundgesetze*. Oxford, Oxford University Press. Forthcoming.
- Heck, R. G. and May, R. (2006). ‘Frege’s contribution to philosophy of language’, in E. Lepore and B. Smith (eds.), *The Oxford Handbook of Philosophy of Language*. Oxford, Oxford University Press, 3–39.
- Hugly, P. and Sayward, C. (2000). ‘Frege on identities’, *History and Philosophy of Logic* 21: 195–205.
- Jeshion, R. (2001). ‘Frege’s notion of self-evidence’, *Mind* 110: 937–76.
- Johnson, K. (2006). ‘On the nature of reverse compositionality’, *Erkenntnis* 64: 37–60.

- 
- Kaplan, D. (1989). 'Demonstratives', in J. Almog, *et al.* (eds.), *Themes from Kaplan*. New York, Oxford University Press, 481–563.
- Kripke, S. (1980). *Naming and Necessity*. Cambridge MA, Harvard University Press.
- Levine, J. (2002). 'Analysis and decomposition in Frege and Russell', *Philosophical Quarterly* 52: 195–216.
- Marcus, R. B. (1995). 'Modalities and intensional languages', in *Modalities: Philosophical Essays*. Oxford, Oxford University Press, 3–24.
- May, R. (2001). 'Frege on identity statements', in C. Cecchetto, *et al.* (eds.), *Semantic Interfaces: Reference, Anaphora, and Aspect*. Stanford, CSLI Publications, 1–62.
- (2006). 'Frege on indexicals', *Philosophical Review* 115: 487–516.
- McDowell, J. (1977). 'On the sense and reference of a proper name', *Mind* 86: 159–85.
- Pagin, P. (2003). 'Communication and strong compositionality', *Journal of Philosophical Logic* 32: 287–322.
- Parsons, T. (1996). 'Fregean theories of truth and meaning', in M. Schirn (ed.), *Frege: Importance and Legacy*. New York, de Gruyter, 371–409.
- Patterson, D. (2005). 'Learnability and compositionality', *Mind and Language* 20: 326–52.
- Perry, J. (1993). 'Frege on demonstratives', in *The Problem of the Essential Indexical, and Other Essays*. New York, Oxford University Press, 3–32.
- Potter, M. and Smiley, T. (2001). 'Abstraction by recarving', *Proceedings of the Aristotelian Society* 101: 327–38.
- (2002). 'Recarving content: Hale's final proposal', *Proceedings of the Aristotelian Society* 102: 351–4.
- Reck, E. H. and Awodey, S., eds. (2004). *Frege's Lectures on Logic: Carnap's Student Notes, 1910–1914*. Peru IL, Open Court. Based upon the German text, edited by Gottfried Gabriel.

- 
- Ricketts, T. (2005). 'Generality, meaning, and sense in Frege', in M. Beaney and E. H. Reck (eds.), *Gottlob Frege: Critical Assessments of Leading Philosophers*, volume IV. London, Routledge, 13–36.
- Robbins, P. (2005). 'The myth of reverse compositionality', *Philosophical Studies* 125: 251–75.
- Rumfitt, I. (1994). 'Frege's theory of predication: An elaboration and defense, with some new applications', *Philosophical Review* 103: 599–637.
- Salmon, N. (1986). *Frege's Puzzle*. Cambridge MA, MIT Press.
- Shapiro, S. (2009). 'We hold these truths to be self-evident: But what do we mean by that?', *Review of Symbolic Logic* 2: 175–207.
- Sluga, H. (1980). *Gottlob Frege*. Boston, Routledge and Kegan Paul.
- Sundholm, G. (2001). 'Frege, August Bebel, and the return of Alsace-Lorraine: The dating of the distinction between *Sinn* and *Bedeutung*', *History and Philosophy of Logic* 22: 57–73.
- Textor, M. (2009). 'A repair to Frege's theory of thoughts', *Synthese* 167: 105–23.
- Thau, M. and Caplan, B. (2001). 'What's puzzling Gottlob Frege?', *Canadian Journal of Philosophy* 31: 159–200.
- van Heijenoort, J. (1967). 'Logic as calculus and logic as language', *Synthese* 17: 324–30.
- Vilkko, R. (1998). 'The reception of Frege's *Begriffsschrift*', *Historia Mathematica* 25: 412–22.
- Wagner, S. (1983). 'Frege's definition of number', *Notre Dame Journal of Formal Logic* 24: 1–21.