

Thoughts about Thoughts: The Structure of Fregean Propositions

Nathan Bice

Submitted in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy  
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2019

© 2019  
Nathan Bice  
All rights reserved

## ABSTRACT

Thoughts about Thoughts: The Structure of Fregean Propositions

Nathan Bice

This dissertation is about the structure of thought. Following Gottlob Frege, I define a thought as the sort of content relevant to determining whether an assertion is true or false. The historical component of the dissertation involves interpreting Frege's actual views on the structure of thought. I argue that Frege did not think that a thought has a unique decomposition into its component senses, but rather the same thought can be decomposed into senses in a variety of distinct ways. I extend Frege's position and use it to develop an account of the hierarchy of senses, the senses expressed by indexicals and demonstratives, and the distinction between logical and non-logical structure. I also discuss various connections with the nature of meta-representation, our capacity for reflective judgment, some aspects of the structure of conscious experience, the way we perceive regions of space and durations of time, and our conscious awareness of our own perceptions and events of thinking.

# Contents

<b>Acknowledgements</b>	<b>iv</b>
<b>Introduction</b>	<b>1</b>
0.1 Content and Thought . . . . .	1
0.2 Summary of the Dissertation . . . . .	3
0.3 A Note on Translation . . . . .	16
<b>1 The Decomposition of Thought</b>	<b>17</b>
1.1 Introduction . . . . .	17
1.2 Logic, Formal Languages, and Conceptual Content . . . . .	19
1.3 Some Remarks on the <i>Grundlagen</i> . . . . .	30
1.4 Function and Object, Sense and Reference, Thought and Truth-Value	41
1.5 A Puzzle . . . . .	50
1.6 The Structure of Thought . . . . .	55
1.7 Parts and Wholes . . . . .	62
1.8 Consequences . . . . .	67
1.9 Appendix . . . . .	72

<b>2</b>	<b>Meta-Representation and the Hierarchy of Senses</b>	<b>81</b>
2.1	Introduction . . . . .	81
2.2	Frege’s Hierarchy of Senses . . . . .	89
2.3	Interpreting Frege . . . . .	93
2.4	Recent Approaches to the Hierarchy . . . . .	99
2.5	Meta-Representation, Consciousness, and Judgment . . . . .	112
<b>3</b>	<b>Demonstratives and Self-Reference</b>	<b>123</b>
3.1	Introduction . . . . .	123
3.2	Frege on Demonstratives and Indexicals . . . . .	133
3.3	Objections to Kripke’s Interpretation . . . . .	142
3.4	Differing Conceptions of Sense . . . . .	148
3.5	Spaces, Times, and Objectivity . . . . .	154
3.6	Modality and Self-Reference . . . . .	161
3.7	Preservation of Thought . . . . .	168
<b>4</b>	<b>Presuppositions of Inference</b>	<b>171</b>
4.1	Introduction . . . . .	171
4.2	A Fregean Conception of Truth . . . . .	175
4.3	Atomic Form . . . . .	184
4.4	Logical Connectives . . . . .	191
4.5	Conclusion . . . . .	202
	<b>Conclusion</b>	<b>204</b>



# Acknowledgements

I would like to thank the members of my dissertation committee: Achille Varzi, Justin Clarke-Doane, John Morrison, and Anubav Vasudevan. I greatly appreciate their stimulating comments and thought-provoking criticisms. Their support throughout my time at Columbia has been a source of great strength.

Early versions of Chapter 1 were presented at Columbia University and New York University and early versions of Chapters 2 and 4 were presented at Columbia University. I would like to thank the audiences at those events for their helpful comments and criticisms.

Various individuals contributed to the writing of this dissertation through conversation, comments on early drafts, and much-needed support. I would especially like to thank David Albert, Daniel Boyd, Simon Brown, David Chalmers, Jessica Collins, Borhane Blili Hamelin, Nicholas Engel, Sidney Felder, Alison Fernandes, Robby Finley, Branden Fitelson, Jeremy Forster, Nemira Gasiunas, Rebecca Harrison, Jeffrey Helzner, Max Hayward, Matthew Heeney, Daniel Hoek, Brian Kim, Patricia Kitcher, Saul Kripke, Jonathan Lawhead, Sharon Li, Yang Liu, Béatrice Longuenesse, Antonella Mallozzi, Oliver Marshall, Robert May, Katharine McIntyre, Jorge Morales,

Matt Moss, Michael Nielsen, Ignacio Ojea Quintana, Asya Passinsky, Christopher Peacocke, Zee Perry, Casey Speer, Rush Stewart, Katja Vogt, Daniel Waxman, Naf-tali Weinberger, Porter Williams, Jack Woods, and Mike Zebrick.

I am especially grateful to my principal supervisor: Haim Gaifman. His support from my arrival at Columbia to the present has been immeasurable. I have been impressed by his philosophical skill since first reading his classic paper “What Gödel’s Incompleteness Result Does and Does Not Show” as an undergraduate. It wasn’t until I arrived at Columbia that I truly became aware of his patience and generosity as a teacher. He has dedicated an incredible amount of time to my philosophical education, and for that I am truly grateful.



To my parents

# Introduction

## 0.1 Content and Thought

This dissertation is about the structure of thought. By ‘thought’ I don’t mean a particular event of thinking, but rather an aspect of the content expressed by such an event. In particular, by ‘thought’ I mean the sort of content relevant to determining whether an assertion (or event of thinking, utterance, sentence, etc.) is true or false. That is, I will use ‘thought’ as a technical notion that can be abstracted from an intuitive notion of ‘content’ or ‘meaning’. This usage stems from Gottlob Frege.<sup>1</sup> Hence, a thought is an abstraction from the complete content of e.g. a declarative sentence, and it fixes a truth-condition.

The value of discussing thoughts is that it allows us to isolate the aspect of meaning relevant to determining truth or falsity. We can avoid getting caught up in debates about the exact nature of linguistic meaning so long as we agree that one aspect of meaning is that which determines the conditions under which what is said is true or

---

<sup>1</sup>For example: “So one has to separate off from the content of a sentence the part that alone can be accepted as true or rejected as false. I call this part the thought expressed by the sentence.” (1906c, PW pg. 197-8)

false.<sup>2</sup>

Thoughts are structured entities. When I compare the assertions ‘John has brown hair’, ‘Mary has brown hair’, ‘Jim has brown hair’, etc., I can recognize that by holding the expression ‘has brown hair’ fixed while allowing the rest of the assertion to vary, I vary the conditions under which the corresponding assertions are true in a systematic way. Hence the expression ‘has brown hair’ makes a systematic contribution to the truth-conditions of assertions containing it, from which we can conclude that the corresponding expressed thoughts have a structural feature in common. This structural feature is the aspect of the content or meaning of the expression ‘has brown hair’ in virtue of which it contributes to the truth-conditions of complete sentences in which it occurs. Following Frege, I call this structural feature the *sense* of the expression, and identify the sense of a complete declarative sentence with the thought expressed by it.<sup>3</sup>

Note that I will not attempt to give individuation-conditions for thoughts within

---

<sup>2</sup>Of course, there is little I can say here to convince the reader who doesn’t think there are any meanings at all, or that nothing ever really means anything, or that meanings don’t fix truth-conditions. Nor will I convince the reader who doesn’t think that the sort of abstraction I suggest is possible (I would be curious, however, how such a reader makes sense of many ordinary properties that seem to be arrived at via abstraction from cases). Such a reader should still find something of value in what follows, since many of the claims I make throughout this dissertation do not depend upon this starting-point.

<sup>3</sup>This is Frege’s official definition of sense in the *Grundgesetze* (note that for Frege complete declarative sentences refer to truth-values): “Every such name of a truth-value *expresses* a sense, a *thought*. For owing to our stipulations, it is determined under which conditions it refers to the True. The sense of this name, the *thought*, is: that these conditions are fulfilled...Now, the simple or complex names of which the name of a truth-value consists contribute to expressing the thought, and this contribution of the individual name is its *sense*. If a name is part of the name of a truth-value, then the sense of the former name is part of the thought expressed by the latter.” (1893, pg. 50-1)

Frege wanted the notion of sense to play various roles, including being the mode of presentation (or way of thinking) of its referent (1892a, pg. 26-7/158) and (in the case of a complete thought) being the object of belief (ibid pg. 37/166). We will discuss these roles in Chapters 2 and 3. I will argue in Chapter 3 that irrespective of Frege’s own views on the matter, the conception of sense (and thought) I have just outlined ought to be considered the most fundamental.

this dissertation. I will, however, argue that thoughts ought to be characterized at a higher level of abstraction than they often are in the literature. I will give various reasons for this, including (1) reflection on cases, (2) the resolution of various issues concerning the hierarchy of senses and the contribution of demonstratives and indexicals to truth-conditions, and (3) connections to various other philosophical positions.

While investigating the structure of thoughts for its own sake is an arguably worthwhile project in the Philosophy of Language and Mind,<sup>4</sup> I argue throughout this dissertation that such an investigation reveals fruitful connections with the nature of meta-representation, our capacity for reflective judgment, some aspects of the structure of conscious experience, the way we perceive regions of space and durations of time, our conscious awareness of our own perceptions and events of thinking, and the distinction between logical and non-logical structure.

## 0.2 Summary of the Dissertation

This dissertation proceeds as follows: Chapter 1 is a purely historical discussion of Gottlob Frege’s actual views on the structure of thought. I begin by tracing Frege’s evolving conception of content from its beginnings in the *Begriffsschrift* to his eventual distinction between sense and reference. I argue that Frege did *not* hold that a thought has a unique decomposition into its constituent senses (as many

---

<sup>4</sup>Note that thoughts (also known as “Fregean propositions”) remain widely discussed in the philosophical literature. Cf. Christopher Peacocke (2008) and Tyler Burge (2010) as paradigm examples.

authors assume), but rather held that the same thought can be decomposed into senses in a variety of distinct ways. I show that this is the natural extension of his earlier views.

Here is a simple example: Frege wanted pairs of sentences such as ‘John loves Mary’ and ‘Mary is loved by John’ to express the same thought. But the “loves” relation is distinct from its inverse: the “is loved by” relation. So the sense expressed by ‘loves’ is distinct from the sense expressed by ‘is loved by’, since their referents are distinct. Hence, if the sentences express the same thought, this thought must be decomposable into senses in at least two distinct ways. In effect, Frege wanted to characterize thoughts at a level of abstraction at which the transition from active to passive voice doesn’t affect the thought expressed. Any difference in meaning between the two sentences is irrelevant to determining their shared truth-value, and hence such differences shouldn’t arise at the level of abstraction we’re considering.<sup>5</sup>

Chapter 1 continues by examining a puzzle discussed in the Frege literature, stemming primarily from Michael Dummett (1981): how could a thought both be built up out of senses as parts and yet be decomposable into senses in distinct ways? If an entity is built up out of parts, then doesn’t it have a unique decomposition into the parts from which it could be constructed?

In response, I first make an interpretive point that clarifies some aspects of Frege’s account of the structure of thought. I show that Frege was committed to using his own function/object distinction to analyze the structure of thoughts. This is

---

<sup>5</sup>This example isn’t intended to convince the reader of the need for multiple decompositions of thought; I merely provide it as an especially simple case of the sort of thing Frege had in mind.

because in an indirect context, Frege believed that expressions refer to the *senses* they ordinarily express (1892a, pg. 28/159). For example, if I assert ‘Copernicus believed that the planetary orbits are circles’, the expressions ‘the planetary orbits’ and ‘are circles’ refer to their ordinary senses. But the expression ‘are circles’ is unsaturated, and hence it must refer to a function on Frege’s view. In particular, it must refer to a function from the referent of ‘the planetary orbits’ to the referent of the complete expression ‘the planetary orbits are circles’. But the referent of this complete expression in this context is the thought ordinarily expressed by it, and hence the ordinary sense of ‘are circles’ not only *refers* to a function, it *is* a function from senses to thoughts which contain it as a part.

I then provide various quotations that decisively show that Frege thought of the part/whole distinction as being such that a whole can be decomposed into (or built up out of) parts in a variety of distinct ways. Thus the puzzle dissolves.

I conclude Chapter 1 by emphasizing the connection between multiple decompositions and potential inferences: grasping a thought in a particular way makes a variety of potential inferences salient while masking others. The Appendix to Chapter 1 includes every quotation that I consider to be relevant to the Multiple Decompositions Thesis, and is meant to decisively establish that this controversial interpretation is correct.

The motivation behind Chapter 1 is not, of course, merely to correct various misinterpretations of Frege. I suggest that the aspects of Frege’s account that I have emphasized are part of a correct account of the structure of thought. This will become especially clear in Chapters 2 and 3, in which I argue that the treatment

of the structure of thought emphasized in Chapter 1 allows us to resolve various well-known issues concerning the hierarchy of senses, demonstratives, and indexicals.

Chapter 2 focuses on the hierarchy of senses. As pointed out above, Frege thought that in indirect context, expressions refer to their ordinary senses. They thereby express *indirect* senses. But in a doubly-embedded indirect context (e.g. ‘Frege believed that Copernicus believed that the planetary orbits are circles’), expressions will refer to their *indirect* senses and express *doubly* indirect senses. Iterating this procedure leads to an infinite hierarchy of senses.

Several authors have criticized this aspect of Frege’s view, with Donald Davidson arguing that it would make language unlearnable (1965, pg. 14-15 in the 1984 reprint). Tyler Burge (2005), Christopher Peacocke (2008), and Saul Kripke (2008) all defend Frege against Davidson and others, and I discuss each author’s resolution in turn (all argue that the hierarchy can be completely specified via a recursive reference rule, and is thus learnable).

I point out, however, that Frege’s own view was more complex than it may appear and in fact relies on the Multiple Decompositions Thesis. In particular, I argue that Frege believed that any thought that  $p$  is identical with the thought expressed by ‘the thought that  $p$  is true’. Hence, any thought can be decomposed into a sense referring to the complete thought and the sense of ‘is true’, a sense which maps the sense of ‘the thought that  $p$ ’ to the thought that  $p$ .

Furthermore, Frege says that the ordinary sense of ‘the thought that  $p$ ’ is the *indirect* sense of ‘ $p$ ’, and hence an indirect sense which refers to a thought is a part of that very thought. I also point out that the structure of an indirect sense mirrors the

structure of the sense it refers to. Compare a map of a region that is located within that very region: e.g. a shopping mall with a map including a point labeled “You are *here*”. I discuss the indexical ‘here’ in the following chapter.

In order to motivate this somewhat technical view of the structure of thoughts about thoughts, I point out connections with the structure of conscious experience. Note that whenever I reflect on the totality of my own conscious experience, simply by virtue of doing so, a part of that experience is representing the whole. Hence this structure exists in meta-representation of our own conscious experience, and hence it should be somewhat unsurprising that it is preserved in thoughts about thoughts.

I also argue that reflection on various cases of indirect contexts suggests that *both* ordinary senses *and* indirect senses are expressed in such contexts, and furthermore are expressed in a way that makes their connection salient. I then attempt an explanation, starting with the following question: why is it so useful to be able to represent our conscious experience as structured in various ways? I suggest the answer is: to aid in conscious decision-making leading to various actions, including mental actions such as judging, inferring, etc.

In particular, if I represent my conscious experience as structured in such a way that aspects of it represent various objects instantiating properties and bearing relations to each other, then I can *compare* how things appear to me with various additional pieces of information I have about how they actually are. Without this meta-representational ability, reflectively judging whether the appearances match reality would be impossible. I suggest that it is in service of this ability that we are capable of meta-level reflection on our own beliefs and the beliefs of others, and hence



in the context of a belief ascription I should not only represent the thought ascribed, but also the relation between that thought and what the thought is about, and that is why both ordinary senses and indirect senses are expressed in a way that makes their connection salient.

I conclude Chapter 2 by demonstrating that the account I've given, which is a modification of Frege's own account, resolves various general issues involving the hierarchy of senses that are not resolved by the accounts of Burge, Peacocke, or Kripke.

Chapter 3 turns to demonstratives and indexicals, traditionally a problematic domain for Fregean theories of content. I begin by developing and criticizing a naïve Fregean theory of the contribution of contextual factors to the expression of thoughts. On such a view, tensed thoughts include senses as components referring to particular times (e.g. the time of utterance), and thoughts have unique decompositions into senses. I point out that such an account would make preservation of thought over time impossible. In effect, one would be unable to grasp a tensed thought at any time other than the time at which it was originally expressed.<sup>6</sup> Naturally, I argue later in the chapter that we can resolve this issue if we allow thoughts to have multiple decompositions into senses.

I also discuss the well-known modal objections stemming from Saul Kripke (1980) and David Kaplan (1989). In effect, demonstratives and indexicals seem to behave as if they have the same reference even when evaluated with respect to other possible

---

<sup>6</sup>This point was made by Kripke (2008), although he (incorrectly) interprets Frege as committed to this view. I extend Kripke's point and consider various possible responses.

states of the world; they appear to be *rigid designators*. It isn't clear how a Fregean theory of context can account for this.

I turn to interpreting Frege's actual views on demonstratives and indexicals. I discuss an interpretation due to Kripke (2008) and defend it against criticism from Wolfgang Künne (2010) and Tyler Burge (2012). On Kripke's interpretation, when I utter 'I am here now', the subject, place, and time of utterance themselves are used as names of themselves, thereby expressing autonomous senses. In effect, one is using aspects of the context as self-referential pieces of language. I show that this interpretation is well-supported by the textual evidence.

Frege wanted the same tensed thought to be expressible at different times, claiming that 'Today it is raining' uttered one day and 'Yesterday it was raining' uttered at the same place the next day express the same thought.<sup>7</sup> I point out that Frege appealed to multiple decompositions of thought here. I then develop this view and show that Frege was committed to e.g. the sense expressed by the combination of 'Today' and the time of utterance having the structure of a *timeline* and the sense expressed by the combination of 'here' and the place of utterance having the structure of a *space*.

Before defending this view as correct, I address a well-known objection due to Kaplan (1989): someone might not realize how many days have passed since their original utterance, and thereby fail to recognize whether their new utterance expresses the same thought or a distinct thought, even if they are an ideally linguistically competent speaker. I point out that this consequence is acceptable provided that

---

<sup>7</sup>Kripke (2008) claims that Frege shouldn't have said this, given his other commitments. I show that Kripke is mistaken, stemming from his failure to recognize the Multiple Decompositions Thesis.

we take the truth-conditional conception of thought as fundamental, as I have been doing throughout the dissertation.

Returning to the view of the structure of senses outlined above, I point out that when I assert that today it is raining here, I am still committed to its having rained at that location on that day when I consider it from the perspective of tomorrow. The truth-conditions of my assertion do not depend on the temporal perspective from which I am viewing the event.

Note that the spatial analogue of a shift in one's temporal perspective on an event as occurring over a particular duration of time is a shift in spatial perspective on an object as having a particular geometric shape. I point out that a perception of an object as a cube carries commitments to how that object would appear from a different perspective. I suggest that an elegant way to capture this idea is to let the representational content expressed by that perception have the structure of a space with that object as origin: any point in that space corresponds to a particular perspective on that object from which I would still perceive it as a cube. This captures my commitments regarding how that object would appear from another angle, and allows me to explain how another perceiver at another point in space could be perceiving the object in the same way.<sup>8</sup>

I claim that the position on the structure of senses I outlined above is the natural extension of this view from the realm of perception to the realm of thought. The sense

---

<sup>8</sup>Obviously, this notion of *way of perceiving* is at a particular level of abstraction. I take it to be clear that there is a level of abstraction at which two people in the same room can perceive a cubic object in the same way, with both carrying commitments to how that object appears from the other's perspective.

expressed by the combination of an utterance of ‘here’ and the place of utterance has the structure of a space in which each point in the space corresponds to a perspective on the place referred to by ‘here’. The shift in perspective is captured at the level of sense via multiple decompositions: the very same sense can be grasped as a function of various points in the space while the full sense remains the same even as I shift my perspective. The analogous remarks with respect to time allow two utterances of ‘Today it is raining here’ and ‘Yesterday it was raining here’ to express the same thought.

I then turn to the modal objections. I point out that on the Fregean view I’ve outlined, indexicals are rigid designators after all, maintaining fixed reference even when evaluated with respect to different possible states of the world. This is due to the autonomous senses expressed by subjects, places, and times of utterances used as names of themselves. I note that this account could be extended to a Fregean treatment of demonstratives and proper names as rigid designators.

This may feel like a mere technical trick. Is there any reason to think that subjects, times, and places are actually used as names of themselves? I suggest that using entities as representations of themselves is ubiquitous in thinking, and (very tentatively) suggest a model of consciousness of our own perceptions under which those perceptions are used as representations of themselves (recall that I discussed consciousness of one’s own perceptions in the previous chapter in connection with the hierarchy of senses).

The tentatively-suggested model is the following: an especially simple way of constructing a self-referential representation is combining a representation of an entity

with that entity. My ability to recall my conscious experience a moment ago implies that my mind is constantly constructing meta-representations of my conscious experience. If such a meta-representation were combined with my conscious experience itself in the appropriate way, each element of that experience could be treated as a representation of itself. This captures the self-referential phenomenology of human consciousness, and in the particular case of perception explains why my perceiving an object and my conscious awareness of that perception have a similar phenomenology (in the latter case I am using that very perception to represent itself). Furthermore, this model also explains why I never have such meta-representations of my own perceptions without having those perceptions, and why there is a sense in which those meta-representations are immune to certain sorts of error.

I close Chapter 3 by returning to the issues of thought preservation outlined at the chapter's beginning, arguing that the view I've defended gives the correct answer in the cases I've discussed.

Chapter 4 discusses logical form. I begin by arguing that my treatment of thoughts as structured entities allows me to characterize deductive inference as a transition between thoughts whose guaranteed truth-preservation can be explained solely by appeal to logical forms instantiated by those thoughts. I point out that part of the value of a contemporary formal language is that it allows us to constrain the semantics so that we can only give an explanation of truth-preservation by appealing to structural features of truth-conditions (and the corresponding thoughts) corresponding to sentences of the formal language. The question, then, is which structural features of thoughts ought to count as *logical* forms.

I turn to defending Frege's position that complete declarative sentences refer to special objects: the True and the False. I begin by asking the question: why is it so cognitively useful for us to see the world as made up of objects?

I point out that in Chapter 3 I argued that a perception of an object as a cube carries commitments to how that object would appear from other points in space. I suggested capturing these commitments by treating the representational content of such a perception as a space of possible perspectives on that object. I then point out that this space has quite a bit of structure to it, since e.g. I am committed to the appearance of the cube shifting in a determinate way as I move closer to it (i.e. it will take up more of my visual field), and this is grounding various expectations of mine regarding how things will appear as I shift my position in space.

With a nod to Immanuel Kant (1781/1787), I argue that the cognitive value of representing something as an object is that it allows me to unite various representations as of the same thing in accordance with particular rules. In the visual perceptual case, this grounds my expectations that my visual experiences a few moments from now will be anything like my visual experiences a moment ago (there is a thing in the world that I'm shifting my perspective on), and that your visual experiences are currently similar to mine (e.g. we're both perceiving the same cube).

Turning back to truth-values, I argue that when I make a sincere assertion, I'm attempting to describe the world as it actually is. I consequently suggest treating the True as the actual world, the intended referent of assertion.<sup>9</sup> I point out that many of

---

<sup>9</sup>I also point out that Frege himself seems to have held this position in his (1892a), although he seems to have given it up later.

our inferences presuppose that we're experiencing the *same* world over time, as does our assumption that someone else's experiences are anything like our own. Much like the more general case of objects, I argue that truth-preserving inferences between thoughts maintain fixed reference to the actual world while varying its representation in accordance with particular rules. These rules are grounded on its being the *same* world that one is continuing to represent.

I then turn to atomic form. I argue that being able to learn from experience requires making a distinction between particular things and *patterns* or *structures* that multiple things have in common. More generally, I argue that learning from experience requires making something like an object/property/relation distinction, which is thus itself unlearnable because any learning presupposes it.

I then give several arguments that the object/property/relation distinction is necessary for the possibility of rational agency. For one, inferential transitions between thoughts require distinguishing between particular thoughts and general structures instantiated by those thoughts, with the relevant structures determined by the rules governing the inferential transitions.

I claim that if a structure is both unlearnable and necessary for the possibility of rational agency then it ought to be considered a logical form. In effect, such structural features correspond to the structure that one would have to put in place if one were designing a rational mind, and in that sense can be thought of as fundamental to the structure of thought. I spend the rest of the chapter considering whether the structures of thoughts corresponding to the sentences of classical first-order logic satisfy these two conditions.

Following a classic argument due to Lewis Carroll (1895) and a more recent argument due to Saul Kripke, I argue that both *Modus Ponens* and Universal Instantiation are such that you couldn't come to recognize them as truth-preserving and make inferences grounded on that recognition without already inferring in accordance with them.<sup>10</sup> I consider whether the corresponding logical connectives are both unlearnable and necessary for the possibility of rational agency, and consider similar arguments regarding the other classical logical connectives.

I close the chapter by returning to the hierarchy of senses discussed in Chapter 2. I argue that this structure is not in fact necessary for the possibility of rational agency, although it is a natural extension of primitive rational agency to the level of meta-representational capacities. I suggest that finding other natural extensions of logical form (and other examples of logical form beyond classical first-order logic) would be a worthwhile enterprise.

Note that while there are many interconnections between the four chapters, they were written in such a way that they can be read independently of each other with minimal loss of understanding. For example, Chapter 3 takes various aspects of the Frege interpretation defended in Chapter 1 for granted, but could still be read independently of Chapter 1 provided the reader is willing to grant me those aspects for the sake of argument. The overarching theme of the dissertation remains, however: the structure of thought.

---

<sup>10</sup>Regarding Kripke's argument, cf. (Padro 2015): imagine trying to teach someone to infer in accordance with Universal Instantiation. They take it on your authority that every instance of Universal Instantiation is truth-preserving. However, when confronted with a particular instance of Universal Instantiation they remain skeptical. The point being, getting from the universal claim to a particular instance is itself an application of Universal Instantiation.



### 0.3 A Note on Translation

Much of this dissertation will involve interpreting the work of Gottlob Frege. When providing quotations, I rely on English translations of Frege’s original German essays and correspondence, contained in (Frege 1979), (Frege 1980), and (Frege 1984). I also use the T. Bynum translation of Frege’s *Begriffsschrift* (1879), the J. L. Austin translation of Frege’s *Die Grundlagen der Arithmetik* (1884), and the P. Ebert and M. Rossberg translation of both Volumes I and II of Frege’s *Grundgesetze der Arithmetik* (1893/1903a). Whenever a work of Frege appeared in print, I will cite that work’s original pagination followed by its pagination within the translation I’m relying on: e.g. ‘(Frege 1892a, pg. 35-6/164-5)’.

Since some basic terms have been translated in different ways, let me note the following: I translate ‘Sinn’ as ‘sense’, ‘Bedeutung’ as ‘referent’ or ‘reference’, and ‘Satz’ as ‘sentence’ or ‘theorem’ (depending on the context). Whenever I deviate from this practice I include a footnote with an explanation. Whenever I have made my own adjustment to a cited translation, I have included that adjustment within brackets: e.g. ‘[referent]’. I occasionally make adjustments in order to account for certain cases of anaphora, e.g. replacing ‘it’ by ‘[language]’. Finally, in some cases I have felt the need to add context to a quotation (e.g. to explain what Frege was referring to). I also use brackets for this purpose, including my initials within the brackets to indicate that it is my own addition: e.g. “Such a designation [‘the negation of the negation of  $A$ ’—NB] is to be regarded as...”.<sup>11</sup>

---

<sup>11</sup>If the reader is confused by any of these uses of brackets, I encourage them to consult either the referenced English translation or the original German.

# Chapter 1

## The Decomposition of Thought

### 1.1 Introduction

Gottlob Frege famously introduced the notion of the *gedanke* or *thought* expressed by a complete declarative sentence as the sort of content relevant to determining whether that sentence is true or false. Here is one of many examples: “So one has to separate off from the content of a sentence the part that alone can be accepted as true or rejected as false. I call this part the thought expressed by the sentence.” (1906c, PW pg. 197-8)<sup>1</sup>

Furthermore, when discussing the structure of thoughts, Frege helped himself to mereological language, often speaking of parts of a thought corresponding to the significant parts of a sentence expressing it.<sup>2</sup> For example, when I utter the sentence

---

<sup>1</sup>In this discussion I follow Frege in treating *content* as an intuitive notion from which the more technical notion of *thought* can be isolated.

<sup>2</sup>For example: “How can [language] achieve so much? By virtue of the fact that thoughts have parts out of which they are built up. And these parts, these building blocks, correspond to groups of sounds, out of which the sentence expressing the thought is built up, so that the construction of

‘John has brown hair’, my utterance expresses the thought that John has brown hair, a thought which contains the sense of ‘John’ and the sense of ‘has brown hair’ as parts.<sup>3</sup>

The goal of this chapter is to further elucidate Frege’s conception of the structure of thoughts. I will argue that Frege held that a thought does not have a unique decomposition into senses, but rather the same thought can be decomposed into senses in many distinct ways. This Multiple Decompositions Thesis will be especially important in later chapters.

This chapter is divided into two parts. In the first part, I will trace Frege’s evolving conception of content from his first reflections in the *Begriffsschrift* of 1879 to his eventual distinction between sense and reference in 1891. In the second part, I will consider some potential problems with the Multiple Decompositions Thesis noted in the literature. The Appendix to this chapter includes a list of every quotation I consider to be relevant to the Multiple Decompositions Thesis, and is meant to decisively establish that this controversial interpretation of Frege is correct.

---

the sentence out of parts of a sentence corresponds to the construction of a thought out of parts of a thought. And as we take a thought to be the sense of a sentence, so we may call a part of a thought the sense of that part of the sentence which corresponds to it.” (1914, PW pg. 225)

Frege explicitly calls such mereological language “figurative” in the following passage: “To be sure, we really talk figuratively when we transfer the relation of whole and part to thoughts; yet the analogy is so ready to hand and so generally appropriate that we are hardly ever bothered by the hitches which occur from time to time.” (1923-6, pg. 36/390) However, this passage should be taken with a grain of salt; this was one of Frege’s last papers, and it is unclear whether Frege’s earlier positions are really compatible with such a noncommittal account of the mereological structure of thoughts.

<sup>3</sup>We will discuss Frege’s notion of “sense” (“Sinn”) throughout this dissertation. Briefly: I take the most fundamental notion of the sense of an expression to be the aspect of the expression’s content in virtue of which it contributes to the truth-conditions of complete declarative sentences which contain it.

## 1.2 Logic, Formal Languages, and Conceptual Content

Frege's primary goal in constructing his *Concept-Script* (*Begriffsschrift*) of 1879 was to demonstrate that arithmetical reasoning, which appeared to depend essentially on intuition or insight, was in fact reducible to a sequence of purely logical inferences. Frege realized that in order to do this he needed to develop a formal language in which all hidden presuppositions are made explicit and even the smallest step in a chain of reasoning is represented; inferences needed to be entirely "free of gaps".

Regarding his formal language, Frege states in the Preface that "I have omitted the expression of everything which is without importance for the chain of inference." (Frege 1879, pg. iv/104) His motivation is clear: Frege intends to ensure that anything that is not relevant to the logical structure of a chain of inferences is simply not expressed by sentences in his formal language. That way, he can avoid the vagaries and imprecision of natural languages, along with the various aspects of linguistic meaning that are irrelevant to logic. Frege calls what *is* expressed by the expressions of his formal language *conceptual content*,<sup>4</sup> where the conceptual content of complete declarative sentences is also known as *judgeable content*.<sup>5</sup>

---

<sup>4</sup>Note that Frege uses this expression more broadly to apply to e.g. *geometric* content as well, although this generalization won't be relevant to our present purposes. I am grateful to Haim Gaifman for pointing this out to me.

<sup>5</sup>Since this is prior to Frege's sense/reference distinction, Frege was a bit sloppy about distinguishing the content *expressed* by an expression from the content *represented* by (or *referred to* by) that expression. Perhaps it would be better to say, as Frege wrote to Husserl, for example (Letter to Husserl 5/24/1891, PMC pg. 63), that conceptual content is a fusion of sense and reference, and hence the content expressed and the content represented are identical. This issue will reappear throughout this section.

Of course, one may think that this distinction between expression and representation is unfair

In Frege's view, past logicians relied too much on the grammar of natural languages and thereby failed to recognize that the *logical* structure of (the content expressed by) a sentence or collection of sentences in an inference is often quite different from its *grammatical* structure (Frege 1879, pg. vi/107). Hence, Frege's formal language is designed so that *its* grammar makes the logical structure of the contents expressed within it transparent. To do so, Frege replaced the traditional distinction between subject and predicate with his new distinction between argument and function. In fact, Frege considered his replacement of the standard subject/predicate analysis of content by an argument/function analysis to be one of his most important advances in logic. Let's briefly consider why.

Why does Frege think that the subject/predicate analysis of content failed to elucidate underlying logical structure? Consider a well-known example of an Aristotelian syllogism: "All men are mortal. Socrates is a man. Therefore, Socrates is mortal." On a subject/predicate analysis, the sentences 'All men are mortal' and 'Socrates is mortal' are distinguished only by their distinct subjects: 'All men' in the former and 'Socrates' in the latter.<sup>6</sup> On a subject/predicate analysis, these two

---

to early Frege, since it nearly takes for granted his later sense/reference distinction. Nothing that follows will hinge on these issues, although I will continue to speak of conceptual content as both expressed and represented.

<sup>6</sup>While it is also true that the two sentences differ in their usage of the verb "to be" (in English, at least), note that the following holds: (1) a sentence with a compound subject such as 'Aristotle and Socrates are mortal' also uses the plural form; (2) Frege is ultimately concerned with the respective roles of 'All men' and 'Socrates' in *deductive inferences*, and he will argue quite convincingly that they play fundamentally different roles.

One might argue that this focus on inferential role is the right approach as follows: why is it so useful for us to *recognize* the sentences we utter as structured entities? It is useful partially because it allows us to make *inferences* to and from an uttered sentence, inferences hinging on the object/property/relation distinction, for example. More generally, it allows us to recognize *relations* between sentences that obtain in virtue of such structure. We shall return to this issue in later chapters.

subjects are of the same logical type. But what is the meaning of the expression ‘All men’? Does it refer to the class of men, or some mereological fusion of all men, or the Platonic form MEN, or, unlike ‘Socrates’, is it simply not the sort of thing that refers at all?<sup>7</sup> Due in no small part to such considerations (although strictly speaking Frege wasn’t yet distinguishing *sense* from *reference*), Frege provides an alternative argument/function analysis of content that treats these two subject-expressions as having fundamentally different logical forms, ultimately stemming from their distinct roles in deductive inference.

In order to elucidate Frege’s alternative argument/function analysis of content, let’s consider the sentence ‘Hydrogen is lighter than carbon dioxide.’ Frege points out that we can replace the term ‘Hydrogen’ by, for example, ‘Oxygen’ in order to get a sentence that represents (expresses) a distinct conceptual content: ‘Oxygen is lighter than carbon dioxide.’ Yet, there is a strong connection between these two sentences. Frege calls the fixed component the *function*, and the replaceable component the *argument*.

Interestingly, Frege says, “This distinction has nothing to do with the conceptual content, but only with our way of viewing it.” (Frege 1879, pg. 15/126) We could have *also* considered ‘carbon dioxide’ to be the argument and ‘being heavier than hydrogen’ to be the function, or *both* ‘hydrogen’ and ‘carbon dioxide’ to be the arguments and ‘is lighter than’ to be the (two-place) function; this makes no difference to the conceptual content itself. Notice the sharp difference from an analysis based upon grammatical

---

<sup>7</sup>Note that although it may seem that Aristotle is the target here (and Frege himself suggests as much, although this explanation of Frege’s motivation is my own), Aristotle did not actually hold such a naïve view of the relation between logical form and natural language.

structure such as the subject/predicate analysis briefly discussed above.

Consider also the sentence ‘Cato killed Cato.’ According to Frege, this could be analyzed in at least three ways (ibid pg. 16/127): taking ‘Cato’ as argument and ‘killing Cato’ as function, taking ‘Cato’ as argument and ‘being killed by Cato’ as function, or taking ‘Cato’ as argument and ‘killing oneself’ as function (in the last case, we treat *both* instances of the term ‘Cato’ in the sentence as replaceable by a single argument). None of this affects the conceptual content, but merely our way of *viewing* or *grasping* it.

But what *are* the functions and arguments: the expressions themselves or the content of the expressions or something else entirely? Also, what do the functions map their arguments *to*? Regarding the first question, there is some textual evidence that in 1879 Frege considered functions and arguments to be parts of *expressions* rather than contents. Let’s consider Frege’s official statement about carving an expression into function and argument:

*If, in an expression (whose content need not be assertible), 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.”* (ibid pg. 16/127)

Here Frege explicitly calls functions and arguments parts of the *expressions* rather than the contents represented (expressed). Such a potential carving of the expression is suggested by the structure of the expression itself.

Further support for this interpretation comes from Frege’s correspondence with

Philip Jourdain. While Frege suggested many corrections to Jourdain’s essay *Gottlob Frege*, he interestingly did *not* correct Jourdain’s claim about Frege’s *Begriffsschrift*-era formalist sympathies: “Notice that here Frege, although he was careful to distinguish the sign ‘*A*’ from what it denoted, spoke of the *expression* as variable, and defined the function as a certain part of the *expression*.” (Jourdain 1912, reprinted in (Frege 1980, pg. 187 fn. 40)) It is possible, however, that this footnote was added to a later draft that Frege didn’t correct. Further evidence comes from Frege’s treatment of identity within the *Begriffsschrift*, which we will discuss shortly.

However, by 1881 he unambiguously treated functions and arguments as parts of the *content* represented (expressed) rather than the expressions themselves.<sup>8</sup>

Regarding the second question, Frege states “For us, the different ways in which the same conceptual content can be considered as a function of this or that argument have no importance so long as function and argument are completely determinate.”

(Frege 1879, pg. 17/128)<sup>9</sup> Frege also states that (outside of the special case of judg-

---

<sup>8</sup>“If, that is, 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.” (Frege 1880-81, PW pg. 16)

<sup>9</sup>Regarding the point about function and argument being completely determinate, Frege proceeds to give an example of a judgment in which the argument is *indeterminate*: “Whatever arbitrary positive integer we take as argument for ‘being representable as the sum of four squares’, the [resulting sentence] is always true.” (ibid pg. 17/128) The point Frege had in mind seems to be the following: Let’s represent this judgment within the *Begriffsschrift* as follows:  $\vdash \frac{\quad}{\quad} \Sigma(a)$ , where

‘ $a$ ’ is a Latin letter used to express generality (as Frege discusses in Sections 1 and 11), ‘ $N$ ’ is an abbreviation for ‘being a positive integer’, and ‘ $\Sigma$ ’ is an abbreviation for ‘being representable as the sum of four squares’. Certain carvings of this expression are ruled out by the content itself, since one cannot, for example, treat the first occurrence of ‘ $a$ ’ as replaceable while leaving the second occurrence fixed (thereby treating it as part of the function), as one could in the ‘Cato killed Cato’ example, since this would allow one to transition to the judgment: “the circumstance that the value of the function ‘being representable as the sum of four squares if 2 is a positive integer’ on *any* argument is a fact is itself a fact”, which implies that *any* argument is representable as the sum of four squares and obviously does not follow from the former judgment.



ments of identity of content, on which more in a moment) “symbols are usually only representatives of their contents—so that each combination [of symbols usually] expresses only a relation between their contents...” (ibid pg. 13/124)

Consider also Frege’s remarks when describing his language, e.g. “ $\vdash$   $\begin{array}{l} \text{---} A \\ | \\ \text{---} B \end{array}$

*stands for* [my emphasis—NB] the judgement that *the third of these possibilities* [i.e. that *A* is denied and *B* is affirmed—NB] *does not occur, but one of the other three does.*” (ibid pg. 5/115) The point being, expressions within Frege’s *Begriffsschrift* do *not* merely represent objects and functions, as they will in Frege’s later work, but rather *conceptual contents grasped in particular ways*. In particular, the expression in the previous quote stands for the following judgment:<sup>10</sup> “the circumstance that the circumstance that *B* bears the material conditional relation to the circumstance that *A* is a fact”. So, varying the carving of an expression into functions and arguments is varying one’s *way of representing* (or *way of expressing*) the conceptual content represented (expressed) by the full expression.

Consider also: “It is easy to see how regarding a content as a function of an argument leads to the formation of concepts.” (ibid pg. vii/107) These quotes suggest that the function maps its argument(s) to a conceptual content, namely, the conceptual content represented (expressed) by the complete expression initially carved into function and argument. Note that even in a case in which one uses an unstruc-

---

<sup>10</sup>Well, it *would* stand for a judgment, provided that one specified what the Greek letters ‘*A*’ and ‘*B*’ are abbreviations *for*, in accordance with Frege’s footnote at the beginning of Section 2. Note also that Frege is referring to the *expression* within his formal language as opposed to the judgment represented (expressed) by it. In this chapter I also slide between using and mentioning expressions involving the judgment stroke in order to avoid unnecessary complications. I expect the reader to determine the intended meaning from the context.

tured expression such as ‘ $A$ ’ in order to represent a judgeable content, one can treat that expression as the argument of the content stroke function, which maps it to the judgeable content represented (expressed) by ‘ $A$ ’ (i.e.  $\ulcorner A \urcorner$ ). Just as the judgment stroke can be seen as an abbreviation for the predicate ‘is a fact’ (ibid pg. 4/113), the content stroke can be seen as an abbreviation for the function ‘the circumstance that’ (ibid pg. 2/112).

With that said, let’s turn to the interesting case of identity judgments. In Section 8, Frege made several seemingly-contradictory claims about identity judgments:

Identity of content differs from conditionality and negation by relating to names, not to contents. Although symbols are usually only representatives of their contents—so that each combination [of symbols usually] expresses only a relation between their contents—they at once appear *in propria persona* [i.e. as themselves—NB] as soon as they are combined by the symbol for identity of content, for this signifies the circumstance that the two names have the same content. Thus, with the introduction of a symbol for identity of content, a bifurcation is necessarily introduced into the meaning of every symbol, the same symbols standing at times for their contents, at times for themselves. (ibid pg. 13-14/124)

Thus, the need of a symbol for identity of content rests upon the following fact: the same content can be fully determined in different ways; but, that *the same content*, in a particular case, is actually given by *two {different} modes of determination* is the content of a *judgement*. (ibid pg. 14/125)

But how could expressions in the context of an identity judgment *both* represent (express) themselves *and* represent (express) distinct modes of determination of the same content?

I claim that Frege held the following position:  $\ulcorner A \equiv B \urcorner$  represents (expresses) the judgment that the circumstance that the *expression* ‘ $A$ ’ and the *expression* ‘ $B$ ’

represent (express) the same conceptual content is a fact (under the presupposition that what these Greek letters are abbreviations for has been specified). Each expression is intended as an abbreviation for a way of determining a conceptual content. In the context of an identity-judgment, these expressions represent (express) themselves *qua* ways of determining conceptual contents. On my interpretation, this is clearer in the case of a structured expression such as ‘7 + 5’. This expression is *itself* a way of determining the number 12, since the function ‘+ 5’ is part of the very expression, and when applied to the remainder as argument (the numeral ‘7’) gives the number 12 as value. In the context of an identity judgment such as  $\vdash 7 + 5 \equiv 12$ , ‘7 + 5’ is used to represent itself, a way of determining a conceptual content.<sup>11</sup>

As mentioned above, this is further evidence that in 1879 Frege considered the function/argument distinction to apply at the level of *expression* rather than content.<sup>12</sup> However, since by 1881 Frege did not hold this view, the correctness of such

---

<sup>11</sup>For further support of this interpretation, consider the opening remarks of “On Sense and Reference”, in which Frege asserts both that in the *Begriffsschrift* he assumed identity was a relation between *names* or *signs* of objects and “In that case the sentence  $a = b$  would no longer refer to the subject matter, but only to its mode of designation; we would express no proper knowledge by its means.” (Frege 1892a, pg. 26/157)

<sup>12</sup>Note also that on this interpretation Frege is committed (circa 1879) to a way of determining the conceptual content of a complex expression proceeding top-down. Given an expression such as ‘(7 + 5) + 7’, the value of the function ‘+ 5’ on input ‘7’ is the conceptual content represented (expressed) by ‘(7 + 5)’. However, the function ‘+’ takes the *expression* ‘(7 + 5)’ as input as opposed to its conceptual content. So, one can’t reach this conceptual content bottom-up by first applying the ‘+ 5’ function to ‘7’ and then applying the ‘+’ function to both the result of the former application and ‘7’. Instead, one arrives at this carving by first carving the complete expression into ‘(7 + 5)’ and ‘7’ as arguments with ‘+’ left as fixed remainder. One can then further carve the expression ‘(7 + 5)’ into ‘7’ as argument and ‘+ 5’ as fixed remainder, or (for example) into ‘7’ and ‘5’ as arguments and ‘+’ as fixed remainder.

In effect, one is initially given the conceptual content of the complete expression. One then considers a way in which that expression can be viewed as determining that conceptual content, and one possible answer is that this content is the value of the function ‘+’ on inputs ‘(7 + 5)’ and ‘7’. The conceptual content of the complete expression also depends upon the conceptual content of ‘(7 + 5)’, and hence one can also look for a way in which *that* expression can be viewed as determining its own conceptual content. The number of times one can continue this process depends upon the structural complexity of the expression under consideration.

a formalist interpretation will be unimportant in the remainder of this chapter.

Now we are in a position to understand some of the advantages of the argument/function analysis of content. Recall that on a subject/predicate analysis, in the sentences ‘All men are mortal’ and ‘Socrates is mortal’, the expressions ‘All men’ and ‘Socrates’ are treated as being of the same logical type. They play very different roles in inference, however. For example, from ‘Socrates is mortal’ I can infer ‘Someone is mortal’ (this is often known as Existential Introduction), but I cannot infer this from ‘All men are mortal’.<sup>13</sup> Furthermore, from ‘All men are mortal’ I can infer ‘If Bob is a man, then Bob is mortal’ (this is often known as Universal Instantiation), and I can make a similar inference involving any other proper name in the place of ‘Bob’, but I cannot infer such things from ‘Socrates is mortal’.<sup>14</sup> As Frege will emphasize after his sense/reference distinction, ‘All men’ doesn’t behave like a referring expression at all.

Frege captures this distinction within his formal language by treating ‘is mortal’ as a function that maps ‘Socrates’ as argument to the judgeable content that Socrates is mortal as value, while the judgeable content represented (expressed) by ‘All men are mortal’ can instead be viewed as the value of the generality function applied to the *function* corresponding to ‘if  $x$  is a man, then  $x$  is mortal’ as argument. The

---

<sup>13</sup>Here I am making the standard contemporary assumption that a universally-quantified conditional doesn’t carry an existential commitment. Hence, one needs to add the premise ‘There are men’ in order to logically imply the conclusion ‘Someone is mortal’. As should be clear, disagreement on that point isn’t relevant to the more general distinction between the inferential roles of the two expressions.

<sup>14</sup>To be fair, one might think that from ‘Socrates is mortal’ one could infer ‘If Bob is Socrates, then Bob is mortal.’ But note the distinct inferential roles of ‘Bob is a man’ and ‘Bob is Socrates’. From ‘Bob is Socrates’ and ‘Jim is Socrates’ I can infer ‘Bob is Jim’, but from ‘Bob is a man’ and ‘Jim is a man’ I *cannot* infer ‘Bob is Jim’.

generality function maps a function of one argument to a judgeable content that is a fact exactly if that function of one argument always maps its argument to a judgeable content that is itself a fact, no matter what that argument is. In this way, Frege can capture the distinct inferential roles of such expressions via his formal language. Furthermore, Frege *ought* to do so since the grammar of his language is intended to make the logical structure of its expressions transparent, and such logical structure is determined entirely by the role such expressions play in logical inference.<sup>15</sup>

Let's return to our discussion of multiple decompositions of content. Notice that Frege's formal language is *not* set up to represent judgeable contents *simpliciter*, but rather judgeable contents plus particular *ways* of carving their representations up into arguments and functions.<sup>16</sup> When expressing 'Cato killed Cato' in Frege's formal language, for example, we must make a choice about what we will treat as the argument(s) and what we will treat as the function(s). For instance, we might choose to represent this content as '—KilledCato(Cato)'.<sup>17</sup> However, not all logical inferences

---

<sup>15</sup>Incidentally, I think that Frege's implied *meta*-level reading of judgments with universal content (as judgments about whether the values of functions on arbitrary arguments are *themselves* always facts, cf. the beginning of Section 11) has been insufficiently discussed. Strictly speaking, such a judgment concerns a judgeable content that is *itself* about the result of applying a function to arbitrary expressions as arguments, namely, about whether or not it will always result in a fact. Hence, such a content is about whether arbitrary expressions of a particular form always represent (express) facts.

Of course, a possible alternative to this reading is that Frege was simply being a bit sloppy, and he certainly *does* revise his views on quantification, functions, etc. later. Still, I think that it is important to recognize areas in which Frege's conception of logic, even merely circa 1879, seems to be quite different from standard contemporary conceptions of logic. Furthermore, given that in Frege's later work the universal quantifier is a 2nd-level function from functions to truth-values, there is a natural continuity from his early conception to his later conception.

<sup>16</sup>One *can* of course leave a judgeable content unanalyzed, like an atomic sentence in propositional/sentential logic. One would then represent the judgeable content in Frege's system as an unstructured entity (Frege himself uses capital Greek letters). However, one would then obviously not be able to formally represent logical inferences dependent upon the structure being ignored.

<sup>17</sup>It is perhaps better to represent it as '— $A(c)$ ', where 'A' and 'c' are treated as abbreviations in the obvious way.

from the judgeable content represented (expressed) by ‘Cato killed Cato’ will be salient once we have chosen a particular way to carve up this representation. For example, although this judgeable content logically entails that there exists someone *killed by* Cato, one cannot represent such an inference *formally* once one has chosen to carve this representation in the way we did above: with ‘Cato’ as argument and ‘killed Cato’ as function. Instead, one could formally represent the inference from this judgeable content to the judgeable content represented (expressed) by the sentence ‘Someone *killed* Cato’. This point applies to Frege’s formal language across-the-board.

Importantly, when representing a conceptual content in such a formal language, one must choose to represent it in such a way that will make the possibility of certain inferences salient and mask the possibility of certain others. Proper formal representation is, at bottom, *relative* to the desired inferential structure of the argument(s) being considered/constructed. *In the context of making an inference, we must grasp the contents involved in particular ways.*

Frege also discusses multiple carvings and their role in concept-formation in his Letter to Marty of 1882:

Now I do not believe that concept formation can precede judgement because this would presuppose the independent existence of concepts, but I think of a concept as having arisen by decomposition from a judgeable content. I do not believe that for any judgeable content there is only one way in which it can be decomposed, or that one of these possible ways can always claim objective preeminence. In the inequality  $3 > 2$  we can regard either 2 or 3 as the subject. In the former case we have the concept ‘smaller than 3’, in the latter, ‘greater than 2’. We can also regard ‘3 and 2’ as a complex subject. As a predicate we then have the concept of the relation of the greater to the smaller. (Frege, Letter to Marty 8/29/1882, PMC pg. 101)

Note that Frege now discusses decomposing judgeable contents *themselves* rather than merely *expressions* for such judgeable contents. Presumably one carves the judgeable content *via* carving its expression.<sup>18</sup> However, it isn't until the *Grundlagen der Arithmetik* that Frege really starts to put this conception of content to work, and it is there that we will turn next.

### 1.3 Some Remarks on the *Grundlagen*

In *The Foundations of Arithmetic* of 1884, Frege gives a detailed philosophical analysis of the concept of number, with the aim of philosophically motivating his project of reducing arithmetic to logic. It would take us too far afield to consider the essay in detail, so instead we'll focus on the aspects of the essay most relevant to our general discussion of Frege's evolving conception of content.

Frege opens the *Grundlagen* by pointing out that although everyone uses and speaks about numbers all the time, it seems that no one, not even a mathematician, can give a satisfactory answer to the question 'What is the number 1?' or 'To what does the numeral '1' refer?' Frege says that one common response is 'The number one is a thing.' Yes, but *which* thing is it? Frege points out that if everyone were allowed to treat the number one as whichever thing he or she liked, then "the same [sentence] about one would mean different things for different people,—such [sentences] would

---

<sup>18</sup>Compare Frege's claim in "Function and Concept" that "We recognize the function in the expression by imagining the latter as split up, and the possibility of thus splitting it up is suggested by its structure." (Frege 1891, pg. 7/141) Consider also the quotation from "Boole's Logical Calculus and the Concept-Script" cited in footnote 8 and Frege's discussion of concept acquisition immediately following that passage.

have no common content.” (Frege 1884, pg. i) Note the underlying interest in the *objective, shareable* content of mathematical assertions in a public language.

Here are two quotations that reveal Frege’s emphasis on the *generality* of laws of thought and his insistence on an anti-psychologistic treatment of the content of mathematical assertions:

Thought is in essentials the same everywhere: it is not true that there are different kinds of laws of thought to suit the different kinds of objects thought about. Such differences as there are consist only in this, that the thought is more pure or less pure, less dependent or more upon psychological influences and on external aids such as words or numerals, and further to some extent too in the finer or coarser structure of the concepts involved...(ibid pg. iii-iv)

Never let us take a description of the origin of an idea for a definition, or an account of the mental and physical conditions on which we become conscious of a [theorem] for a proof of it. A [theorem] may be thought, and again it may be true; let us never confuse these two things.<sup>19</sup> (ibid pg. vi)

At the end of the Introduction, Frege states his three fundamental principles (ibid pg. x):

- (1) always to separate sharply the psychological from the logical, the subjective from the objective;

---

<sup>19</sup>I have chosen to translate ‘satz’ as ‘theorem’ in this case, since that seems closer than ‘sentence’ or ‘proposition’ to Frege’s intended meaning. It is also worth noting that the word being translated here as ‘idea’ is ‘vorstellung’, which is often found in Kant and translated as either ‘representation’ or ‘presentation’. This choice of vocabulary reflects Frege’s primary criticism of Kant’s account of arithmetic; namely, that it is far too subjective. However, I think it is clear that Frege misunderstood Kant’s account, although perhaps he understood the so-called “Neo-Kantians” he interacted with quite well (cf. his “Dialogue with Pünjer on Existence” (before 1884)). In any case, Frege states that he will always use this word in the psychological, subjective sense.



- (2) never to ask for the meaning<sup>20</sup> of a word in isolation, but only in the context of a [sentence];
- (3) never to lose sight of the distinction between concept and object.

Principle (2) is known as the *Context Principle*. Frege immediately motivates it as follows: “If the second principle is not observed, one is almost forced to take as the meanings of words mental pictures or acts of the individual mind, and so to offend against the first principle as well.” (ibid pg. x). Later in this essay (ibid pg. 71), Frege says “Only in a [sentence] have the words really a meaning. It may be that mental pictures float before us all the while, but these need not correspond to the logical elements in the judgement. It is enough if the [sentence] taken as a whole has a sense; it is this that confers on its parts also their content.”

Note the suggestion once again that the content expressed by a complete, declarative sentence is more *objective* than the content expressed by a word or phrase in isolation. Why does Frege think this? Only the former sorts of content can be *true* or *false*, and truth and falsity are the primary markers of objectivity for Frege. Frege does not want to waste time considering what ideas different people associate with the word ‘chair’ (or what have you); he is only interested in its *objective* content in the public language, and he suggests that the content of an arbitrary expression is determined by the contribution it makes to the content expressed by complete,

---

<sup>20</sup>Although Frege uses ‘*bedeutung*’ here, I’ve chosen to follow both Michael Beaney (1997) and J. L. Austin in translating ‘*bedeutung*’ as ‘meaning’, since Frege had not yet made his sense/reference distinction. Tyler Burge (2005) argues that Frege would later accept *both* versions of the updated Context Principle: one version about *reference* and one version about *sense*.

declarative sentences containing that expression in such a language.<sup>21</sup>

Regarding Principle (3), Frege states: “it is a mere illusion to suppose that a concept can be made an object without altering it.” (ibid pg. x) Frege means that concepts are *essentially* predicative while objects are not. Objects are complete; concepts are *unsaturated* and must be saturated by an argument in order to form a judgeable content with a determinate truth-value.

Importantly, Frege also distinguishes the *objective* from the *actual*. By the latter, Frege means something like “that which is capable of acting on the senses.”<sup>22</sup> As Frege says:

The axis of the earth is objective, so is the center of mass of the solar system, but I should not call them actual in the way the earth itself is so...If to be recognized were to be created, then we should be able to say nothing positive about the equator for any period earlier than the date of its alleged creation...What is objective in [space] is what is subject to laws, what can be conceived and judged, what is expressible in words. What is purely intuitable is not communicable. (Frege 1884, pg. 35)

Frege then argues that *numbers*, like the center of mass of the solar system, are *objective* but non-actual; they are neither physical nor mental. (ibid pg. 37-38)

With such preliminaries out of the way, let’s consider Frege’s account of number in detail, since it reveals quite a bit about Frege’s conception of content circa 1884. Now, Frege is trying to learn about the concept of number by analyzing the content of numerical expressions. Frege is trying to determine the content of numerical expressions by reflecting on the content of complete, declarative sentences containing those

---

<sup>21</sup>I remind the reader of Frege’s top-down approach to the determination of content in the *Begriffsschrift*, as we discussed in footnote 12 above.

<sup>22</sup>Cf. (Frege 1897, PW pg. 138).

expressions, in line with Principle (2) above. Now, what is the difference between the content of the sentence ‘Bob has two apples’ and the content of the sentence ‘Bob has one pair of apples’? While the two contain distinct numerical expressions, they certainly don’t seem to be making distinct claims about the physical world, but rather merely reflect distinct ways of *conceptualizing* it. As Frege puts it (referring to different examples):

Now what changes here from one judgement to the other is neither any individual object, nor the whole, the agglomeration of them, but rather my terminology. But that is itself only a sign that one concept has been substituted for another. This suggests as the answer to the first of the questions left open in our last paragraph, that the content of a statement of number is an assertion about a concept. (ibid pg. 59)

For example, if I say that Venus has 0 moons, I have asserted that the concept “being a moon of Venus” has the property of having nothing fall under it. Note that this is still objective and factual. Once the concept  $F$  is held fixed (and provided it doesn’t have vague extension), there is only *one* correct answer to the question, ‘How many  $F$ s are there?’ Frege argues that we can now explain why many have thought that numbers are arrived at via *abstraction* (a view which Frege criticized earlier): many *concepts* are acquired via abstraction, and a number is then discovered in the concept. According to Frege, some other concepts are acquired by combining simple concepts into a more complex concept (e.g. *positive whole number*) via what are roughly the standard Boolean operations of conjunction, disjunction, and negation. Frege also speaks of a *third* way of acquiring new concepts, suggested earlier, and we’ll turn to this topic in a moment.

Frege next turns to the task of *defining* the numbers 0 and 1. He first gives an analysis of the content of sentences of the form ‘The number  $n$  belongs to the concept  $F$ ’: ‘The number 0 belongs to the concept  $F$ ’ is true if, for *any* object  $a$ ,  $a$  does not fall under the concept  $F$ . ‘The number 1 belongs to the concept  $F$ ’ is true if 0 does *not* belong to the concept  $F$  and whenever both  $a$  and  $b$  fall under the concept  $F$ , it follows universally that  $a$  is identical to  $b$ . More generally, ‘The number  $n + 1$  belongs to the concept  $F$ ’ is true if there is an object  $a$  which falls under  $F$  such that the number  $n$  belongs to the concept “ $F$ , but not  $a$ ”. (ibid pg. 67)

Frege is unsatisfied with this, however. He famously points out that we could never “decide by means of our definitions whether any concept has the number *Julius Caesar* belonging to it, or whether that same familiar conqueror of Gaul is a number or not. Moreover we cannot by the aid of our suggested definitions prove that if the number  $a$  belongs to the concept  $F$  and the number  $b$  belongs to the same concept, then necessarily  $a = b$ .” (ibid pg. 68) What Frege is getting at here is that all we’ve done is given the sense of the phrases ‘the number 0 belongs to’, ‘the number 1 belongs to’, etc.; we aren’t yet able to show that a number belonging to a concept is *unique*, nor that nothing *else* is a number belonging to a concept besides what falls under our recursive definition. In contemporary parlance, we haven’t yet given *identity-conditions* for the natural numbers.

Recalling Principle (2), Frege states:

Since it is only in the context of a [sentence] that words have any meaning, our problem becomes this: To define the sense of a [sentence] in which a number word occurs...If we are to use the symbol  $a$  to signify an object, we must have a criterion for deciding in all cases whether  $b$  is

the same as  $a$ , even if it is not always in our power to apply this criterion...When we have thus acquired a means of arriving at a determinate number and of recognizing it again as the same, we can assign it a number word as its proper name. (ibid pg. 73)

Given all this, Frege intends to define the numbers by means of the concept of identity, which he takes as both purely logical and already understood (via Leibniz's definition of identity: "Things are the same as each other, of which one can be substituted for the other without loss of truth", quoted by Frege on pg. 76). The way that Frege intends to go about this has been thoroughly discussed in recent philosophy (e.g. (Hale and Wright, 2001)). Let's consider it in detail.

Frege begins with a different example. Consider the judgment 'Line  $a$  is parallel to line  $b$ ' (in symbols: ' $a//b$ '). Frege says that this judgment

can be taken as an identity. If we do this, we obtain the concept of direction, and say: "The direction of line  $a$  is identical with the direction of line  $b$ ." Thus we replace the symbol  $//$  by the more generic symbol  $=$ , through removing what is specific in the content of the former and dividing it between  $a$  and  $b$ . We carve up the content in a way different from the original way and this yields us a new concept. (ibid pg. 74-75)

This is the *third* way of acquiring new concepts that Frege discusses (the other two being abstraction and Boolean combination). Recall our discussion above of carving a conceptual content in different ways. Frege also suggests a metaphor to explain the difference between this method of acquiring new concepts and the method I called the Boolean combination of simple concepts to make a more complex concept: the latter is like using lines already given to demarcate an area in a new way; the former is like drawing new boundary lines that were not there in the first place. (ibid pg. 100-101)

Anyway, Frege then tries to give a more formal treatment of the intuitive story just discussed. To begin with, Frege suggests we *stipulate* that the sentence ‘The direction of line  $a$  is identical with the direction of line  $b$ ’ is to mean the same as ‘Line  $a$  is parallel to line  $b$ ’. Since Frege accepts Leibniz’s Law of Identity, he suggests that we check the adequacy of this definition by determining whether, whenever line  $a$  is parallel to line  $b$ , ‘the direction of line  $a$ ’ can always be substituted for ‘the direction of line  $b$ ’ (and vice versa) in a complete, declarative sentence without loss of truth.

Frege doesn’t stop there, however. He says that we still can’t use our semi-formal definition to determine whether England, for example, is equal to the direction of the Earth’s axis. What we’re still missing is the concept of *direction* (the acquisition of which was the point of Frege’s intuitive story quoted above), which would allow us to say that our definition applies only to directions. Frege solves this problem to his satisfaction by defining ‘the direction of line  $a$ ’ as the *extension* of a particular concept: namely, the concept *line parallel to line  $a$* .

Why this concept? Because of obvious facts about its extension such as: its extension is identical to the extension of another concept of the form *line parallel to line  $b$  exactly if lines  $a$  and  $b$  are in fact parallel*. This is due to the nature of the relation *being parallel to*. In contemporary mathematical parlance, it is an *equivalence relation*.<sup>23</sup> He can then show that everything he argued above holds of this object, and hence fix the content expressed by any sentence containing the phrase “the direction

---

<sup>23</sup>In a letter to Russell written nearly twenty years later, Frege makes it clear that he believed that such a method works exactly if the relation under consideration is an equivalence relation. Cf. (Frege, Letter to Russell 7/28/1902, PMC pg. 141). Incidentally, this letter also suggests that Frege still considered such a move to be legitimate in his later work.

of line  $a$ ".

How in particular does this story involve carving a content in a new way? Here is a reconstruction of the reasoning involved: One begins with the content expressed by the sentence 'line  $a$  is parallel to line  $b$ ', carved in the way suggested by that sentence (with lines  $a$  and  $b$  as objects bearing the parallel relation to each other). From there, one can recarve this content by holding one of the relata fixed while allowing the other to vary, as suggested by the sentence ' $a$  has the property of being a line parallel to line  $b$ '. Here, one recognizes that one could have just as easily carved this content in the way suggested by the sentence ' $b$  has the property of being a line parallel to line  $a$ '. This fact, along with facts about equivalence relations, implies that the extension of the concept *line parallel to line  $b$*  is identical to the extension of the concept *line parallel to line  $a$* . Hence, Frege exploited the transition from a two-place relation to a one-place property in which a particular relata is held fixed while the other is allowed to vary. (Frege 1884, pg. 79)

Frege brings these ideas to bear on the concept of number. First of all, the arithmetical analogue of being parallel will be being *equinumerous*: two concepts are *equinumerous* exactly if the objects which fall under them can be placed in one-to-one correspondence (this is *Hume's Principle*). Think of trying to determine whether you have just as many apples as you do oranges if you have no idea how to count: you can successfully do so by lining up each apple next to a unique orange and making sure that there are just enough oranges for this to be possible and no more. In doing so, you have defined what mathematicians call a *bijective function* between the set of apples and the set of oranges. As Frege would say, you've shown that the concepts

*being one of my apples* and *being one of my oranges* are equinumerous.

Just as in the direction case, Frege defines ‘the number that belongs to the concept  $F$ ’ as the extension of a particular concept: namely, the concept *equinumerous to the concept  $F$* . He later defines the individual numbers as the numbers belonging to particular concepts (0, for example, is the number which belongs to the concept *not identical with itself*, which works as a definition because everything is self-identical). Frege then proceeds to prove that these definitions both satisfy the desiderata he listed previously and can be used to derive all of the desired arithmetical properties of number (he even extends his definitions into the transfinite). The formal details are unimportant for our present purposes.

It *is* important to note, however, that Frege considered his *particular* choice of definition of number to be inessential: one could have defined the numbers in another way, maybe even as something other than extensions of concepts (but for the logicist reduction to work, it would have to be a “logical object”).<sup>24</sup> This is a mathematician’s definition, by which I mean that Frege *isn’t* arguing that numbers were these particular extensions of concepts *all along*, but rather that we can formally *define* numbers in this way and thereby capture all of the essential properties that

---

<sup>24</sup>On page 117, Frege says, “This way of getting over the difficulty cannot be expected to meet with universal approval, and many will prefer other methods of removing the doubt in question. I attach no decisive importance even to bringing in the extensions of concepts at all.” Note also Frege’s remark that it seems odd to say that a Number is more inclusive than some other extension of a concept (like the extension of *is male* is more inclusive than the extension of *is a bachelor*), but “neither is there anything to prevent us speaking in this way, if such a case should ever occur.” (ibid pg. 81) That is to say, this is an inessential consequence of our chosen modeling, nothing more. It’s just like the number two being *an element of* the number three on the standard set-theoretic definition of the natural numbers in ZFC. Robert May has suggested to me that Frege changed his mind about this in his later work, eventually finding extensions of concepts to be essential to his logical project.



we want numbers to have (for mathematical purposes, at least). This allows us to fix the content expressed by *every* complete, declarative sentence containing numerical expressions. Furthermore, the particular definitions we have chosen (extensions of concepts) are *logical* objects on Frege's view. Hence, it seems that once the details have been worked out, Frege's goal of reducing arithmetic to logic will have been achieved.

Compare a popular set-theoretic definition of an ordered pair:  $(a, b)$  is defined as the set  $\{\{a\}, \{a, b\}\}$ . Now, few would accept that this set-theoretic definition tells us what ordered pairs were all along. Instead, it is sufficient as a definition simply because this set has the main property we want our ordered pairs to possess; namely,  $(a, b) = (c, d)$  exactly if  $a = c$  and  $b = d$ , and hence  $(a, b) \neq (b, a)$  unless  $a = b$ . Other properties of whatever ordered pairs *really* are, whatever those properties may be, seem to be irrelevant for mathematical purposes, provided our formal definition shares the appropriate structure with the notion we're attempting to model. Provided an appropriate structural isomorphism exists, we can let the set  $\{\{a\}, \{a, b\}\}$  act as a "stand-in" for the ordered pair  $(a, b)$ .<sup>25</sup>

There is much more that could be said about the adequacy of Frege's account of number, but we've already dwelt on this topic long enough. A final remark on the *Grundlagen*: Frege calls the laws of number *laws* of laws of nature, since "They assert not connexions between phenomena, but connexions between judgements; and among judgements are included the laws of nature." (ibid pg. 99) Numbers are

---

<sup>25</sup>Haim Gaifman has pointed out to me that this point is also made by W.V. Quine in his (1960, pg. 257-62).

neither physical nor psychological, but *logical*.

With his philosophical motivation complete and his argument sketched, Frege still needs to fill in all the formal details (this would become the *Grundgesetze der Arithmetik* (*Basic Laws of Arithmetic*) Volumes I (1893) and II (1903a)). As is well-known, this led to contradiction.<sup>26</sup> Before that, however, Frege turned to philosophy and wrote a series of highly influential essays on more general issues in the philosophy of language. We will now turn to the first of these essays.

## 1.4 Function and Object, Sense and Reference, Thought and Truth-Value

The centrality of the mathematical concept of *function* in Frege's thought led him to consider functions in mathematics more generally and develop a particular position on the nature of functions. This, coupled with worries about the content of identity-statements and worries about substitution failures in indirect discourse, arguably led him to his famous distinction between sense and reference. Let's consider Frege's first published discussion of this distinction: his "Function and Concept" of 1891.

Frege begins "Function and Concept" by considering a standard answer to the

---

<sup>26</sup>I will not extensively discuss the contradiction in Frege's formal system in this dissertation. Frege himself (and most scholars) trace the contradiction to Frege's Basic Law V, which (when translated into a more contemporary style) states that for any first-level function  $f$  of one argument and any first-level function  $g$  of one argument, the course-of-values (or value-range) of  $f$  is identical to the course-of-values of  $g$  if and only if for any object  $a$ ,  $f(a) = g(a)$ .

Strictly speaking, one can retain Basic Law V while making other adjustments to Frege's formal system in order to avoid the contradiction, but it is highly doubtful that any formal adjustments will recover Frege's attempt to reduce arithmetic to logic (see (Hale and Wright 2001) for a well-known attempt). Frege himself gave this project up.

question ‘What is a function?’: namely, a *function* of  $x$  is a mathematical expression containing  $x$ . Hence, ‘ $2(x^3) + x$ ’ is a function of  $x$  and ‘ $2(2^3) + 2$ ’ is a function of 2. Frege responds, “This answer cannot satisfy us, for here no distinction is made between form and content, sign and thing signified...a mere expression, the form for a content, cannot be the heart of the matter; only the content itself can be that.” (Frege 1891, pg. 2-3/138)<sup>27</sup>

Frege continues his criticism of the opposing formalist view: “Difference of sign cannot by itself be a sufficient ground for difference of the thing signified.” (ibid pg. 3/138) According to Frege, this confusion is what leads to silly claims such as  $2 + 5$  and  $3 + 4$  are “equal but not the same”. Frege motivates these points in several ways. First of all, he points out that one isn’t going to discover the property of being the multiplicative identity by investigating the *numeral* ‘1’, just like one isn’t going to discover whether London is pretty by staring at the *word* ‘London’. Marks on paper have chemical and physical properties that can be investigated, but one isn’t thereby investigating *numbers*. Introducing new numerals doesn’t give us new numbers with new properties to be investigated any more than calling my laptop ‘Richard’ gets me a new object with new, interesting properties to discover.

A natural thing to say is that ‘ $2 + 5$ ’, ‘ $3 + 4$ ’, and the numeral ‘7’ are different ways of picking out the same object; they are different expressions with the same *reference*. Frege points out that the conception of different expressions having the same reference has been underlying our talk of numbers *all along*. One says the

---

<sup>27</sup>Note that on my interpretation, Frege is criticizing his own view from the *Begriffsschrift*, although this controversial interpretive claim won’t be important in what follows. Also, Frege obviously never made the mistake of confusing *numbers* with expressions.

equation ' $x^2 = 4$ ' has two roots, but it would have indefinitely many if, e.g. 2, 1 + 1, and 6/3 counted as distinct. Based upon these considerations, Frege concludes that expressions such as ' $2(2^3) + 2$ ' refer to *numbers* (in this case, the number 18), not functions. The expression ' $2(x^3) + x$ ', on the other hand, *indefinitely* indicates a number: it both *isn't* a function and doesn't *refer to* a function. The *function*, says Frege, is "what is present in ' $2.x^3 + x$ ' over and above the letter ' $x$ .'" (ibid pg. 6/140) We recognize the same function in the expressions ' $2(1^3) + 1$ ', ' $2(4^3) + 4$ ', ' $2(5^3) + 5$ ', etc. The essential feature of the function here is revealed by what these expressions have in common; each expression can be understood as referring to the result of applying the same function to different arguments. Frege uses ' $2.( )^3 + ( )$ ' in order to refer to this function (although he isn't completely satisfied with this expression either).

Frege then states, "I am concerned to show that the argument does not belong with a function, but goes together with the function to make up a complete whole; for a function by itself must be called incomplete, in need of supplementation, or 'unsaturated'. And in this respect functions differ fundamentally from numbers." (ibid pg. 6/140) These suggestive remarks form the basis of Frege's fundamental ontological distinction between *objects* and *functions* (including *concepts*, as we shall see), as alluded to in Principle (3) of the *Grundlagen*. Functions are essentially *incomplete* in a way that objects are not.

Later in this essay, Frege briefly discusses what he means by 'Object': "I regard a regular definition as impossible, since we have here something too simple to admit of logical analysis. It is only possible to indicate what is meant. Here I can only say

briefly: An object is anything that is not a function, so that an expression for it does not contain an empty place.” (ibid pg. 18/147)<sup>28</sup> Let’s consider another of Frege’s suggestive remarks, mentioned in footnote 18 above: “We recognize the function in the expression by imagining the latter as split up, and the possibility of thus splitting it up is suggested by its structure.” (ibid pg. 7/141)

Compare carving up the content of a complete, declarative sentence into objects and functions/concepts. The contents expressed by ‘London is pretty’, ‘Paris is pretty’, ‘Venice is pretty’, etc., can each be grasped as asserting that a different object falls under the same concept, namely, *being pretty*. One might think that what Frege is doing here is carving up that which is expressed by an expression (including expressions that are not sentences), what he’ll later call the *sense* of the expression, into argument-senses and function-senses. Each such carving will thus have particular objects and functions as referents corresponding to these argument-senses and function-senses. This is the natural extension of his *Begriffsschrift*-era views on carving expressions into arguments and functions, and we’ll see below that Frege was committed to it.

The *value* of a function on a particular argument is the referent of the expression formed by replacing each of the empty spaces in an expression which refers to the function by a *name of* (i.e. *expression referring to*) that argument. Hence, in the example above, ‘ $2(1^3) + 1$ ’ can be thought of as referring to the result of applying the following function to the number 1: cube the argument, multiply the result by 2, and add this result to the original argument. Note that we could *also* carve up the sense

---

<sup>28</sup> *Pace* Frege, I think we can and should say more. This will be a topic of Chapter 4.

of the expression ‘ $2(1^3) + 1$ ’ into the sense of the expression ‘2’ and the sense of the remainder, where the remainder refers to the function that takes an argument (the number 2 in this case), multiplies it by  $1^3$ , and adds 1 to the result. Frege doesn’t address this point in this paper, but it seems to fit nicely with his earlier views on conceptual content.

In fact, Frege provides us with an excellent geometric metaphor to elucidate his conception of the object/function distinction, a metaphor which immediately suggests the possibility of multiple decompositions of content:

We may compare this [the splitting up of a mathematical expression into argument-expression and function-expression—NB] with the division of a line by a point. One is inclined in that case to count the dividing-point along with both segments; but if we want to make a clean division, i.e. so as not to count anything twice over or leave anything out, then we may only count the dividing-point along with one segment. This segment thus becomes fully complete in itself, and may be compared to the argument; whereas the other is lacking in something – viz. the dividing-point, which one may call its endpoint, does not belong to it. Only by completing it with this endpoint, or with a line that has two endpoints, do we get from it something entire. (ibid pg. 7/141)

Recall that Frege began his career as a geometer, and geometry provides many obvious examples of multiple decompositions: e.g. carving a square into two triangles or two rectangles.

Frege’s notion of function is very general. For instance, he speaks of the function  $x^2 = 1$ .<sup>29</sup> But what is the *value* of this function on a particular argument? Equivalently, what does the corresponding completed expression *refer to*? In this

---

<sup>29</sup>Frege allows himself to use such expressions for functions provided that it is kept in mind that the letter ‘ $x$ ’ merely serves to indicate where a name of an argument would go in order to construct an expression which refers to the *value* of the function on that particular argument.

case, replacing the ‘ $x$ ’ in this expression by a *numeral*, for example, will result in a complete, declarative sentence. Frege points out that the resulting sentence will be *true* if ‘ $x$ ’ is replaced by ‘1’ or ‘ $-1$ ’ and *false* if ‘ $x$ ’ is replaced by ‘0’ or ‘2’. The point is, as we vary the argument while holding the function fixed, the truth-value of the corresponding sentences also varies. Frege concludes, “I now say: ‘the value of our function is a truth-value’, and distinguish between the truth-values of what is true and what is false. I call the first, for short, the True; and the second, the False.” (ibid pg. 13/144) This reveals part of Frege’s motivation for regarding sentences as referring to *truth-values*. There is more to be said here, but since Frege’s treatment of truth-values will be a major topic of Chapter 4, let’s move on to Frege’s stated motivation for the sense/reference distinction.

I’ll reproduce Frege’s first published explanation of the sense/reference (Sinn/Bedeutung) distinction here in full:

The objection here suggests itself that ‘ $2^2 = 4$ ’ and ‘ $2 > 1$ ’ nevertheless tell us quite different things, express quite different thoughts; but likewise ‘ $2^4 = 4^2$ ’ and ‘ $4.4 = 4^2$ ’ express different thoughts; and yet we can replace ‘ $2^4$ ’ by ‘ $4.4$ ’, since both signs have the same [reference]. Consequently, ‘ $2^4 = 4^2$ ’ and ‘ $4.4 = 4^2$ ’ likewise have the same [reference]. We see from this that from sameness of reference 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 one as true and the other as false. And yet both sentences must [refer to] the same thing; for it is just a matter of interchange of the words ‘Evening Star’ and ‘Morning Star’, which [refer to] the same thing, i.e. are proper names of the same heavenly body. We must distinguish between sense and [reference]. ‘ $2^4$ ’ 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, ‘ $2^4 = 4^2$ ’ and ‘ $4.4 = 4^2$ ’ [refer to] the same thing, but have not the same sense (i.e.,

in this case: they do not contain the same thought). (ibid pg. 13-14/144-145)

This paragraph reveals a great deal about Frege's underlying motivation. First of all, Frege is implicitly assuming a principle of the *Compositionality of Reference*: the reference of an expression is determined by (is a *function* of) the referents of *parts* of the expression. Hence, if one replaces a part of an expression by a distinct expression with the *same* reference as that part, the reference of the larger expression remains the same. Hence, since many such replacements result in expressions which intuitively express quite different contents although reference is preserved, we must distinguish such content, which Frege calls *sense*, from reference. Again, Frege calls the sense expressed by a complete, declarative sentence a *thought*. He thus replaces his old notion of judgeable content by *thought* and *truth-value*.<sup>30</sup>

The last sentence also suggests that Frege is implicitly assuming a principle of the *Compositionality of Sense*: the sense of an expression is determined by (is a function of) the senses of parts of the expression. Hence, if one replaces a part of an expression by a distinct expression with the same sense as that part, the sense of the resulting expression remains the same.

---

<sup>30</sup>Frege makes this point explicitly in his Letter to Husserl of 5/24/1891. (PMC pg. 63) Recall our discussion above of Frege's argument/function distinction in the *Begriffsschrift*: there, a function was a meaningful *expression* that mapped other *expressions* (arguments) to the conceptual content represented/expressed by the complete expression (of which the function and arguments were parts). Now, one might see Frege as specifying *two* sorts functions which determine the content, one at the level of *sense* and one at the level of *reference*: namely, (1) a function-sense that maps argument-senses to the complete sense of the corresponding complete expression (including the case of a complete thought), (2) the function corresponding to that function-sense (as its referent) that maps objects (corresponding to the referents of the argument-senses) to objects (corresponding to the referent of the complete expression, which will be a truth-value in the case of complete, declarative sentences). I will argue that Frege was committed to this view below.



A final remark on this paragraph: Frege suggests a way of determining that two sentences express *distinct* thoughts: consider whether it is possible for a (linguistically competent) agent to regard one as true and the other as false. If so, the thoughts expressed are distinct. This criterion isn't specific enough to be especially useful, but we'll consider it further in Chapter 3.

Frege defines a *concept* as a function whose value is always a truth-value. He motivates this by considering again the function  $x^2 = 1$ : it maps arguments to the True and the False. So, for instance, the value of this function on argument  $-1$  is the True. Frege says we can express this as ' $-1$  is a square root of 1' or ' $-1$  falls under the concept: square root of 1'. Similarly, we can express the fact that this function takes 2 to the False by saying ' $2$  does not fall under the concept: square root of 1'. For Frege, a *function* seems to be (roughly speaking) anything with which one could (in principle) associate a unique, definite object (the "output") given any object whatsoever (as "input"). So, there seems to be an intimate connection between functions from objects to truth-values and *concepts*. Hence, Frege identifies them. A *concept* is just a function whose range consists of the two truth-values: the True and the False. So, the *extension* of a concept is subsumed under Frege's more general notion of the *course-of-values* or *value-range* of a function.<sup>31</sup>

---

<sup>31</sup>Note that Frege does *not* identify a function with its *graph*, as in modern Set Theory. Note that a course-of-values is not the same thing as a set of ordered pairs (the *graph* of a function). In particular, since Frege requires that every function be defined over absolutely everything, *including its own course-of-values*, every course-of-values will be radically non-well-founded. Frege was *not* working with a background set theory here, at least as understood in accordance with the iterative conception of set. He was against the idea of defining sets/classes/extensions via "aggregating" their elements, and instead took concepts to be the prior notion. For example, "But it is surely a highly arbitrary procedure to form concepts merely by assembling individuals, and one devoid of significance for actual thinking unless the objects are held together by having characteristics in common. It is precisely these which constitute the essence of the concept." (Frege 1880-81, PW pg.

With respect to his formal system, Frege insists that the *domain* of a function be absolutely everything whatsoever. Frege insists on this so that an expression such as ‘ $a + b$ ’ has a reference no matter what objects ‘ $a$ ’ and ‘ $b$ ’ refer to, even if in some cases one picks the reference of the resulting expression by an arbitrary rule. For example, we can stipulate that ‘ $2 + \text{the Moon}$ ’ refers to the False, and hence the addition function can still be defined when we take the Moon as one of its arguments. Similarly, Frege requires that within a formal system concepts have sharp boundaries: for any argument whatsoever, a concept should have a truth-value as its value. Otherwise, in certain cases the concept would take no value at all, and in particular the corresponding sentences wouldn’t refer to truth-values. For example, if the function  $x + 1$  had no value on some argument, the concept  $x + 1 = 10$  wouldn’t take a truth-value as value on that argument either.

Why does Frege place such draconian requirements upon functions and concepts within his formal system? Frege is well aware that many expressions in natural language fail to refer, and many (perhaps *most*) concepts have vague extensions. Recall that Frege’s original motivation for studying language was to develop a precise, formal language for expressing conceptual content in a way that elucidates the structure of logical inference, particularly in arithmetical reasoning. As soon as sentences in his formal language can lack truth-value, Frege’s logical laws no longer apply universally.

---

34) See also Frege’s discussion in section 150 of Volume II of the *Grundgesetze*.

Since we aren’t considering Frege’s formal system or the nature of its inconsistency in detail, these distinctions aren’t really important for our purposes. For Frege, the course-of-values is an *object*, while the function is not an object. Frege calls this notion of function the logically prior one. One wouldn’t have come up with, e.g., the static, set-theoretic definition without some intuitive notion of a *rule* or *procedure* that produces a unique output for every input. Compare our discussion of ordered pairs above.

More importantly, such sentences fail to be *about the world* at all, on Frege's view. For instance, if in a simple sentence of the form ' $F(a)$ ', ' $a$ ' lacks a referent, then what is the concept referred to by the expression ' $F(\quad)$ ' predicated *of*? What is it saturated *by*? At best, such sentences may express thoughts in the realm of fiction, such as the thought expressed by 'Odysseus was set ashore at Ithaca while sound asleep.' When we utter such sentences, we aren't interested in truth-value, but rather merely the thought being expressed.<sup>32</sup> Hence, Frege insists that all the expressions in his formal language (and any "scientific" language) have reference, all functions are defined for *any* argument, and all concepts have precise extensions.

Thus we have seen how Frege's conception of content naturally evolved from the *Begriffsschrift* to his eventual distinction between sense and reference. Let's now discuss the structure of senses in more detail. We will do so in the context of a puzzle for Frege's conception of content that has been widely discussed in the literature.

## 1.5 A Puzzle

Recall that Gottlob Frege considered the *thought* expressed by a complete declarative sentence to be the sort of content relevant to determining whether that sentence is true or false. Here is one of many examples in which he made this claim: "So one has to separate off from the content of a sentence the part that alone can be accepted as true or rejected as false. I call this part the thought expressed by the sentence."

(1906c, PW pg. 197-8)

---

<sup>32</sup>Frege makes these points in "On Sense and Reference" (1892a), which we will address directly in the following chapter.

Furthermore, when discussing the structure of thoughts, Frege helped himself to mereological language, often speaking of parts of a thought corresponding to the significant parts of a sentence expressing it.<sup>33</sup> For example, when I utter the sentence ‘John has brown hair’, my utterance expresses the thought that John has brown hair, a thought which contains the sense of ‘John’ and the sense of ‘has brown hair’ as parts.

Since a thought is a more abstract entity than the full content of a complete declarative sentence (a content which may include elements irrelevant to the truth or falsity of the sentence), it is no surprise that Frege wanted distinct sentences to be capable of expressing the same thought. For example, on Frege’s view the sentences ‘John loves Mary’ and ‘Mary is loved by John’ express the same thought; the transition from active to passive voice does not affect the thought expressed. Discussing a similar example, Frege says “Although in actual speech it can certainly be very important where the attention is directed and where the stress falls, it is of no concern to logic.” (1897, PW pg. 141) Intuitively, the sentence ‘John loves Mary’ carries no more and no less information than the sentence ‘Mary is loved by John’; any difference between the two is irrelevant to determining their (shared) truth-value.

Many Frege scholars have found this aspect of Frege’s view to be in tension with his view that thoughts have parts corresponding to the significant parts of sentences

---

<sup>33</sup>Here is the example from the Introduction once again: “How can [language] achieve so much? By virtue of the fact that thoughts have parts out of which they are built up. And these parts, these building blocks, correspond to groups of sounds, out of which the sentence expressing the thought is built up, so that the construction of the sentence out of parts of a sentence corresponds to the construction of a thought out of parts of a thought. And as we take a thought to be the sense of a sentence, so we may call a part of a thought the sense of that part of the sentence which corresponds to it.” (1914, PW pg. 225)

expressing them. Returning to our simple example, how could ‘John loves Mary’ and ‘Mary is loved by John’ express the same thought, given that the thought expressed by the former sentence contains a part corresponding to ‘loves’ while the latter does not? Surely the thought expressed by the latter sentence instead contains a part corresponding to ‘is loved by’. These parts cannot be the same, since the former denotes the “loves” relation while the latter denotes its obviously distinct inverse: the “is loved by” relation.<sup>34</sup>

There are plenty of other seemingly-problematic examples in Frege’s writings, some of which involve multiple analyses of the same sentence. Here is one: “Now if the same proper name occurs in both consequent and antecedent, we may regard the hypothetical thought as singular if we think of it as being analysed into the complete part that corresponds to the proper name and the unsaturated part left over.” (Frege 1906b, PW pg. 188) Thus, the thought expressed by the sentence ‘If John loves Mary, then John is happy’ can not only be analyzed into the two thoughts expressed by ‘John loves Mary’ and ‘John is happy’ connected by the sense corresponding to the material conditional; it can also be analyzed into the sense expressed by ‘John’ and the predicative sense corresponding to the remainder of the sentence (with both instances of ‘John’ removed: ‘If ( ) loves Mary, then ( ) is happy’). Yet these analyses give distinct logical forms: the former corresponds to a material conditional between atomic sentences (e.g. ‘ $L(j, m) \longrightarrow H(j)$ ’) while the latter corresponds to an atomic sentence (e.g. ‘ $K(j)$ ’).

---

<sup>34</sup>I trust the reader familiar with Frege’s “concept horse” problem can rewrite this sentence in their preferred style, if they so choose.

It follows that on Frege's view, the structure of a thought does not determine a unique logical form. However, when one constructs a formal language for *expressing* thoughts (in order to give a precise account of various inferential transitions between thoughts, for example), one may choose to construct one's language in such a way that it satisfies *unique readability*: for every well-formed formula of the language, there is exactly one procedure for generating that formula from its component parts (e.g.  $((p \wedge q) \wedge r)$  is uniquely generated by applying the operator  $\wedge$  to atomic sentences  $p$  and  $q$ , followed by again applying  $\wedge$  to  $(p \wedge q)$  and  $r$ ). This gives a unique decomposition of any well-formed formula into its components. The point being, on Frege's conception of thought, the same thought can be expressed by logically distinct sentences and in distinct formal languages, even if each choice of expression has a unique decomposition.

Interestingly, Frege's own formal language, as developed in the *Grundgesetze* (1893), does not satisfy unique readability. For example, the formula  $\neg f(a)$  can be read as the application of the horizontal to the result of applying the function-name  $f(\xi)$  to the object-name  $a$ , but it can also be read as the application of the horizontal to the result of applying the two-place function-name  $\phi(\xi)$  to the function-name  $f(\xi)$  and the object-name  $a$ , where this two-place function-name refers to the relation between an object and a concept under which it falls. (ibid pg. 39)

Frege was quite explicit that the same thought can be analyzed in distinct ways. For example, in the essay quoted above he says, "If several proper names occur in a sentence, the corresponding thought can be analysed into a complete and unsaturated

part in different ways. The sense of each of these proper names can be set up as the complete part over against the rest of the thought as the unsaturated part.” (1906b, PW pg. 192)<sup>35</sup>

But how could a thought have multiple analyses, given that a thought is built up out of its constituent senses as parts?<sup>36</sup> Won’t the structure of a thought be uniquely determined by its constituent senses and the way they are combined? Wouldn’t any distinct analysis of a thought force either distinct senses as constituents or a distinct manner of combination? While it is true that, as I pointed out in footnote 2, in one of his last papers Frege explicitly called his mereological treatment of thoughts “figurative”, this doesn’t give us a satisfactory answer to these questions.<sup>37</sup> Frege treated thoughts as composed of senses, and hence similar questions arise however one understands such composition.

In the rest of this chapter, I’ll undertake a careful reading of Frege’s remarks on the structure of thought and address the controversies that have appeared in recent scholarship. I will first discuss the sort of structure that thoughts should have on Frege’s conception, and explain why (*pace* Dummett, Bell, Penco, and others) this treatment of thoughts doesn’t conflict with his other views. I will then argue that a proper appreciation of Frege’s understanding of the part/whole relation will allow us

---

<sup>35</sup>There are many additional examples. Prior to making his sense/reference distinction, Frege spoke of multiple analyses of judgeable content in (Letter to Marty 8/29/1882, PMC pg. 101) and (1884, pg. 74-5), among others. After making this distinction, Frege spoke of multiple analyses in (1892b, PW pg. 107-8), (1906b, PW pg. 187), and (1906c, PW pg. 201-2), among others. Once again, see the Appendix for a comprehensive list of every Frege quotation that I consider to be relevant to the Multiple Decompositions Thesis.

<sup>36</sup>See footnote 33.

<sup>37</sup>Indeed, as I pointed out in footnote 2, it isn’t clear that this retreat from mereology is really compatible with Frege’s earlier views.

to reconcile his remarks on multiple analyses of the same thought with his remarks on a thought being composed of senses as parts that correspond to the significant parts of a sentence expressing it.

## 1.6 The Structure of Thought

In Chapter 15 of his *The Interpretation of Frege's Philosophy* (1981), Michael Dummett argues that Frege held the following four theses:

- A1.** A thought may be analyzed in distinct ways.
- A2.** A thought is not built up out of its component [senses];<sup>38</sup> rather, the constituents of the thought are arrived at by analysis of it.
- B1.** The senses of the parts of a sentence are parts of the thought expressed by the whole.
- B2.** A thought is built up out of its constituents, which correspond, by and large, to the parts of the sentence expressing it. (ibid pg. 261)

As Dummett points out, the A theses seem to contradict the B theses. How can a thought be analyzed in distinct ways into distinct constituents if a thought is built

---

<sup>38</sup>Dummett actually writes 'concepts' rather than 'senses' here, claiming to follow Frege's *early* usage of 'concept' ('Begriff'). Dummett claims that early Frege used 'concept' in roughly the sense in which he would later use 'sense' ('Sinn'). (ibid pg. 261-2) As is well-known, after making his sense/reference distinction, Frege did not use the word 'concept' to denote a component of a thought, but rather the referent of a predicate. *Pace* Dummett, I do not think early Frege used 'concept' to in effect mean thought-constituent, since a judgeable content is partly a union of thought and truth-value (sense and reference), and have replaced 'concepts' by 'senses' in order to avoid confusion.



up out of its constituents? How could a thought both be and not be built up out of its component senses? Could Frege have failed to notice this?

According to Dummett, the A theses and B theses correspond to two different relations of a whole thought to its parts: the A theses assume a relation much like that of a country that can be subdivided into regions in multiple ways for various purposes, with no unique best subdivision, while the B theses assume a relation much like that of a molecule to its atoms, with a unique analysis of that molecule into its constituent atoms. (ibid pg. 263-4)

Dummett mentions a possible solution: Peter Geach has defended the view that for Frege the structure of a thought is best analyzed using Frege's own argument/function distinction (Geach 1975, pg. 149). On such an interpretation, just as e.g. the number 16 is the value of the function  $f(x) = x^2$  on the argument 4 and the value of the function  $g(x) = 4^x$  on the argument 2, the same thought can be the value of multiple functions on multiple arguments. So, the thought that John loves Mary could be seen as the value of the function-sense expressed by 'loves Mary' with the object-sense expressed by 'John' as argument. But it could also be seen as the value of the function-sense expressed by 'John loves' with the object-sense expressed by 'Mary' as argument.

For Geach, this explains why Frege often spoke of one part of a thought needing to be *incomplete* or *unsaturated* in order for the parts of that thought to hold together.<sup>39</sup>

Just as the referent of an incomplete expression is itself incomplete or unsaturated for

---

<sup>39</sup>Frege says this in multiple places. For example: "For not all the parts of a thought can be complete; at least one must be 'unsaturated', or predicative; otherwise they would not hold together." (1892c, pg. 205/193)

Frege, the *sense* of an incomplete expression is incomplete as well. Given that Frege treats the referent of an incomplete expression as a function that must be saturated by an argument (or arguments) in order to result in a value, it is unsurprising that he treats the sense of an incomplete expression as incomplete in the same way: as a function that must be completed by an argument (or arguments) to result in a complete sense as value.<sup>40</sup>

Dummett rejects this analysis of the structure of thought. Recall thesis B1: the senses of the parts of a sentence are parts of the thought expressed by the whole. Hence, given Geach's interpretation, a function-sense will be a part of its value along with its argument. For example, given that the sense of 'loves Mary' is a function from the sense of 'John' to the thought that John loves Mary, this function-sense, being the sense of 'loves Mary', is also a part of the thought that John loves Mary. This is quite different from the case of reference. Consider again the number 16. Few would find it plausible that the function  $f(x) = x^2$  is a *part of* 16 in addition to mapping the number 4 to it. Functions over the natural numbers are not *parts* of those numbers.

Dummett notes this difference while primarily criticizing the interpretation for leaving us unable to say how the sense of a predicate contributes to the condition under which a sentence containing that predicate is true. (1981, pg. 251-2) Recall that Frege often introduced thoughts as the sort of content of a complete declarative sentence relevant to determining whether that sentence is true or false. Hence, a

---

<sup>40</sup>Recall that I also suggested this interpretation above, which I developed independently. I will argue for it below.

thought fixes a truth-condition.<sup>41</sup> According to Dummett, if the sense of a predicate is a function from the sense of a name to a thought, then we can provide no account of how it contributes to fixing that truth-condition. The way the sense of a predicate contributes to fixing a truth-condition must be via determining a condition that objects either satisfy or fail to satisfy. This sense *just is* a way of determining such a condition, where that condition is a function from objects to truth-values. But if this sense is itself a function from object-senses to thoughts, then it isn't a way of determining such a condition. (1981, pg. 270)

Dummett also complains that in order to understand such a function, we would already need to know what thoughts are and how to identify particular thoughts, so this account is of no help in explaining what thoughts are or how we grasp them. (ibid pg. 267ff) I see little reason to accept Dummett's insistence that to form a conception of a function, we must already know its range of possible values. Consider a function corresponding to a process of construction. I may have no idea what the output will be until after I have engaged in the process. I discover the output by engaging in the process relative to a particular input, and I conceive of the output as the result of that process. Frege's own geometric example meant to help explain the nature of a function, which we discussed above and will return to below, can be viewed in this way.

---

<sup>41</sup>Consider the following famous passage from Volume I of the *Grundgesetze*: "Every such name of a truth-value *expresses* a sense, a *thought*. For owing to our stipulations, it is determined under which conditions it refers to the True. The sense of this name, the *thought*, is: that these conditions are fulfilled...Now, the simple or complex names of which the name of a truth-value consists contribute to expressing the thought, and this contribution of the individual name is its *sense*. If a name is part of the name of a truth-value, then the sense of the former name is part of the thought expressed by the latter." (1893, pg. 50-1)

Returning to Dummett's primary criticism, there are two separate questions here. The first is whether Geach's position on the senses of predicates was Frege's actual view. The second is whether this view is tenable. Regarding the first question, it is clear that Frege was in fact committed to the senses of predicates being functions from the senses of names to thoughts. Note that on Frege's view, when I assert the sentence 'Copernicus believed that the planetary orbits are circles', the expressions 'the planetary orbits' and 'are circles' refer to their customary senses, and the complete expression 'the planetary orbits are circles' refers to the thought customarily expressed by it.<sup>42</sup> Hence, the referent of this complete expression ought to be a function of the referents of 'the planetary orbits' and 'are circles', and since the former expression is complete while the latter is unsaturated, the latter must refer to a function on Frege's view, in particular, a function mapping the referent of the former to the complete thought.<sup>43</sup> Hence, the customary sense of 'are circles' not only *refers* to a function; it *is* a function from senses to thoughts. In this case, it maps the customary sense of 'the planetary orbits' to the complete thought expressed by 'the planetary orbits are circles', a thought which contains both senses as parts.

In my view, much of the resistance to treating the sense of a predicate as a function from senses to thoughts while at the same time being a part of the thoughts it maps to comes from various prejudices concerning functions that have little to do

---

<sup>42</sup>Cf. (1892a, pg. 28/159). We will discuss this in great detail in the next chapter; here I am merely pointing out a consequence of this view.

<sup>43</sup>In support of this, note that in "Function and Concept", Frege says that an object is "anything that is not a function, so that an expression for it does not contain an empty place." (1891, pg. 18/147) Given that Frege's ontology divides completely into objects and functions, a function is therefore anything whose expression *does* contain an empty place. Compare also his definition of function-names in Volume I of the *Grundgesetze*. (1893, pg. 43-4)

with Frege’s conception of them. As I pointed out above, in Frege’s “Function and Concept” (1891, pg. 7/141), he uses the following analogy to explain the sense in which functions are unsaturated while objects are not: Consider a line divided by a single point. Now treat the line as divided into two segments, one which includes the point and one which doesn’t. The segment which doesn’t include the point is supposed to be analogous to a function while the other segment, being “complete in itself”, is analogous to the object. Referring to the former, Frege says, “Only by completing it with this endpoint, or with a line that has two endpoints, do we get from it something entire.” (loc. cit.) Extending Frege’s analogy, what does this function map the other line-segment *to*? The whole line, of course. And yet at the same time, this function is a part of that line. More generally, it combines its arguments with itself in order to form a longer line-segment.

I think that this is a perfectly acceptable notion of “function”, and find it plausible that this is how Frege was thinking of the unsaturated senses of predicates as both functions from complete senses to thoughts and parts of those thoughts.<sup>44</sup> This is not to say that *all* functions map their arguments to values which include those very functions and arguments as parts, which Frege rejected explicitly in his (1919, PW pg. 255), nor is this meant to somehow fully explain the nature of understanding a predicate. The point is merely that nothing in Frege’s conception of functions implies that a predicative sense can’t both be a part of its value (given a complete sense as

---

<sup>44</sup>In addition to Frege’s geometric example, consider a formal syntax in which predicate symbols are functions from constant symbols and variables to atomic formulas. These functions will thus be components of their values, since their values are atomic formulas containing them. I trust the reader recognizes both this example and the previous geometric example as coherent, no matter how strange this conception of functions may seem. Haim Gaifman has suggested the possibility of such a formal syntax in unpublished work from 2006. I discuss set-theoretic objections below.

argument) and count as a function.<sup>45 46</sup>

Furthermore, in response to Dummett, granting that the sense of a predicate is a *way of thinking* of a condition that objects either satisfy or fail to satisfy (where that condition is a function from objects to truth-values), there is nothing to stop it from *also* being a function from object-senses to thoughts, as indeed it must be in order to remain consistent with Frege's views on the referents of expressions in indirect contexts, the referents of incomplete expressions, and the compositionality of reference (i.e. the referent of a whole expression is a function of the referents of its significant parts). One might even say that the sense of a predicate is *fundamentally* a way of thinking of a condition on objects while adding that it is itself incomplete, much like its referent, and must be saturated by an object-sense to form a complete thought. The logical form of a predicative sense doesn't constrain its content to such an extent that it can't be a way of thinking of a function. Dummett has given us no reason to think otherwise, and hence we have no reason to reject the view Frege was clearly committed to.<sup>47</sup>

---

<sup>45</sup>In fact, as I pointed out above, prior to making his sense/reference distinction Frege seemed to hold this view of conceptual contents carved into arguments and functions. Cf. his (1880-81, PW pg. 16) in combination with Section 9 of his (1879).

<sup>46</sup>The reader might still be bothered by such a conception of a function, since this function's values for appropriate arguments will contain the function itself as a part. This may seem somehow incompatible with the standard axioms of set theory (ZFC), since a set of ordered pairs can't include itself as an element of the second component of an ordered pair without violating the axiom of foundation. Let me make a few brief points in response: (1) There is no reason that one's formal representation of a function as a set of ordered pairs must also represent the mereological structure of the function's values; set theory and mereology are distinct enterprises. (2) Even if one wants one's formal representation of such a function to be such that its values are sets containing that function as an element, one can easily do so in a non-well-founded set theory, including ZFC itself without the axiom of foundation. (3) For Frege, a function is fundamentally an unsaturated entity while a set of ordered pairs is saturated (it is an object). Hence, such a formal representation doesn't respect Frege's conception of functions in any case.

<sup>47</sup>Heck and May (2011) raise a similar complaint to Dummett's: "Saying that thoughts are

## 1.7 Parts and Wholes

The function/argument analysis of thought is still insufficient to solve our original interpretive problem, however. As Dummett points out, even if this interpretive position could be made tenable (as I argued it could in the last section), it would only explain why Frege held the A theses, not the B theses, which seem to presuppose a different conception of the structure of thought that is manifestly incompatible with the A theses. (1981, pg. 266)

In fact, for us the problem is worse, since I have argued above that Frege held (or, more carefully: was committed to) *both* the function/argument analysis of thought *and* a part/whole conception of thoughts, so that the function-sense expressed by ‘loves Mary’ is both a function from the sense of ‘John’ to the thought that John loves Mary and a part of that thought, thus leaving us seemingly unable to use the function/argument analysis to explain why the same thought can be analyzed in different ways. If such analysis can be made sense of, it is not simply on the model of distinct numerical functions mapping distinct arguments to the same value: as I

---

the values of sense-functions in no way explains how the composition of senses determines truth-conditions.” (ibid pg. 147) It isn’t meant to. All that is needed at the level of *form* is a structural distinction between senses denoting objects and senses denoting functions (and senses denoting second-level functions, etc.), which the function/argument analysis obviously provides. This is in no way meant to completely explain the *content* of a thought, and in particular doesn’t provide a complete explanation of the way in which a thought fixes a truth-condition, or of the contribution the parts of a thought make to fixing that truth-condition.

Heck and May also complain that such a conception can’t explain why ‘The Morning Star is a planet’ and ‘The Evening Star is a planet’ express distinct thoughts. (ibid pg. 147) While I won’t address their complaint in detail here, it is important to recognize, as I argued above, that Frege held *both* the function/argument analysis of thought and a part/whole conception of thought: the predicative sense expressed by ‘is a planet’ is both a function from the sense expressed by ‘The Morning Star’ to the thought expressed by ‘The Morning Star is a planet’ and a part of that thought (along with its argument). This is why the sense of ‘is a planet’ can’t be a function of an arbitrary argument to the thought that the Morning Star is a planet; it can only be a function of the sense of ‘The Morning Star’ to that thought, since that thought will contain both this function and its argument as parts.

stated above, the function  $f(x) = x^2$  may map 4 to 16, but  $f$  is not a part of 16.

Dummett himself attempts to solve the interpretive problem by making a distinction between *analysis* and *decomposition*, arguing that thoughts have a unique *analysis* but can be *decomposed in* multiple ways. Dummett argues that although Frege never says so explicitly, he was clearly working with such a distinction, and the apparent contradiction in his remarks can be resolved by interpreting the B-theses in terms of analysis and the A-theses in terms of decomposition. (ibid pg. 271) For Dummett, analysis reveals the ultimate constituents of the thought (and how they are combined), while decomposition reveals a pattern that is shared by multiple thoughts that can be exploited for the purposes of inference.<sup>48</sup>

This attempt at a resolution has been roundly criticized by Frege scholars. Many point out that it seems to have little to no textual support.<sup>49</sup> If Frege really distinguished analysis from decomposition, he never seemed to say so. It would be preferable to find a resolution that corresponds to what Frege actually says.

There is another way of solving this interpretive problem. Note that Dummett assumes that when Frege spoke of a thought as being built out of parts, Frege was implying that there is a unique analysis of that thought into its component parts.

---

<sup>48</sup>To see what Dummett has in mind here, consider his example: ‘If anyone killed Brutus, then he was an honorable man.’ From this sentence one can infer ‘If Brutus killed Brutus, then Brutus was an honorable man.’ This inference is plausibly explained in virtue of the shared pattern ‘If ( ) killed Brutus, then ( ) was an honorable man’ present in both sentences. But this pattern will not correspond to a constituent of the thought expressed by the latter sentence on Dummett’s view, since the latter thought is uniquely analyzed into the sense corresponding to the material conditional and the two thoughts expressed by ‘Brutus killed Brutus’ and ‘Brutus was an honorable man’, each having a unique analysis of their own. Recognizing such a pattern in a thought that is not a constituent of it corresponds to decomposing it in a particular way. (ibid pg. 273ff)

<sup>49</sup>Two examples are Currie (1985, pg. 286) and Garavaso (1991, pg. 201). There are various other objections to Dummett’s interpretation that we needn’t discuss here, since we will see that Dummett’s distinction between analysis and decomposition is unmotivated.



That is, Dummett assumes that Frege thought of the part/whole distinction as being such that a whole can only be built up out of parts in a unique way. But (1) there is ample textual evidence that Frege thought of wholes as being analyzable into parts in multiple, distinct ways. (2) Frege's discussion of thoughts being built up out of their parts does not imply that thoughts are the result of a temporal process, since for Frege thoughts are timeless, abstract entities. Furthermore, Frege's own examples provide cases in which a whole is not only analyzable in distinct ways; it can be built up out of parts in distinct ways as well.

Once we recognize these interpretive points, there is no need to postulate an analysis/decomposition distinction, for Frege's treatment of thoughts as built up out of parts in no way contradicts his treatment of thoughts as having multiple analyses. Given that Frege thought of wholes more generally as having multiple decompositions into parts, Frege can quite consistently speak of a thought as built up out of parts and as having multiple decompositions.<sup>50</sup>

Let's first consider the textual evidence for (1). Here are four quotations which demonstrate that Frege thought of wholes as divisible into parts in a variety of nonequivalent ways, each of which occurred in the context of Frege distinguishing extensions of concepts from aggregates (a distinction which, he argued, the set theorists failed to appreciate):

---

<sup>50</sup>Levine (2002, pg. 202ff) also made essentially this point against Dummett, although unfortunately it has been overlooked by many Frege scholars. In fairness to Dummett, in his earlier work (e.g. (Dummett 1981)) he mentions the possibility of a part/whole relation in which the whole can be decomposed into parts in multiple ways (I pointed this out above). His example is that of a country that can be divided into component regions in different ways depending on one's purposes in doing so. (ibid pg. 264) The essential issue is Dummett's insistence that a thought can only be *built up out of* parts in a unique way, thus making the B-theses nontrivially incompatible with the A-theses. I address this issue below.

What Mr. Schröder calls ‘inclusion’ or ‘subsumption’ is here, properly speaking, nothing but the part–whole relation, extended in such a way that every whole is to be treated as a part of itself. From the point of view we are now adopting, we do not need the words ‘individual’ and ‘single thing’. Divisibility can be imagined as going on *ad infinitum*...and we have no need at all to assume there are parts insusceptible of further division; so perhaps it is better at this stage not to talk about elements at all. (1895, pg. 434-435/211)

[I]f we are given a whole, it is not yet determined what we are to envisage as its parts. As parts of a regiment I can regard the battalions, the companies or the individual soldiers, and as parts of a sand pile, the grains of sand or the silicon and oxygen atoms. On the other hand, if we are given a class, it is determined what objects are members of it. (Letter to Russell 7/28/1902, PMC pg. 140)

It is determined by a concept which objects fall under it; it is not determined by an aggregate what should count as its parts, whether, for example, in a regiment they are the individual soldiers, the companies, or the battalions; or whether, in a chair, they are the atoms, the molecules, or the artificially joined pieces of wood. (1903a, pg. 150)

The aggregate is composed of its parts. Whereas the extension of a concept is not composed of the objects that belong to it...Now of course it can happen that all objects which belong to the extension of a concept are at the same time parts of an aggregate and what is more in such a way that the whole being of the aggregate is completely exhausted by them...A grain of sand is an aggregate. And it can be that the extension of the concept *silicic acid molecule contained in this grain of sand* apparently coincides with the aggregate which we call this grain of sand. But we could just as well let the extension of the concept *atom contained in this grain of sand* coincide with our aggregate. But in that case the two extensions of concepts would coincide, which is impossible. From which it follows that neither of the two extensions of concepts coincides with the aggregate, for if one of them were to do so, then the other could with equal right be said to do so. (1906a, PW pg. 183)

This establishes that a whole can be decomposed into parts in multiple ways, on Frege’s view. Let’s turn to point (2): can a whole be *built up out of* parts in multiple ways? Dummett’s claimed inconsistency in Frege hinges on a negative

answer to this question. But as these examples already suggest, there is no reason to think that Frege thought a whole must be built up out of parts in a unique way.<sup>51</sup> More importantly, note Frege's remark in one of the very papers that Dummett cites as evidence for an inconsistency: "Compound Thoughts". Frege says, "By filling the gaps [in the unsaturated expression] with expressions of thoughts, we form the expression of a compound thought of the second kind. But we really should not talk of the compound thought as originating this way, for it is a thought and thoughts have no origin." (1923-26, pg. 40/394). The point being, thoughts are timeless entities and shouldn't be thought of as the result of a temporal process. That is merely a helpful metaphor and should not be taken so literally.

Finally, here is a direct quotation from Frege on the matter: "We must notice, however, that one and the same thought can be split up in different ways and so can be seen as put together out of parts in different ways." (1906c, PW pg. 201-2)

Hence, the apparent inconsistency in Frege dissolves. Frege can perfectly well speak of a thought as built up out of parts while denying that a thought has a unique analysis into its components, for each distinct analysis of the thought will reveal distinct parts of the thought that combine into the whole.<sup>52</sup> Consider again Frege's example of a line divided by a point. One can see this point as dividing the line into two parts, one which contains the point and one which doesn't. And yet,

---

<sup>51</sup>Take the third example. Constructing a chair by arranging its atoms is a different procedure than constructing it by arranging its component pieces of wood.

<sup>52</sup>Returning to Dummett's four theses, we can accept all four once we remove the insistence in thesis A2 that a thought is not built up out of its component senses, which is in no way implied by the latter half of the thesis: the constituents of a thought being arrived at by analysis of it. One should instead say that a thought can be built up out of constituents in multiple ways.

if we divided the very same line at another point, we would have a distinct pair of parts that combine to form the very same line. These parts can further be divided into parts *ad infinitum*. Crucially, one should *not* think of the line as unstructured. Indeed, it is in virtue of the structure of the line that it is capable of being divided into parts in this way.<sup>53</sup> <sup>54</sup>

Frege held an analogous view of the structure of thoughts. To return to our simple example, the thought that John loves Mary can be seen as composed of the function-sense expressed by ‘( ) loves Mary’ and the object-sense expressed by ‘John’, but it can also be seen as composed of the function-sense expressed by ‘John loves ( )’ and the object-sense expressed by ‘Mary’. Furthermore, it can be seen as composed of the function-sense expressed by ‘( ) is loved by John’ and the object-sense ‘Mary’, and in various other ways as well. Each way of decomposing it corresponds to a particular analysis of the whole thought into senses as parts.

## 1.8 Consequences

As we have seen, Gottlob Frege held that a thought does not have a unique decomposition into parts, but rather the same thought can be decomposed into parts in

---

<sup>53</sup>As any elementary topology textbook makes clear, the real line has quite a bit of topological structure that distinguishes it from other topological spaces. Analogously, a thought can have quite a bit of structure and yet still be decomposable in multiple ways.

<sup>54</sup>Several Frege scholars have modified Frege’s account of thoughts in order to resolve its seeming-inconsistency, including (Currie 1985), (Bell 1996), (Bermudez 2001), (Penco 2003), (Textor 2009), and (Kemmerling 2011). We lack the space to discuss each scholar’s resolution in detail, and must instead simply note that none of these scholars recognize Frege’s treatment of the part/whole distinction, nor how it combines with Frege’s function/argument analysis of thought and his elegant geometric metaphor regarding the nature of functions into a tenable position on the structure of thought.

various ways. We have seen two primary examples of this: (1) the transition from a sentence in active voice to the equivalent sentence in passive voice does not affect the thought expressed, and (2) any sentence containing names<sup>55</sup> expresses a thought that can be decomposed into a complete sense corresponding to one of those names and a predicative sense corresponding to the remainder of the sentence. One could also decompose a thought with multiple names into complete senses corresponding to two or more names (or instances of the same name) and a predicative sense corresponding to a relational expression. Since sentences often contain multiple names or multiple instances of the same name, examples of type (2) will often suggest multiple decompositions of the same thought.

How far do examples in Frege of multiple decompositions extend? In Section 5 I showed that many examples of type (2) lead to multiple decompositions of the same thought that have distinct logical forms. In fact, Frege pointed out that distinct analyses may have distinct logical forms explicitly in (1906b, PW pg. 187), (1906c, PW pg. 201-2), and (Letter to Linke 8/24/1919, PMC pg. 98).

A more controversial example of multiple decompositions is Frege's famous re-carving of the judgeable content that line  $a$  is parallel to line  $b$  into the judgeable content that the direction of line  $a$  is identical with the direction of line  $b$ , a transition meant to allow the thinker to acquire the concept of direction (1884, pg. 74-5). It is unclear whether Frege would have still accepted this type of example of multiple decompositions after he made his sense/reference distinction, since it is not mentioned

---

<sup>55</sup>Here (following Frege) a *name* is understood broadly as any expression meant to denote an entity, and hence a definite description counts as a name on Frege's view. The reader unhappy with this terminology may prefer the less-natural expression 'entity-expression'.

in the *Grundgesetze*.<sup>56 57</sup>

Another radical example is Frege's assertion that the sentences 'There is at least one square root of 4', 'The concept *square root of 4* is realized', and 'The number 4 has the property that there is something of which it is the square' all express the same thought. (1892c, pg. 199/188) Decomposing this thought in the manner suggested by the first sentence results in a logical form with an existential quantifier as its main connective, while the latter two suggested decompositions have distinct atomic logical forms, with the first including an object-sense referring to a proxy-object for the concept *square root of 4*<sup>58</sup> while the second includes an object-sense referring to the number 4.<sup>59</sup>

In "Compound Thoughts" (1923-26) Frege identifies the thoughts expressed by several logically equivalent types of sentences, including 'A and B' and 'B and A'

---

<sup>56</sup>Frege *did*, however, bring up similar examples in a letter to Bertrand Russell of 7/28/1902, stating that so long as the relation in question is an equivalence relation (to use the modern expression), "this relation can be transformed into an equality (identity)..." (PMC pg. 141). Whether this counts as a case of multiple decompositions of the same thought depends upon what Frege meant by "can be transformed into" or "can be replaced by".

<sup>57</sup>Here is another example where Frege may have changed his mind: in "Function and Concept", Frege speaks of the two sides of a particular case of Basic Law V as "express[ing] the same sense, but in a different way." (1891, pg. 10-11/143) However, there is no evidence in the *Grundgesetze* that Frege continued to think of Basic Law V in this way.

<sup>58</sup>Note that for Frege the expression 'the concept *square root of 4*' actually denotes a proxy-object rather than a concept. This isn't the place to discuss the "concept horse" problem beyond noting this point and the resulting different decompositions of the expressed thought on Frege's view.

<sup>59</sup>Near the end of the same paper ("On Concept and Object"), Frege makes several remarks about the saturated/unsaturated distinction that suggest he thought of e.g. 'The number 2 is a prime number', 'The number 2 falls under the concept *prime number*', 'The *falls under* relation holds between the number 2 and the concept *prime number*', 'The *holds between* relation holds among...', etc., as all expressing the same thought. (1892c, pg. 204-05/193) However, he does not say so explicitly, instead stating "It is thus easy for us to see that the difficulty arising from the 'unsaturatedness' of one part of the thought can indeed be shifted, but not avoided." (ibid pg. 205/193) If Frege meant that the unsaturatedness of one part of the thought can be shifted to another part of the same thought, then this would indeed count as a case of multiple decompositions of the same thought. Furthermore, such cases fit naturally with Frege's explicit examples discussed above.

(ibid pg. 39/393), ‘A and A’ and ‘A’ (ibid pg. 39 fn. 5/393 fn. 21), ‘not (not B)’ and ‘B’ (ibid pg. 44/399), and ‘not[(not A) and A]’ and ‘If A then A’ (pg. 50/405). He even suggests in a letter to Husserl that so long as two sentences don’t contain logically self-evident component parts, if both sentences follow logically from each other, they express the same thought. (Letter to Husserl 12/9/1906, PMC pg. 70)<sup>60</sup>

61

These various remarks on multiple decompositions suggest that thoughts are grasped in particular *ways* corresponding to particular decompositions of those thoughts into parts. And note that some decompositions of a thought will make various inferences from that thought salient while masking others. For example, if I grasp the thought that John loves Mary as consisting of the object-sense expressed by ‘John’ composed with the predicative sense expressed by ‘loves Mary’, this will immediately make salient a logical inference to the thought that *someone* loves Mary, while if I grasp this very same thought as consisting of the object-sense expressed by ‘Mary’ composed with the predicative sense expressed by ‘John loves’, this will instead immediately make salient a logical inference to the thought that John loves *someone*, while masking the inference to the thought that someone loves Mary.

As I emphasized above, the point is the following: *In the context of making an*

---

<sup>60</sup>Why the qualification? I suggest Frege was thinking of the following sort of case: compare the sentences ‘John loves Mary and either Jim has brown hair or Jim does not have brown hair’ and ‘John loves Mary’. Although the two sentences are logically equivalent, the former contains additional truth-conditional content not found in the latter, and hence expresses a distinct thought. The thought expressed by the former must include an object-sense referring to Jim, while the latter thought contains no such component. Hence, logical equivalence is not a sufficient condition for thought identity.

<sup>61</sup>There are additional cases of multiple decompositions stemming from Frege’s remarks on truth (e.g. 1892a, pg. 34/164), indexicals (e.g. 1918-19a, pg. 64/358), and (arguably) the hierarchy of senses (1892a, pg. 28/159, 34/164, and 37/166). We will discuss these cases in Chapters 2 and 3.

*inference, we must grasp the thoughts involved in particular ways.* Each way of grasping a thought makes various possible inferences salient while masking others. Hence Dummett was right to link multiple decompositions to possible inferences, but still wrong to treat analysis as a separate enterprise revealing the true structure of the thought. Much like the real line, the thought is structured in such a way that it can be decomposed into parts in many distinct ways, each of which reveals a different aspect of its underlying structure.

In the next chapter, we'll continue our discussion of the structure of sense by considering Frege's hierarchy of senses. In both Chapter 2 and Chapter 3 we will see how the Multiple Decompositions Thesis is not only the correct interpretation of Frege; it also resolves various issues concerning the structure of thought.



## 1.9 Appendix

I call the first component a function, the second its argument. This distinction has nothing to do with the conceptual content, but only with our way of viewing it. Although, in the mode of consideration just indicated, “hydrogen” was the argument and “being lighter than carbon dioxide” the function, we can also apprehend the same conceptual content in such a way that “carbon dioxide” becomes the argument and “being heavier than hydrogen” the function. (Frege 1879, pg. 15-16/126)

Consider now as our example: “the circumstance that the centre of mass of the solar system has no acceleration, if only internal forces act on the solar system”. Here, “solar system” occurs in two places. We can thus consider this as a function of the argument “solar system” in various ways, according as we think of “solar system” as replaceable by other arguments at the first or the second or at both places—but in the last case, replaceable by the same thing both times. These three functions are all different. The proposition that Cato killed Cato shows the same thing. Here, if we think of “Cato” as replaceable at the first occurrence, then “killing Cato” is the function. If we think of “Cato” as replaceable at the second occurrence, then “being killed by Cato” is the function. Finally, if we think of “Cato” as replaceable at both occurrences, then “killing oneself” is the function. (ibid pg. 16/127)

For us, the different ways in which the same conceptual content can be considered as a function of this or that argument have no importance so long as function and argument are completely determinate. (ibid pg. 17/128)

*If we imagine that in a function a symbol, which has so far been regarded as not replaceable, is now replaceable at some or all of the places where it occurs, we then obtain, by considering it in this way, a function with another argument besides the one it had before. In this way, functions of two or more arguments arise. Thus, for example, “the circumstance that hydrogen is lighter than carbon dioxide” can be considered a function of the two arguments “hydrogen” and “carbon dioxide”. (ibid pg. 17-18/128)*

We can also consider a symbol already previously regarded as replaceable [at some places] as now further replaceable at those positions where it was previously considered constant. (ibid pg. 17 fn./128 fn.)

As opposed to this, I start out from judgements and their contents, and not from concepts...We may now express  $2^4 = 16$  by the sentences ‘2 is a fourth root of 16’ or ‘the individual 2 falls under the concept “4th root of 16”’ or ‘belongs to the class of 4th roots of 16’. But we may also just as well say ‘4 is a logarithm of 16 to the base 2’. Here the 4 is being treated as replaceable and so we get the concept ‘logarithm of 16 to the base 2’:  $2^x = 16$ . (Frege 1880-81, PW pg. 16-17)

The fourth example gives us the concept of a multiple of 4, if we imagine the 12 in [Frege’s notation for 12 being a member of a series beginning with 0 and continuing to add 4 at each stage—NB] as replaceable by something else; the concept of the relation of a number and its multiple if we imagine the 4 as also replaceable; and the concept of a factor of 12 if we imagine the 4 alone as replaceable. (ibid pg. 32-33)

I do not believe that for any judgeable content there is only one way in which it can be decomposed, or that one of these possible ways can always claim objective pre-eminence. (Frege, Letter to Marty 8/29/1882, PMC pg. 101)

The judgement “line  $a$  is parallel to line  $b$ ”, or, using symbols,  $a // b$ , can be taken as an identity. If we do this, we obtain the concept of direction, and say: “The direction of line  $a$  is identical with the direction of line  $b$ .” Thus we replace the symbol  $//$  by the more generic symbol  $=$ , through removing what is specific in the content of the former and dividing it between  $a$  and  $b$ . We carve up the content in a way different from the original way and this yields us a new concept. (Frege 1884, pg. 74-75)

In the same way with the definitions of fractions, complex numbers and the rest, everything will in the end come down to the search from a judgement-content which can be transformed into an identity whose sides precisely are the new numbers. In other words, what we must do is fix the sense of a recognition-judgement for the case of these numbers. (ibid pg. 114-15)

If we understand ‘ $x^2 - 4x = x(x - 4)$ ’ in the same sense as before, this expresses the same sense, but in a different way. (Frege 1891, pg. 11/143)

Thus, we split up this sign for the True,  $3 > 2$ , into ‘3’ and ‘ $x > 2$ ’. We can further split up the ‘unsaturated’ part ‘ $x > 2$ ’ in the same way,

into '2' and  $x > y$ , where 'y' enables us to recognize the empty place previously filled up by '2'. (ibid pg. 27/154)

One can, indeed, say: 'The thought that 5 is a prime number is true.' But closer examination shows that nothing more has been said than in the simple sentence '5 is a prime number'. The truth claim arises in each case from the form of the assertoric sentence, and when the latter lacks its usual force, e.g., in the mouth of an actor upon the stage, even the sentence 'The thought that 5 is a prime number is true' contains only a thought, and indeed the same thought as the simple '5 is a prime number'. (Frege 1892a, pg. 34/164)

The sense of the sentence, 'After Schleswig-Holstein was separated from Denmark, Prussia and Austria quarrelled' can also be rendered in the form 'After the separation of Schleswig-Holstein from Denmark, Prussia and Austria quarrelled.' (ibid pg. 42 fn. 9/170 fn. 14)

Thus the sense of the sentence previously used ['If a number is less than 1 and greater than 0, then its square is less than 1 and greater than 0'—NB] can be given in the form 'The square of a number which is less than 1 and greater than 0 is less than 1 and greater than 0'. (ibid pg. 44/172)

The thought of our sentence ['If the Sun has already risen, the sky is very cloudy'—NB] might also be expressed thus: 'Either the Sun has not risen yet or the sky is very cloudy' – which shows how this kind of sentence connection is to be understood. (ibid pg. 46 fn. 13/173 fn. 17)

Even the expressions 'singular content of possible judgement', 'particular content of possible judgement' are not quite accurate in that they ascribe to the content itself an attribute which, strictly speaking, belongs to it only under a certain form—a certain way of analysing it into subject and predicate. (Frege 1892b, PW pg. 107-08)

In the sentence 'There is at least one square root of 4' we assert that the first level concept *square root of 4* falls under a concept of second level, whereas in the sentence 'The concept *square root of 4* is realized' we assert that the object *the concept square root of 4* falls under the first level concept *concept that is realized*. We do indeed have the same thought in the two concepts, [sic.] but this, being analysed differently, is construed in a different way. (ibid PW pg. 110)

In the sentence ‘there is at least one square root of 4’, we have an assertion, not about (say) the definite number 2, nor about  $-2$ , but about a concept, *square root of 4*; viz. that it is not empty. But if I express the same thought thus: ‘The concept *square root of 4* is realized’, then the first six words form the proper name of an object, and it is about this object that something is asserted. But notice carefully that what is asserted here is not the same thing as was asserted about the concept. This will be surprising only to somebody who fails to see that a thought can be split up in many ways, so that now one thing, now another, appears as subject or predicate. The thought itself does not yet determine what is to be regarded as the subject. If we say ‘the subject of this judgment’, we do not designate anything definite unless at the same time we indicate a definite kind of analysis; as a rule, we do this in connexion with a definite wording. But we must never forget that different sentences may express the same thought. For example, the thought we are considering could also be taken as saying something about the number 4: ‘The number 4 has the property that there is something of which it is the square’. (Frege 1892c, pg. 199/107)<sup>62</sup>

If we expressed this thought in the way that we gave above, we should have [Frege’s notation for the function  $x^2 = 1$  having the same course-of-values as the function  $(x + 1)^2 = 2(x + 1)$ —NB]. What we have here is that second level relation which corresponds to, but should not be confused with, equality (complete coincidence) between objects. If we write [Frege’s notation for the value of the function  $x^2 = 1$  being identical to the value of the function  $(x + 1)^2 = 2(x + 1)$  on any argument—NB], we have expressed what is essentially the same thought, construed as an equation between values of functions that holds generally. We have here the same second level relation; we have in addition the sign of equality, but this does not suffice on its own to designate this relation: it has to be used in combination with the sign for generality: in the first line we have a general statement but not an equation. (Frege 1892-95, PW pg. 121)

If I assert that the sum of 2 and 3 is 5, then I thereby assert that it is true that 2 and 3 make 5. So I assert that it is true that my idea of Cologne Cathedral agrees with reality, if I assert that it agrees with reality. (Frege 1897, PW pg. 129)

---

<sup>62</sup>Given that ‘square root of 4’ in the first sentence refers to a concept and ‘the concept square root of 4’ in the second sentence refers to an object, and given that concepts are *not* objects, it follows from these two sentences expressing the same thought that this thought must be decomposable into senses in different ways, each resulting in distinct senses as “parts” of the thought.

The distinction between the active and passive voice belongs here too. The sentences ‘M gave document A to N’, ‘Document A was given by N [to] M’, ‘N received document A from M’ express exactly the same thought; we learn not a whit more or less from any one of these sentences than we do from the others. Hence it is impossible that one of them should be true whilst another is false. It is the very same thing that is here capable of being true or false. (ibid PW pg. 141)

Again in the two sentences ‘Frederick the Great won the battle of Rossbach’ and ‘It is true that Frederick the Great won the battle of Rossbach’, we have, as we said earlier, the same thought in a different verbal form. (ibid PW pg. 141)

From the linguistic point of view, what is to be considered the subject is determined by the form of the [sentence]. The situation is different when considered from the logical point of view. We may decompose the [sentence] ‘ $8 = 2^3$ ’ either into ‘8’ and “is the third power of 2,” or into ‘2’ and “is something whose third power is 8,” or into ‘3’ and “is something which, when the power of 2, yields 8.” (Frege 1903b, pg. 372 fn. 12/281 fn. 13)

A sentence such as ‘Two is a prime’ can be analysed into two essentially different component parts: into ‘two’ and ‘is a prime’...In the sentence ‘Two is a prime’ we find a relation designated: that of subsumption. We may also say the object falls under the concept *prime*, but if we do so, we must not forget the imprecision of linguistic expression we have just mentioned. (Frege 1906a, PW pg. 177-8)

We may speak of the negation of a thought before we have made any distinction of parts within it...We may, metaphorically speaking, regard the predicative component of a thought as a covering for the subject-component. If further coverings are added, these automatically become one with those already there. (Frege 1906b, PW pg. 185)

We should mention that, strictly speaking, it is not in itself that a thought is singular, but only with respect to a possible way of analysing it. It is possible for the same thought, with respect to a different analysis, to appear as particular. (ibid PW pg. 187)

Now if the same proper name occurs in both consequent and antecedent, we may regard the hypothetical thought as singular if we think

of it as being analysed into the complete part that corresponds to the proper name and the unsaturated part left over. (ibid PW pg. 188)

So we are assuming that we have a hypothetical thought—one which can at the same time be construed as a singular thought—from which we, as was said above, always obtain a true thought by keeping the unsaturated part fixed, whatever complete part we saturate it with. (ibid PW pg. 189)

If several proper names occur in a sentence, the corresponding thought can be analysed into a complete and unsaturated part in different ways. The sense of each of these proper names can be set up as the complete part over against the rest of the thought as the unsaturated part. (ibid PW pg. 192)

In fact at bottom the sentence ‘it is true that 2 is prime’ says no more than the sentence ‘2 is prime’. (ibid PW pg. 194)

Let us now go further and split the proper name ‘3 – 2’ up into the proper name ‘2’ and the unsaturated part ‘3 – ’. Now we may also split the original sentence ‘3 – 2 > 0’ up into the proper name ‘2’ and the unsaturated part ‘3 – > 0’. The [referent] of this is the concept of something that yields a positive number when subtracted from 3. (ibid PW pg. 195)

We must notice, however, that one and the same thought can be split up in different ways and so can be seen as put together out of parts in different ways. The word ‘singular’ does not apply to the thought in itself but only with respect to a particular way of splitting it up. Each of the sentence-parts ‘1 is greater than 2’ and ‘1<sup>2</sup> is greater than 2’ can also be seen as put together out of the proper name ‘1’ and an unsaturated part. The corresponding holds for the related thoughts. (Frege 1906c, PW pg. 201-02)

It seems to me that an objective criterion is necessary for recognizing a thought again as the same, for without it logical analysis is impossible. Now it seems to me that the only possible means of deciding whether [sentence] *A* expresses the same thought as [sentence] *B* is the following, and here I assume that neither of the two [sentences] contains a logically self-evident component part in its sense. If *both* the assumption that the content of *A* is false and that of *B* true *and* the assumption that the

content of  $A$  is true and that of  $B$  false lead to a logical contradiction, and if this can be established without knowing whether the content of  $A$  or  $B$  is true or false, and without requiring other than purely logical laws for this purpose, then nothing can belong to the content of  $A$  as far as it is capable of being judged true or false, which does not also belong to the content of  $B$ ; for there would be no reason at all for any such surplus in the content of  $B$ , and according to the presupposition above, such a surplus would not be logically self-evident either. In the same way, given our supposition, nothing can belong to the content of  $B$ , as far as it is capable of being judged true or false, except what also belongs to the content of  $A$ . Thus what is capable of being judged true or false in the contents of  $A$  and  $B$  is identical, and this alone is of concern to logic, and this is what I call the thought expressed by both  $A$  and  $B$ . (Frege, Letter to Husserl 12/9/1906, PMC pg. 70-1)

$2^3 = 8$  can be decomposed in different ways into a saturated and unsaturated part. (Carnap's Student Notes, Winter Semester 1910-11, in: Awodey, Reck, and Gabriel 2004, pg. 19/67)<sup>63</sup>

If the *definiens* occurs in a sentence and we replace it by the *definiendum*, this does not affect the thought at all. It is true that we get a different sentence if we do this, but we do not get a different thought. (Frege 1914, PW pg. 208)

Whenever a proper name occurs in an assertoric sentence, we can regard the remaining part as a concept-sign. (ibid PW pg. 229)

'The thought that  $3 > 2$  is true' can be more simply said by the sentence '3 is greater than 2'. (ibid PW pg. 233)

For the same reason as before I put this also in the dependent form 'that it is true that sea-water is salt[y]'. The thought expressed in these words coincides with the sense of the sentence 'that sea-water is salt[y]'. So the sense of the word 'true' is such that it does not make any essential contribution to the thought. If I assert 'it is true that sea-water is salt[y]', I assert the same thing as if I assert 'sea-water is salt[y]'. (Frege 1915, PW pg. 251)

---

<sup>63</sup>Carnap also included a diagram of such decompositions on page 27/75.

It is also worth noticing that the sentence ‘I smell the scent of violets’ has just the same content as the sentence ‘It is true that I smell the scent of violets’. (Frege 1918-19a, pg. 61/354)

If someone wants to say today what he expressed yesterday using the word ‘today’, he will replace this word with ‘yesterday’. Although the thought is the same its verbal expression must be different in order that the change of sense which would otherwise be effected by the differing times of utterance may be cancelled out. The case is the same with words like ‘here’ or ‘there’. (ibid pg. 64/358)

Such a designation [‘the negation of the negation of  $A$ ’—NB] is to be regarded as directly composed of the parts: ‘the negation of —’ and ‘the negation of  $A$ ’. But it may also be regarded as made up of the parts: ‘the negation of the negation of —’ and: ‘ $A$ ’. (Frege 1918-19b, pg. 156/387)

I compare that which needs completion to a wrapping, e.g. a coat, which cannot stand upright by itself; in order to do that, it must be wrapped round somebody. The man whom it is wrapped round may put on another wrapping, e.g. a cloak. The two wrappings unite to form a single wrapping. There are thus two possible ways of looking at the matter; we may say either that a man who already wore a coat was now dressed up in a second wrapping, a cloak, or, that his clothing consists of two wrappings – coat and cloak. (ibid pg. 157/388)

So I do not begin with concepts and put them together to form a thought or judgement; I come by the parts of a thought by analysing the thought. (Frege 1919, PW pg. 253)

A saturated part obtained by analysing a thought can sometimes itself be split up in the same way into a part in need of supplementation and a saturated part. The sentence ‘The capital of Sweden is situated at the mouth of Lake Mälär’ can be split up into a part in need of supplementation and the saturated part ‘the capital of Sweden’. This can further be split up into the part ‘the capital of’, which stands in need of supplementation, and the saturated part ‘Sweden’. Splitting up the thought expressed by a sentence corresponds to such a splitting up of the sentence. (ibid PW pg. 255)

If a thought is an equation, this does not exclude its also being a subsumption. The [sentence] ‘Napoleon is the loser of Waterloo’ can be



changed into ‘Napoleon is identical with the loser of Waterloo’; and here Napoleon is subsumed, not indeed under the loser of Waterloo, but under the concept ‘identical with the loser of Waterloo’. (Frege, Letter to Linke 8/24/1919, PMC pg. 98)

That ‘B and A’ has the same sense as ‘A and B’ we may see without proof by merely being aware of the sense. (Frege 1923-26, pg. 39/393)

Another case of the sort is that ‘A and A’ has the same sense as ‘A’. (ibid pg. 39 fn. 5/393 fn. 21)

If, therefore, ‘not [A and B]’ expresses a compound thought, then ‘not [B and A]’ express the same compound of the same thoughts. (ibid pg. 40/394)

Compound thoughts of the first four kinds have this in common, that their component thoughts may be interchanged. (ibid pg. 43/397)

But since ‘not (not B)’ has the same sense as ‘B’... (ibid pg. 44/399)

But the sense of ‘A and (not B)’ is the same as that of ‘(not B) and A’... (ibid pg. 48/403)

For any ‘A’ that is a sentence proper, ‘A and A’ expresses the same thought as ‘A’; the former says no more and no less than the latter. (ibid pg. 49/404)

We can also render this compound thought [expressed by ‘not [(not A) and A]’—NB] verbally by the expression ‘If A, then A’... (ibid pg. 50/405)

Thus, the compound expressed by ‘(A and B) and C’ is composed of the thoughts expressed by ‘A and B’ and ‘C’. But we can also treat it as composed of the thoughts expressed by ‘A’ and ‘B’ and ‘C’. (ibid pg. 50-51/405-06)

## Chapter 2

# Meta-Representation and the Hierarchy of Senses

### 2.1 Introduction

When I make a typical assertion while speaking or writing, whether my assertion is true or false will depend upon which entities I'm referring to and whether or not they actually bear the properties and relations I'm ascribing to them. When I say "John has brown hair", for example, whether my assertion is true or false will depend upon whether the person referred to by the name 'John' has the property referred to by the predicate 'has brown hair' at the time at which I make the assertion. In fact, its truth-value seems to be a *function of* the referent of the proper name, the referent of the predicate, and the time of utterance; no further information about the utterance

is needed in order to determine the truth-value of the assertion.<sup>1</sup> In particular, the truth-value would have remained the same even if I had used another expression with the same referent as ‘John’ when making my assertion. Similar remarks apply to a relational claim such as “John loves Mary.”

However intuitive such a Fregean approach to giving the truth-conditions of an assertion may seem in the case of uttering such a simple tensed subject-predicate sentence, things get quite a bit more complicated when we start considering cases in which I ascribe a *belief* to someone. When I say “Mary believes that John has brown hair”, for example, I may say something true even though if I had substituted the expression ‘John’ with ‘the man sitting in the corner’ I would have said something false, even if John is in fact the man sitting in the corner. Mary might not realize that John is the man sitting in the corner, having just dyed his hair blonde earlier today.<sup>2</sup> When I ascribe a belief to someone in an act of assertion, whether my assertion is true or false seems to depend on that person’s *way of thinking* of the entities their belief is about rather than merely the entities themselves.

---

<sup>1</sup>We may suppose that the predicate refers to a *property*, namely, having brown hair, although this assumption can of course be disputed. The important point for present purposes is that the truth-value of the assertion seems to depend upon the *referent* of the predicate as opposed to its sense, syntactic structure, etc. If I had used another predicate that picks out the property of having brown hair in a different way, the truth-value of my assertion would have been the same. If one thinks that predicates refer to *extensions*, one can accept the same point with respect to distinct predicates having the same extension. So we needn’t get too caught up in debates about the nature of predicates or properties (or Frege’s saturated/unsaturated distinction) for now.

<sup>2</sup>To remove some of the artificiality of the example, let’s suppose that Mary has just gone to the bar to get a drink, and I’m pointing out to my friend that Mary hasn’t yet realized that the blonde man sitting in the corner, who we were all just discussing, is actually our good friend John. I hope that this clarifies the intuition that if I had instead asserted “Mary believes that the man sitting in the corner has brown hair”, I would have said something false (in this context, both assertions would be intended with a *de dicto* rather than a *de re* reading (on which more later), and hence the former could be true while the latter is false). Mary is well aware that the man sitting in the corner has blonde hair; we were just discussing him a few minutes ago.

Gottlob Frege himself thought it obvious that when I make an assertion ascribing a belief to someone, by uttering “Copernicus believed that the planetary orbits are circles”, for example, expressions such as ‘the planetary orbits’ refer to the *senses* or *modes of presentation* they ordinarily express rather than the entities they ordinarily refer to: “It is quite clear that in this way of speaking words do not have their customary [reference] but designate what is usually their sense.” (1892a, pg. 28/159)

Indeed, if we accept that belief is a relation between a person and a proposition, where a proposition is the *representational content* or *Fregean thought* expressed by an utterance of a sentence in an act of assertion,<sup>3</sup> and a proposition is composed of *concepts* or *ways of thinking* of objects, properties, and relations, then Frege’s position can seem quite plausible.<sup>4</sup> When I made that assertion about Copernicus, I asserted that he bore the belief relation to the proposition that the planetary orbits are circles. Hence, it seems natural to say that the expression ‘Copernicus’ referred to Copernicus, the expression ‘believed’ referred to the belief relation, and the expression ‘that the planetary orbits are circles’ referred to that proposition, the very proposition

---

<sup>3</sup>By the *representational content* of an assertion I mean the sort of content relevant to determining whether the assertion is true or false. This is the sense in which I intend to use both ‘proposition’ and ‘thought’ in this chapter. Compare Frege: “So one has to separate off from the content of a sentence the part that alone can be accepted as true or rejected as false. I call this part the thought expressed by the sentence.” (1906c, PW pg. 197-8)

Note that there are many conceptions of ‘proposition’ in the literature and we needn’t take a stance on whether any of these is *the* proposition expressed in an assertion. I *will* be assuming, in line with the discussion above about belief ascriptions needing to take the ascriber’s ways of thinking into account, that belief is a relation to a content from which a thought can be abstracted.

<sup>4</sup>Here I am using ‘concept’ in the sense used by many contemporary philosophers of mind (as roughly equivalent to Frege’s use of ‘Sinn’ (‘sense’) rather than Frege’s use of ‘Begriff’ (‘concept’)), and will continue to do so throughout this dissertation. Cf. Peacocke (2008) or Burge (2010), for example. Peacocke in particular uses ‘sense’ and ‘concept’ interchangeably, as does Kripke in his essay on Frege (2008, pg. 259 fn. 16 in the 2011 reprint), and I will follow this practice here. Once again, I will *not* be following Frege’s own usage of ‘concept’ for the rest of this dissertation, and will instead follow the contemporary usage to avoid confusion.

that I would have expressed and asserted to be true if I had instead uttered “The planetary orbits are circles”. Given that that proposition is composed of senses, it seems natural to further say that the expressions ‘the planetary orbits’ and ‘are circles’ referred to the *senses* ordinarily expressed by them.

This position on belief ascription allowed Frege to continue to treat the truth-value of such an assertion as a function of the *referents* of the expressions uttered.<sup>5</sup> Since expressions *refer to* their ordinary senses in the context of belief ascriptions, the truth-value of such an assertion can depend upon the ascriber’s *way of thinking* of the entities the belief ascribed to them is about and still be a function of the referents of the uttered expressions. According to Frege, in such a context the expressions used have *indirect* reference and express *indirect* senses, which can be described as modes of presentation of modes of presentation (1892a, pg. 28/159).

Furthermore, since I can embed a belief ascription within another belief ascription, by asserting “Frege believed that Copernicus believed that the planetary orbits are circles”, for example, it seems that in such contexts expressions will express *doubly* indirect senses in order to refer to their *doubly* indirect referents, the latter being their (singly) indirect senses.<sup>6</sup> Given that I can in principle continue embedding belief ascriptions within other belief ascriptions to any finite level of depth, this Fregean picture seems to be committed to every expression that expresses a sense in an unembedded context also potentially expressing an infinite number of *distinct* senses,

---

<sup>5</sup>It may also depend upon various contextual factors such as the time of utterance. We will discuss such complications in the next chapter.

<sup>6</sup>Frege considers expressions having doubly indirect reference (“indirect reference of the second degree”) explicitly in his (Letter to Russell 12/28/1902, PMC pg. 154).

with the particular sense expressed depending upon the number of embeddings. This is Frege's infinite hierarchy of senses.

Donald Davidson famously complained that if a natural language really attached infinitely many possible senses to the same expression in the manner suggested, it would be unlearnable since "there is no rule that gives the reference in more complex contexts on the basis of the reference in simpler ones." (1968, pg. 99 in the 1984 reprint)<sup>7</sup> In the same essay, Davidson claimed that "If we could recover our pre-Fregean semantic innocence, I think it would seem to us plainly incredible that the words 'The earth moves', uttered after the words 'Galileo said that', mean anything different, or refer to anything else, than is their wont when they come in other environments." (ibid, pg. 108)<sup>8</sup>

I already gave some reasons above for thinking that if so, our pre-Fregean selves would be mistaken. And yet, when we consider cases like my asserting "Mary believes that John has brown hair, but he actually doesn't", the most natural treatment of the anaphoric reference of the expression 'he' seems to be that it inherits its reference from the expression 'John', but this can't be right if the expression 'John' refers to its ordinary sense (senses don't have hair).<sup>9</sup> On the other hand, if I had asserted

---

<sup>7</sup>See also his (1965, pg. 14-5 in the 1984 reprint), where he makes the unlearnability point explicitly. All of the approaches to the hierarchy we'll consider below attempt to address this objection directly by specifying a recursive reference rule.

<sup>8</sup>Burge (2005, pg. 170ff) and Kripke (2008, pg. 267) both point out that the hierarchy of indirect senses has an analogue in direct quotation (e.g. the word 'John' refers to John, the expression 'John' refers to the word 'John', etc.), and both suggest that consideration of the analogous direct quotation hierarchy ought to have removed some of the force of the unlearnability objection. I would add that in ordinary speech, speakers often use an expression to refer to itself rather than its ordinary referent in a way that listeners must determine from the context. Compare the two verbal utterances "John only has one child" and "John only has one syllable". This puts pressure on Davidson's "semantic innocence" objection as well.

<sup>9</sup>I may owe this point to Gary Ostertag.

“Mary believes that the man sitting in the corner is suspicious, but she doesn’t realize that he is John”, I couldn’t have simply replaced ‘he’ by ‘John’ while preserving the truth-value, since Mary is well aware that John is John. Setting delicate issues about anaphora aside, it still seems that in making such an assertion, I am thinking about *both* Mary’s way of thinking about John *and* John himself, and this latter fact is what allows me to immediately and effortlessly transition to other thoughts about John. If we want to hold onto our Fregean view, we’re going to need a more complex account of reference in such cases; losing our semantic innocence certainly comes with a cost.<sup>10</sup>

We seem to run into similar issues once we start considering *inferences*. Suppose I think to myself, “Frege believed that Copernicus believed that the planetary orbits are circles, but it’s false that the planetary orbits are circles; hence some proposition that Frege believed that Copernicus believed is false.” This seems like a valid inference, and yet if the expression ‘that the planetary orbits are circles’ in the second premise refers to the proposition that the planetary orbits are circles, while that same expression in first premise refers to a *way of thinking* of that proposition rather than the proposition itself, then how could this inference be valid?<sup>11</sup> Replacing the second

---

<sup>10</sup>With that said, this paper is not an attempt to treat the semantics of natural language, including rules governing anaphoric reference. As I suggested (following Frege) in footnote 3 above, I’m interested in the sort of content relevant to determining truth, not ordinary linguistic meaning. Senses are components of such content, and I would further suggest that a sense is the aspect of the content of an expression in virtue of which it contributes to the truth-conditions of complete declarative sentences which contain it. These cases are merely meant to suggest that even in the context of a belief ascription, reflection on the reference of expressions gives us some reason to think that customary senses are expressed in addition to indirect senses. I give further reasons in what immediately follows.

<sup>11</sup>This is a modification of an example from Peacocke (1996, pg. 143), who has revised his views on the hierarchy since that publication. We’ll consider his more recent views when we discuss his (2008) below.

premise by the more natural-sounding ‘but the planetary orbits aren’t circles’ seems to make the problem even worse, since this expression *expresses* the negation of that proposition and doesn’t seem to refer to a proposition at all, although perhaps the intuition that we have an instance of a valid argument-form without the need for additional premises is weaker in this case.<sup>12</sup>

Furthermore, many transitions in thought beyond deductive inference seem to involve thinking about the *customary* referents of expressions (i.e. their referents in *direct* contexts) rather than their *indirect* referents, even when those expressions occur embedded in belief ascriptions. When I think to myself “Mary believes that John has brown hair”, this seems to immediately make salient transitions to various other thoughts about John, with no need for me to first think about a reference-relation holding between Mary’s way of thinking of John (the purported referent of the previous use of the expression ‘John’) and John himself. Furthermore, it doesn’t seem like someone could understand the proposition expressed by that sentence without also understanding the proposition that John has brown hair.<sup>13</sup> Since this isn’t true of other expressions that might refer to that very same proposition, e.g. ‘Bob’s favorite proposition’ or ‘the proposition expressed by the sentence written on the blackboard in room 314’, the indirect sense of ‘John has brown hair’ had better present its referent,

---

<sup>12</sup>More carefully: replacing the second premise in this way results in the following on the Fregean view we’re considering: ‘the planetary orbits are circles’ in premise 1 expresses its doubly indirect sense and refers to its indirect sense. However, ‘the planetary orbits aren’t circles’ in premise 2 expresses the *negation of* the former expression’s customary sense (i.e. the proposition that the planetary orbits are circles) and refers to a *truth-value* (well, at the very least, it doesn’t seem to refer to the proposition it is the negation of). Finally, in the conclusion we’ve existentially quantified over the customary sense of ‘the planetary orbits are circles’, which (as we’ve just pointed out) isn’t referred to in *either* premise. So we seem forced to conclude we’ve engaged in an especially egregious fallacy of equivocation. And yet, the argument still seems intuitively valid as it stands.

<sup>13</sup>As we’ll see below, the latter intuition is shared by Burge, Peacocke, and Kripke.



the customary sense of that same sentence, in a very “transparent” or “revelatory” way.<sup>14</sup>

All of these examples suggest that in the context of a belief ascription, *both* the indirect sense of an expression *and* its customary sense are somehow expressed in a way that both reveals their special connection and makes either available for use in various transitions in thought. In this chapter, I will attempt to explain how that could be, and how such seemingly-technical issues might relate to substantive issues concerning the role of meta-representation in both perceptual experience and more abstract thought.

To accomplish this, I will develop a theory of the structure of senses relevant to the hierarchy, grounded in my interpretation of Frege. In order to do so, I will begin by discussing what Frege actually says about the hierarchy of senses in “On Sense and Reference” (1892a), understood in the light of Frege’s commitment to the Multiple Decompositions Thesis as defended in Chapter 1. I will then give a more careful formulation of Frege’s actual views on the hierarchy. In the following section, I will consider other recent approaches to the hierarchy due to Tyler Burge, Christopher Peacocke, and Saul Kripke. In the final section I will turn toward more general reasons for thinking of senses as having the sort of structure I’ve argued for, including giving various reasons why one might expect such an account given the role of meta-representation in first-person judgments and inferences and the phenomenology of conscious perceptual experience.

---

<sup>14</sup>I owe this point to (Kripke 2008), which we’ll discuss below.

## 2.2 Frege's Hierarchy of Senses

Many philosophers writing about Frege complain that when he introduces indirect senses in “On Sense and Reference” (1892a)—to account for shifts in the reference of expressions in belief ascriptions and other indirect contexts—he never tells us what those indirect senses actually are. Michael Dummett wrote, “What, then, is the indirect sense of an expression? Frege has told us what its indirect referent is, namely its ordinary sense; but that is not enough to determine what its indirect sense is to be.” (1981, pg. 267) In particular, there are presumably infinitely many senses referring to the ordinary sense of an expression ‘A’ (e.g. the sense expressed by ‘the sense of the expression ‘A’’, the sense expressed by ‘Bob’s favorite sense’, etc.); which of these will be *the* indirect sense expressed by that expression in an indirect context, and *referred to* by that expression in a doubly indirect context?

Actually, Frege does give us some idea of what the indirect sense of a complete sentence is supposed to be: “So here [the case of indirect quotation—NB], the subordinate clause has for its [referent] a thought, not a truth-value and for its sense not a thought, but the sense of the words ‘the thought that (etc.)’, which is only a part of the thought in the entire complex sentence.” (1892a, pg. 37/166)<sup>15</sup> Here Frege quite explicitly asserts that the indirect sense of an assertoric sentence ‘*p*’ is the customary sense of the expression ‘the thought that *p*’.<sup>16</sup> In his illuminating essay on Frege, Saul Kripke acknowledges this while adding “Well *of course* this is Frege’s view; but

---

<sup>15</sup>Recall that for Frege the customary referent of an assertoric sentence is its truth-value: the True or the False.

<sup>16</sup>Here I am using ‘*p*’ as a schematic letter, meant to be replaced by an assertoric sentence in a particular case, and will continue to do so.

it says nothing as to what the indirect senses of the significant parts of the sentence are, nor how they combine to give the sense of the whole.” (2008, pg. 259)

Furthermore, it isn't yet clear what sort of relation the indirect sense has to the customary sense beyond simply referring to it. Is the customary sense a *part* of the indirect sense? This might explain the intuition that one can't grasp the indirect sense without grasping the customary sense, but it quickly leads to problems.<sup>17</sup> In particular, the most natural way for Frege to treat the customary sense of a sentence as a part of its indirect sense is to treat the sense of 'the thought that ( )' (where, just as Frege used it above, this incomplete expression is meant to be completed by an assertoric sentence) as a function from thoughts to indirect senses, much like he treats the sense of a predicate as a function from the sense of a singular term to a thought, that is, the thought expressed by the sentence resulting from composing that predicate with that singular term.<sup>18</sup> But this simply won't work, since then the expression 'the thought that ( )' would have to *refer to* a function from *truth-values* to thoughts, and yet as soon as we apply that expression to two sentences expressing distinct thoughts with the same truth-value, this "function" won't be well-defined, since it would thereby have to map that truth-value to both thoughts.<sup>19</sup> This had

---

<sup>17</sup>This idea about the structure of senses is also naturally suggested by several popular notations (such as that used by (Burge 2005) and (Peacocke 2008), as we'll discuss below) in which an expression expressing the indirect sense of another expression includes the latter expression as a component: e.g. '<e>' expresses the indirect sense of the expression 'e' and refers to its customary sense.

<sup>18</sup>This interpretation has been controversial. I defended it in the previous chapter and will discuss it further immediately below.

<sup>19</sup>More carefully: if 'the thought that ( )' expresses a function-sense that is itself a function from thoughts to the indirect senses referring to those thoughts, then it *refers to* a "function" from truth-values to thoughts, since on such a Fregean view thoughts refer to truth-values and the indirect senses in question (the outputs of the purported sense of 'the thought that ( )') refer to thoughts (namely, the corresponding inputs). But consider e.g. 'the thought that snow is white' and 'the

better not be what Frege had in mind.

Why think that Frege held that the sense of a predicate is *itself* a function from a sense to a complete thought containing both senses as parts? I defended this interpretation in the previous chapter and will repeat that defense here. On Frege's view, when I assert the sentence 'Copernicus believed that the planetary orbits are circles', the expressions 'the planetary orbits' and 'are circles' refer to their customary senses, and the complete expression 'the planetary orbits are circles' refers to the proposition customarily expressed by it. Hence, the referent of this complete expression ought to be a function of the referents of 'the planetary orbits' and 'are circles', and since the former expression is complete while the latter is unsaturated, the latter must refer to a function on Frege's view, in particular, a function mapping the referent of the former to the complete thought. Hence, the customary sense of 'are circles' not only *refers* to a function; it *is* a function from senses to thoughts. In this case, it maps the customary sense of 'the planetary orbits' to the complete thought expressed by 'the planetary orbits are circles', a thought which contains both senses as parts.

In my view, much of the resistance to treating the sense of a predicate as a function from senses to thoughts while at the same time being a part of the thoughts it maps to comes from various prejudices concerning functions that have little to do with Frege's conception of them. In Frege's "Function and Concept" (1891, pg. 7/141), he uses the following analogy to explain the sense in which functions are unsaturated while objects are not: Consider a line divided by a single point. Now treat the line as divided

---

thought that grass is green'. Both 'snow is white' and 'grass is green' are true, hence this "function" will have to map the True to both the thought that snow is white and the thought that grass is green. But these are distinct thoughts. Hence our "function" isn't well-defined.

into two segments, one which includes the point and one which doesn't. The segment which doesn't include the point is supposed to be analogous to a function while the other segment, being "complete in itself", is analogous to the object. Referring to the former, Frege says, "Only by completing it with this endpoint, or with a line that has two endpoints, do we get from it something entire." (loc. cit.) Extending Frege's analogy, what does this function map the other line-segment *to*? The whole line, of course. And yet at the same time, this function is a part of that line. More generally, it combines its arguments with itself in order to form a longer line-segment.

I think that this is a perfectly acceptable notion of "function", and find it plausible that this is how Frege was thinking of the unsaturated senses of predicates as both functions from complete senses to thoughts and parts of those thoughts. This is not to say that *all* functions map their arguments to values which include those very functions and arguments as parts, which Frege rejected explicitly in his (1919, PW pg. 255), nor is this meant to somehow fully explain the nature of understanding a predicate. The point is merely that nothing in Frege's conception of functions implies that a predicative sense can't both be a part of its value (given a complete sense as argument) and count as a function.<sup>20</sup>

With that said, it had better not have also been Frege's view that the sense of 'the thought that ( )' maps thoughts to their indirect senses, for the reasons we gave above. Indeed, it wasn't. In the last chapter I defended the view that Frege did *not* think that a thought has a unique decomposition into senses, but rather the same

---

<sup>20</sup>In fact, prior to making his sense/reference distinction, Frege seemed to hold this view of conceptual contents carved into arguments and functions. Cf. his (1880-81, PW pg. 16) in combination with Section 9 of his (1879).

thought can be decomposed into senses in a variety of distinct ways. More specifically, grasping a thought in a particular way makes salient a variety of possible inferences to other thoughts while masking others. Now we can put that interpretation to work.

## 2.3 Interpreting Frege

Let's consider one of the most puzzling cases in which Frege claims that two distinct sentences express the same thought. There are many places in his writings in which Frege says that any assertoric sentence ' $p$ ' expresses the very same thought as the sentence 'the thought that  $p$  is true' or 'it is true that  $p$ '. He does so in "On Sense and Reference", for example: "...even the sentence 'The thought that 5 is a prime number is true' contains only a thought, and indeed the same thought as the simple '5 is a prime number'." (1892a, pg. 34/164)<sup>21</sup> Note that the expression 'the thought that  $p$ ' refers to, well, *the thought that  $p$* , and hence if ' $p$ ' and 'the thought that  $p$  is true' always express the same thought, then on Frege's view, *any* thought can be decomposed into a complete sense *referring to* that whole thought and a predicative sense (expressed by 'is true') mapping that complete sense to its referent, also the whole thought. Any thought can be grasped as including a sense referring to that thought as a part.<sup>22</sup>

---

<sup>21</sup>There are many other examples throughout his later work. Cf. (1897, PW pg. 141), (1906b, PW pg. 194), (1914, PW pg. 233), (1915, PW pg. 251), and (1918-19a, pg. 61/354).

<sup>22</sup>The reader may feel that this structure is somehow circular. As a useful analogy (first suggested to me by Sidney Felder in a general discussion of self-reference), consider a map of a region located within that very region (e.g. a map inside a shopping mall labeled "you are here"). The region includes a representation of itself as a part. One can easily imagine someone first designing the map and then placing it within the region, with no circularity involved, even if the map has enough detail to represent its own location as occupied by itself. We will discuss this structure further in the final

Could Frege simply have been a bit sloppy in making this claim? This seems difficult to believe, given the number of times he said this over the course of nearly 30 years (see footnote 21). Notice also that Frege never says that a sentence ‘ $p$ ’ expresses the same thought as ‘the *sentence* ‘ $p$ ’ is true’. This would require the thought that  $p$  to contain a part referring to a *sentence* expressing it (since the expression ‘the sentence ‘ $p$ ’ ’ refers to ‘ $p$ ’, obviously), and given that any thought can be expressed by indefinitely many distinct sentences, the thought that  $p$  would have to include distinct senses as parts referring to each possible vehicle for expressing it. This was obviously not Frege’s view, and hence it is no surprise that he carefully avoided asserting it, even though he recognized that we often ascribe truth to sentences.<sup>23</sup>

Furthermore, as I pointed out above, in the same essay in which Frege first makes this claim, he uses the expression ‘the thought that  $p$ ’ explicitly to *refer to* the thought customarily expressed by ‘ $p$ ’ while *expressing* the indirect sense of ‘ $p$ ’ (1892a, pg. 37/166), making it very unlikely that he intended this expression to refer to anything other than the thought that  $p$  (or to express a sense distinct from the indirect sense of ‘ $p$ ’) in the context of a sentence of the form ‘the thought that  $p$  is true’. I think the evidence strongly supports our taking Frege at his word, which implies that the predicative sense of ‘is true’ maps the sense of ‘the thought that  $p$ ’ to the thought that  $p$ .<sup>24</sup>

---

section.

The reader may have also noticed that this example makes essential use of the truth predicate. We will discuss Frege’s views on truth immediately below.

<sup>23</sup>Burge mistakenly ascribes this view to Frege in his “Postscript to “Frege on Truth”” (2005, pg. 144), in part (I suspect) because he does not recognize Frege’s commitment to the Multiple Decompositions Thesis.

<sup>24</sup>This is also supported by Frege’s claim that when I assert that the thought that  $p$  is true,

Why might Frege have said this? Every time he makes this claim, it is in the context of arguing that the sense of the predicate ‘is true’ doesn’t seem to add anything to the thought expressed, and yet he often also suggests that the word ‘true’ reveals the essence of logic. This is especially clear in his unpublished “My Basic Logical Insights” (1915, PW pg. 251-2):

If I assert ‘it is true that sea-water is salt[y]’, I assert the same thing as if I assert ‘sea-water is salt[y]’. This enables us to recognize that the assertion is not to be found in the word ‘true’, but in the assertoric force with which the sentence is uttered...So the word ‘true’ seems to make the impossible possible: it allows what corresponds to the assertoric force to assume the form of a contribution to the thought...‘true’ only makes an abortive attempt to indicate the essence of logic, since what logic is really concerned with is not contained in the word ‘true’ at all but in the assertoric force with which a sentence is uttered.

As Frege often points out, even though uttering the word ‘true’ might seem to indicate that one is asserting something (imagine uttering “It is true that John is our prime suspect”), an actor on a stage could utter the very same words without actually asserting them.

And just how does the word ‘true’ do this? For Frege, to make an assertion is to express or communicate a judgment, and a *judgment* is a transition from a thought to a truth-value.<sup>25</sup> And as I argued above, the sense of ‘is true’ in the context

---

I’m asserting that the relation between the sense of a sign and its referent holds as opposed to the subsumption relation between an object and a property: “If we say ‘the thought is true’, we seem to be ascribing truth to the thought as a property. If that were so, we should have a case of subsumption. The thought as an object would be subsumed under the concept of the true. But here we are misled by language. We don’t have the relation of an object to a property, but that of the sense of a sign to its [referent].” (1906b, PW pg. 194) Once again, this is a straightforward consequence of my interpretation, under which the predicative sense of ‘is true’ maps the sense of ‘the thought that  $p$ ’ to the thought that  $p$ , and thus refers to a function from the thought that  $p$  to its referent.

<sup>25</sup>Cf. (Letter to Husserl 5/24/1891, PMC pg. 63) and (1892a, pg. 34-5/164-5).



of a sentence of the form ‘the thought that  $p$  is true’ must map the sense of ‘the thought that  $p$ ’ to the thought that  $p$ , since this sentence expresses the very same thought as ‘ $p$ ’ itself, carved in a different way. Hence, it *refers to* a function from thoughts to their truth-values. *This* is the sense in which the word ‘true’ “makes an abortive attempt to indicate the essence of logic”: when attached to a sentence, it represents the transition from the expressed thought to its truth-value characteristic of a judgment without actually bringing about that transition. When applying the sense of ‘is true’, we remain at the level of thoughts, just as Frege said, rather than passing to the level of reference. More specifically, we pass from the second level of the hierarchy of senses to the first.

As is probably clear by now, I claim that Frege held the following thesis:

*The indirect sense of a sentence ‘ $p$ ’ is the customary sense of the expression ‘the thought that  $p$ ’, where the customary sense of the latter expression is a part of the thought customarily expressed by ‘ $p$ ’ made salient via carving that thought in the manner suggested by the sentence ‘the thought that  $p$  is true’.<sup>26</sup>*

Because this indirect sense is a *part* of that thought first grasped by carving it in this way, Frege can thereby explain the intuition that one can’t grasp the indirect sense of a sentence without also grasping the thought it refers to. Furthermore, there is nothing to stop him (or us) from treating e.g. the predicative sense of the expression ‘Copernicus believes’ as a function from the *thought* that the planetary orbits are circles (as well as the indirect sense referring to that thought) to the thought that Copernicus believes that the planetary orbits are circles, and hence he can further

---

<sup>26</sup> “Carving a thought” is of course a metaphor meant to correspond to grasping that thought as decomposed into senses in a particular way (compare (Frege 1884, pg. 74-5)).

explain the intuition that in indirect contexts *both* the customary sense *and* the indirect sense are expressed in a way that makes either available for use in judgment and inference.<sup>27</sup>

What of Kripke's complaint that Frege hasn't told us what the indirect senses of significant parts of the sentence are, nor how they are combined to give the indirect sense of the whole sentence? By an argument analogous to the one I gave at the end of the previous section, Frege is committed to the indirect sense of a sentence being a function of the indirect senses of its parts, and in a way that suggests how he might have answered Kripke: Consider again the sentence "Frege believed that Copernicus believed that the planetary orbits are circles." In this context, the sentence 'the planetary orbits are circles' refers to its indirect sense and expresses its doubly indirect sense. Hence, on Frege's view its referent ought to be a function of the referents of its parts. Since 'the planetary orbits' is a complete expression while 'are circles' is an unsaturated expression, the latter must refer to a function on Frege's view, and in particular a function from the referent of the former to the referent of the whole. Hence, the indirect sense of 'are circles' is itself a function from the indirect sense of 'the planetary orbits' to the indirect sense of 'the planetary orbits are circles'. In this way, the structure of the indirect sense of a sentence mirrors the structure of the thought that indirect sense refers to, and parts of the indirect sense correspond to the significant parts of the sentence expressing it.

Furthermore, since the structure of the indirect sense of a sentence mirrors the

---

<sup>27</sup>We will discuss this issue in more detail after first considering other recent approaches to the hierarchy.

structure of the thought it refers to, one might expect the doubly indirect sense of a sentence customarily expressing that thought to itself be a part of the (singly) indirect sense of that sentence. Consider again a sentence of the form ‘the thought that  $p$  is true’, which Frege claimed expresses the same thought as ‘ $p$ ’. Since this is also an assertoric sentence, it ought to express the same thought as ‘the thought that the thought that  $p$  is true is true’. Hence, in this context ‘the thought that the thought that  $p$  is true’ ought to express the same sense as is customarily expressed by ‘the thought that  $p$ ’, since the sense it expresses will be mapped by the customary sense of ‘is true’ to the thought customarily expressed by ‘ $p$ ’. Might the customary sense of that expression contain the doubly indirect sense of ‘ $p$ ’ as a part?

Note that the expression ‘the thought that’ creates an indirect context in which the expressions within its scope *refer* to the senses they would have expressed if that expression hadn’t been attached to them. But the customary sense of ‘the thought that  $p$ ’ is, as Frege said, the indirect sense of ‘ $p$ ’. Hence, in the context of the expression ‘the thought that the thought that  $p$  is true’, the expression ‘the thought that  $p$ ’ *refers* to the indirect sense of ‘ $p$ ’, the very same sense that is *expressed* by the complete expression. Hence, in that context, the expression ‘the thought that  $p$ ’ *expresses* a sense that is mapped by the indirect sense of ‘is true’ to the customary sense of ‘the thought that  $p$ ’, which is itself the indirect sense of ‘ $p$ ’. By identifying this sense (the indirect sense of ‘the thought that  $p$ ’) with the doubly indirect sense of ‘ $p$ ’ we ensure that the structure of the indirect sense of a sentence mirrors the structure of the thought it refers to, and in particular the doubly indirect sense of that sentence is a *part* of its (singly) indirect sense, a part which is mapped by the

*indirect* sense of ‘is true’ to the (singly) indirect sense of that sentence, as desired. By repeating this procedure, we can ascend up the hierarchy as high as we like.

Hence, Frege has in fact told us enough about the hierarchy of senses to determine quite a bit of its structure, given Frege’s other commitments regarding the structure of thoughts and the nature of judgment and truth. Still, one is left with the feeling that a significant part of the story is missing. *Why* might Frege have held that any thought can be carved in such a way that it includes a sense as a part referring to that whole thought? Can we motivate such a view independently of both the role it plays in Frege’s conception of judgment and truth and its ability to resolve more technical issues involving the hierarchy of indirect senses? These questions will occupy us for the remainder of this chapter. In order to make progress on this front, let’s turn to three other recent approaches to Frege’s hierarchy of senses, all of which emphasize connections to broader philosophical issues.

## 2.4 Recent Approaches to the Hierarchy

In his “Postscript to “Frege and the Hierarchy”” (2005, pg. 167-210), Tyler Burge defends Frege’s infinite hierarchy of senses against a variety of objections. He points out that when I use a that-clause to make a belief ascription, I rely on my *understanding* of the customary senses of the expressions within the that-clause in order to fix those senses as the referents of those expressions, and in grasping their indirect senses I have little else to go on than that understanding of the customary senses. (ibid pg. 170) He suggests that recognizing this is part of the key to understanding

what indirect senses are.

Burge claims that there is a function from the customary sense of a sentence to the *canonical sense* that determines that customary sense, where this canonical sense is expressed by a canonical name of that customary sense. In English, this canonical name will be a that-clause formed by attaching the expression ‘that’ to the sentence in question. Understanding the indirect sense of a sentence (the canonical sense that determines its customary sense) requires no more than understanding its customary sense and knowing how to use a that-clause-forming expression. Hence, the indirect sense of a sentence is uniquely fixed by its customary sense, the doubly indirect sense is uniquely fixed by its (singly) indirect sense, etc., and Burge concludes that there is no particular problem with learning such an infinite hierarchy. (ibid pg. 172) *Contra* Davidson, one doesn’t need to grasp infinitely many primitive senses in order to learn such a language; the hierarchy can be recursively specified via such a canonical sense function.

Note that Burge uses the expression ‘determines’ to refer to the relation between the sense of an expression and its referent (or *denotation*, as Burge prefers) in order to distinguish it from the relation between the expression itself and its referent, while I used ‘refers’ above to refer to both. If the reader feels that these relations ought to be distinguished, I trust that they can determine which relation ought to be referred to via the context.

Frege himself considered the relation between *sense* and referent more fundamental than that between *expression* and referent. (Cf. his (1892a, pg. 27/159), and recall his oft-repeated claim that it is really thoughts that have truth-values rather

than sentences, e.g. (1918-19a, pg. 60/353).) The expression refers in virtue of its sense rather than vice versa.

Burge accepts the following *Principle for Canonical Names of Senses*: *The canonical name of a sense can be understood only if the sense that it names is understood.* (ibid pg. 174)

With respect to “relevant” canonical names, he also accepts a *Stronger Principle for Canonical Names of Senses*: *To think the sense of a canonical name of a sense, one must simultaneously think the lowest-level (ultimately, customary) sense in the downward hierarchy associated with the canonical name.* (ibid pg. 174)

Burge states these principles without argument. They are meant to elucidate how he is thinking of his canonical sense function, denoted by a functional expression ‘C’ in his sketch of a formal language. Note that in Burge’s formal language, he takes a first level of canonical names of customary senses *as primitive*, later following the convention of letting an expression of the form ‘<e>’ be the canonical name of the customary sense expressed by the expression ‘e’. *Given* such a canonical name ‘<e>’ of the customary sense of an expression ‘e’, one can construct a canonical name of the canonical sense *expressed by* ‘<e>’ using the functional expression ‘C’ (namely: ‘C(<e>)’), and continue applying this functional expression to ascend up the hierarchy, with each iteration expressing the canonical sense that determines the sense expressed by the previous iteration.

One does not apply the functional expression ‘C’ to produce the canonical name at the first level, however. In particular, we can’t treat the *sense* of ‘C’ as a function from arbitrary senses to canonical senses of those senses, since the “function” referred

to by ‘C’ would consequently not be well-defined for the same reasons I gave above when discussing a possible Fregean treatment of the sense of ‘the thought that ( )’ (consider what would happen if you were to apply ‘C’ to two expressions with the same reference but distinct senses). This may be why Burge takes the first level of canonical names as primitive, and furthermore why he explains the sense he intends ‘C’ to express by appealing to various principles as opposed to specifying its structure more directly.

However, this leaves it somewhat mysterious how Burge’s formal language is supposed to correspond to his account of the hierarchy. One might have thought that since ‘C(<Igor>)', for example, expresses a sense that is composed of the sense of ‘C’ and the sense of ‘<Igor>', one can’t grasp this sense (the canonical sense of the sense of ‘<Igor>’) without also simultaneously grasping the sense of ‘<Igor>' (the canonical sense of the sense of ‘Igor’), thus reflecting Burge’s commitment to the *Stronger Principle for Canonical Names of Senses*. But what about the sense of ‘<Igor>', or any canonical sense of a sense at the first level of the hierarchy? Since this expression is taken as primitive, one shouldn’t think of it as expressing a sense composed of the sense of ‘< >’ and the sense of ‘Igor’. Indeed, it isn’t clear how this could work anyway, since ‘< >’ can’t refer to a function from *objects* to senses for the reasons I just gave above. One wishes that Burge had said a bit more about how his formal language is meant to be understood. *Why* can’t I grasp the canonical sense of a sense without also grasping the sense it determines? Is this somehow reflected in the structure of the senses themselves?<sup>28</sup> Not if we take Burge’s formal language at

---

<sup>28</sup>Burge seems to suggest this in his more recent work: “In propositional cases, the lower-level

face-value, since the first level of canonical names of senses are (appearances to the contrary) *primitive expressions*.<sup>29</sup>

Burge also points out that although a thinker's being capable of grasping canonical senses of customary senses requires grasping the customary senses determined by them, it also involves a higher level of intellectual sophistication: "It is one thing to be able to think a thought and another to be able to attribute the thought to someone else, or to oneself." (ibid pg. 175) This point is surely correct. Indirect senses, on Burge's view, mark this additional attributive capacity. Burge later states that this additional layer of intellectual capacity is constituted by the sort of understanding involved in comprehending attributions of ways of thinking expressed in natural language, an understanding partially dependent on grasping those ways of thinking themselves. He goes on to claim that the senses of the initial layer of canonical names cannot be explained in any further way, and in fact such further explanation is unnecessary. (ibid pg. 177)

However, it is difficult to see this as much of an *explanation* at all, as opposed to a restatement of the standard Fregean view of the senses expressed in indirect contexts,

---

thought content is always preserved in the meta-representational thought by being *used*—thought—*in* the canonical name for the content." (2011, pg. 208 in the 2013 reprint)

<sup>29</sup>Not that Burge really suggests otherwise, although he could easily be misread as such. He initially uses the unstructured term 's' as a canonical name of a customary sense (2005, pg. 174), only later adopting the bracket-notation convention (ibid pg. 178), which he follows for the rest of the essay.

Notice that the view I argued above was Frege's own won't have this problem, since the indirect sense determining a particular thought will be a part of that thought rather than vice versa, with parts of the indirect sense corresponding to (and determining) parts of the thought it determines. In particular, since the indirect sense itself is a part of the thought it determines, it will include a part (the *doubly* indirect sense) corresponding to (and determining) the (singly) indirect sense itself. This nested pattern repeats up the hierarchy of senses, with each sense at level  $n$  (for  $n > 1$ ) being a part of the sense at the previous level  $n - 1$ . We will discuss why senses ought to have this structure in the next section.



coupled with Burge's brief remarks about grasping such senses requiring a capacity to attribute thoughts grounded on a capacity to think those thoughts. Furthermore, it seems quite implausible that no further explanation can be given. Burge has said very little about the additional intellectual capacity needed to think about one's own thoughts or the thoughts of others.<sup>30</sup> But note that such a capacity must play a central role in reflective judgment, for example. It is one thing to represent an object. It is quite another to reflectively consider whether a particular representation of an object is veridical for the purposes of making a judgment about it. The latter presupposes that I am capable of *representing* that representation, and furthermore representing it in such a way that reveals the *way* in which that representation is representing its object. I must represent that representation as having the sort of content that can be assessed as veridical or non-veridical. It seems clear that such a capacity for meta-representation of one's own representations as expressing such contents is the key to understanding the sorts of capacities we are presupposing thinkers to have when we judge them capable of grasping indirect senses.<sup>31</sup>

Christopher Peacocke, in his *Truly Understood* (2008), gives a more robust account

---

<sup>30</sup>Burge may have since recognized this need, since he does attempt to say more in more recent work. For example, in (Burge 2009) he ties our knowledge of our own thought contents to "reflexive, meta-representational specification of what we discursively understand" (ibid pg. 316), where such discursive understanding gives us a way of referring to those contents *de re* (ibid pg. 314-6). In the Dewey Lectures that I cited above, he also treats our meta-representational capacities as constitutive of being a critical reasoner. (2011, pg. 178ff) So, in his more recent work, Burge seems sympathetic to a position broadly in line with the one I defend below, although perhaps not its details.

<sup>31</sup>For example, when I reflectively consider whether I'm perceiving an object veridically, I'm relying on an ability to *represent* my perception as distinct from the object it is a perception of, and furthermore to represent it in a way that allows me to assess what sorts of properties and relations that perception is representing its object as possessing. Otherwise, I wouldn't be able to make sense of a perceptual illusion, for example, or even consider whether the world is as it appears to be. Comparing a representation with what it is purportedly a representation of is constitutive of reflective judgment, and my ability to do so relies upon an ability to form representations of representations.

of the capacities involved in grasping indirect senses. Peacocke's account is part of a more general theory of senses (or concepts) developed and defended throughout the book. According to Peacocke, concepts are individuated by their Fundamental Reference Rules, where a *Fundamental Reference Rule* (FRR) for a concept is "the rule that specifies what makes something the reference of the concept." (ibid pg. 56) For example, the FRR for the first-person concept *I* is the following: "what makes someone the reference of the first-person concept *I* in a thinking is that he or she is the thinker (the producer of that thinking)." (ibid pg. 56) That is, what makes the first-person concept the concept it is is that in order for someone to be its referent when it is employed by a thinker at a time, that person has to *be* the thinker of that thought at that time. The nature of the concept is given by its FRR. Furthermore, *grasping* a concept consists in having tacit knowledge of its FRR.<sup>32</sup>

This position allows Peacocke to give an elegant treatment of canonical concepts of concepts (i.e. indirect senses) in the final chapter of the book: since the essence of a concept is given by its FRR, what makes a concept the same concept as e.g. the concept *F* is, ultimately, having the same FRR as *F*. Hence, the FRR of the canonical concept of the concept *F* ought to be something like the following: what makes an arbitrary concept *C* the referent of the canonical concept of the concept *F* is having the very same FRR as *F*. This concept is arguably *canonical* since it is individuated in terms of the very FRR that individuates *F* itself. As Peacocke puts

---

<sup>32</sup>Regarding tacit knowledge, compare the tacit knowledge a young child has of a basic grammatical rule, knowledge which allows the child to distinguish grammatical from ungrammatical sentences without being able to articulate the rule itself. Peacocke intends to use tacit knowledge of FRRs analogously to explain a thinker's intuitions about particular cases of an object's falling under or not falling under a concept, and in such a way that can explain the rationality of the thinker's corresponding judgments.

it (with ‘can(F)’ referring to the canonical concept of F): “For an arbitrary concept C to fall under can(F) is for the fundamental condition for something to fall under C to be the same as the fundamental condition for something to fall under the concept F.” (ibid pg. 291)<sup>33</sup>

Peacocke notes that one can ascend to the next level of the hierarchy using an FRR of the same general form: simply replace ‘F’ by ‘can(F)’ throughout. This suggests a generating principle for uniquely determining the canonical concept of a concept from that concept itself, and Peacocke claims, *pace* Davidson, “Tacit grasp of this generating principle will allow thinkers to understand arbitrary embeddings of propositional attitude-verbs.” (ibid pg. 297)

What about the problematic inferences I mentioned in the Introduction, examples of which were discussed in Peacocke’s earlier work (1996, pg. 143) as part of an argument *against* the infinite hierarchy? Peacocke points out that one could simply add premises stating reference-relations between e.g. canonical concepts of propositions and those propositions themselves, but it seems that natural language and natural inference simply don’t work this way. As he puts it, “the language behaves as if “that *p*” refers to the same thing, however embedded, functioning as a complex term for the same Thought wherever it occurs.” (2008, pg. 299)<sup>34</sup>

---

<sup>33</sup>Peacocke adds a few qualifications to address both someone’s failing to know which concept F is (which shouldn’t be possible if that person is thinking of it via its canonical concept) and partial understanding of F (which is perfectly possible), but we needn’t worry too much about those qualifications here. In effect, on Peacocke’s full account a thinker possessing the canonical concept of F must have tacit knowledge of both the quoted condition and “enough” of the actual FRR of the concept F to be capable of propositional attitudes with contents containing F.

<sup>34</sup>Peacocke capitalizes ‘Thought’ in order to make it clear that he is referring to a Fregean thought, a “propositional conceptual intentional content”, rather than a particular mental event of thinking. (ibid pg. 284) As I stated in the Introduction, for the purposes of this chapter I’m using ‘proposition’ in this sense as well. See footnote 3 and the discussion immediately surrounding it.

Peacocke's response to his earlier self is to claim that in the case of concepts  $s$  of abstract objects  $x$  that are determined by what individuates  $x$ , if someone believes e.g.  $F(s)$ , then she believes *of* the object  $x$  that it is  $F$ . Since canonical concepts of Thoughts are determined in this way on Peacocke's account, if e.g.  $a$  believes that John believes that  $p$ , then it follows that concerning the Thought  $p$ :  $a$  believes John believes it. With a nod to Quine (1956), Peacocke calls this "exportation of concept". (ibid pg. 299-300) This allows him to (semi)-formalize the problematic inferences in a way that makes their validity explicit without adding additional premises. For example, the premise "John believes that Karl believes that Bush is powerful" becomes "Concerning the Thought  $\langle p \rangle$ , namely  $\langle \text{Bush} \rangle^{\wedge} \langle \text{is powerful} \rangle$ ,  $\text{Bel}(\text{John}, \langle \text{Bel} \rangle^{\wedge} \langle \text{Karl} \rangle^{\wedge} \langle p \rangle)$ ", and various seemingly-problematic inferences from it can then be (semi)-formalized in the obvious way. (ibid pg. 304-5)<sup>35</sup>

This response is a bit problematic, however. Most obviously, accepting it commits one to a strong metaphysical position on the nature of concepts of certain abstract objects, in addition to a strong position on the quite thorny issues of exportation and relational belief (and related issues involving *de re* attitudes). The reader is welcome to make of that what they will. More importantly, it seems insufficiently general.

---

<sup>35</sup>Like Burge, Peacocke uses bracket notation in his semi-formal language, where e.g. the expression ' $\langle p \rangle$ ' *refers to* the Thought that  $p$  and *expresses* the canonical concept of this Thought. Also, 'Bel' stands for the relevant belief relation and '^' stands for "predicational combination of concepts". (ibid pg. 284)

I should also note that in the example Peacocke is still treating John as thinking of this Thought under its canonical concept, even though this isn't reflected in the notation. It is instead a consequence of a further background assumption that having a *de re* attitude about a Thought is only possible if the thinker is thinking about it under its canonical concept. He points out that if we don't want to make this assumption, we can still capture John's way of thinking of this Thought while maintaining this position on exportation via a more complex notational scheme. (ibid pg. 305-7)

As I argued in the Introduction, it isn't merely expressions of the form 'that  $p$ ' that behave as if they refer to the same thing even when embedded in indirect contexts; expressions such as 'John' exhibit similar behavior.

For example, consider the following inference: "Mary believes that John has brown hair. John is a good friend of mine. Hence, John is both believed by Mary to have brown hair and a good friend of mine." Here again we have a case in which the language behaves as if 'John' refers to the same thing throughout the inference.<sup>36</sup> However, adopting Peacocke's strategy in such cases would amount to allowing for *unrestricted* exportation, something that he wants to avoid (as do I).<sup>37</sup> It would be preferable to have a more general account.<sup>38</sup>

Unlike Burge or Peacocke, Saul Kripke does not explicitly endorse Frege's hierarchy of senses in his essay "Frege's Theory of Sense and Reference: Some Exegetical

---

<sup>36</sup>One could attempt to capture this inference in classical logic by "cheating", since one could represent the conclusion as the conjunction of two sentences, one in which the indirect sense of 'John' is expressed (the first premise) and the other in which the customary sense is expressed (the second premise). But (1) it seems obvious that the conclusion is ascribing two relations to John himself, and that reference to John is preserved from the two premises. And (2) it's implicit in the conclusion that we're using the same name twice, which couldn't be done on this suggested formalization in classical logic. That's why it's cheating. I am grateful to Achille Varzi for helpful discussion of this case.

<sup>37</sup>Kripke gives various arguments against unrestricted exportation and in defense of the intuitive distinction between *de re* and *de dicto* belief (or, if one prefers: between *relational* and *notional* belief) in his "Unrestricted Exportation and Some Morals for the Philosophy of Language" (2011, pg. 322-50). In particular, Kripke argues (via a cleverly-designed definite description) that if we allow for unrestricted exportation, so long as someone has a single false belief, it will be true of that person that they believe e.g. *of the Eiffel Tower* that it has all sorts of ridiculous properties. (ibid pg. 328-30)

<sup>38</sup>It is also worth noting that Peacocke does not commit himself to Burge's principle that one can't think a canonical concept of a concept without also simultaneously thinking that concept, stating instead the weaker thesis that one can't *possess* a canonical concept of a concept without also *possessing* the latter concept. (2008, pg. 301) Relatedly, it is unclear whether he would think of a canonical concept of a concept as structured in such a way that it *contains* the concept it refers to (and is individuated in terms of) as a component, although his choice of notation might suggest this. Furthermore, unlike Frege, Peacocke does not treat predicative concepts as functions from concepts to Thoughts, and remains non-committal regarding how to think of "predicational combination of concepts".

Notes” (2008). Instead, he is primarily interested in explicating Frege’s position and defending it against various objections. One of Kripke’s main points is that whatever indirect senses are, they must be *revelatory* of their referents, where “a sense is revelatory of its referent if one can figure out from the sense alone what the referent is.” (ibid pg. 259)

In support of this, Kripke notes that the indirect sense of e.g. ‘the planetary orbits are circular’ can’t be the customary sense of the expression ‘the sense of ‘the planetary orbits are circular’ ’, since then someone could understand the sentence ‘Copernicus believed that the planetary orbits are circular’ without having any idea what ‘the planetary orbits are circular’ means, so long as they could recognize indirect quotation contexts in English and knew that in them words have indirect senses. According to Kripke, “this consequence is plainly absurd.” (ibid pg. 258-9) One shouldn’t be able to understand an indirect sense without also immediately knowing which sense it refers to.<sup>39</sup>

In response to worries about the hierarchy due to Davidson and others, and with several nods to Kaplan (1968) and (1989), Kripke points out various analogies with direct quotation (which I mentioned above but am omitting here), in addition to stating the standard Fregean rule for determining reference in indirect contexts:

( $\beta$ ): When words appear in indirect contexts, that is, “says that”, “believes that”, and so on, they refer to their senses in the clause following the “that”. (ibid pg. 268)

---

<sup>39</sup>Kripke defines a sense as *immediately* revelatory if “anyone who understands it must immediately know its referent”, and he is in fact arguing that indirect senses must be *immediately* revelatory. (ibid pg. 261) Compare Burge’s principle that one must simultaneously think a concept whenever one is thinking its canonical concept, and my arguments in the Introduction that both the customary sense and the indirect sense of a sentence are expressed in an indirect context.

Kripke emphasizes that this rule is to be understood iteratively, so that the reference can be determined recursively when we are dealing with iterative indirect quotations. This rule, applied in the simple uniterated case, gives us the indirect references of the words (their ordinary senses). So what are the indirect senses? Kripke continues, “Here we should surely say that the rule ( $\beta$ ) itself, applied to each indirect case, gives the indirect sense. For we have conceded that it determines the reference in each particular case, and whatever determines a reference is a Fregean sense. This consideration can be applied iteratively, and we appear to have a theory of the entire hierarchy.” (ibid pg. 268-9)<sup>40</sup>

So, for Kripke, the rule ( $\beta$ ) itself, each time it is used to specify a reference, implicitly gives the indirect sense, which can then be specified as a reference by another recursive application of the rule. Thus the infinite hierarchy of senses is determined by an explicit recursive rule.<sup>41</sup>

But how will an individual speaker actually apply rule ( $\beta$ ) in an iterated case? Yes, applying the rule to determine indirect reference also determines an indirect sense, but how will the speaker actually know what this indirect sense is, as would be needed in order to apply the rule again and refer to it? Kripke considers the analogous

---

<sup>40</sup>While it is true on Frege’s view that “once one has specified the reference one has specified the sense...because the sense is the way the reference or *Bedeutung* is specified” (ibid pg. 255), I do not think Frege would have agreed that whatever determines a reference *is* a Fregean sense. Senses are abstract objects, while reference can be determined by quite concrete actions such as gestures (that is, in conjunction with some sort of conventions for interpreting the gestures in the appropriate way). Such actions will *determine* a sense, but shouldn’t be *identified with* that sense. So, I would amend Kripke’s point and instead say that whatever determines a reference *gives* a Fregean sense. This doesn’t seem to affect his following points, and may have been a minor slip.

<sup>41</sup>Kripke follows this discussion by mistakenly attributing the unique decomposition thesis to Frege, albeit with some hesitancy. (ibid pg. 270-1, 277) I will address this in the next chapter when I discuss Kripke’s interpretation of Frege’s treatment of contextual factors.

*direct* quotation rule ( $\alpha$ ), according to which quoted words refer autonomously to themselves.<sup>42</sup> We know which words are referred to because in a particular case of writing or a particular utterance we see or hear (or produce) the token-words that make up the entire sentence-token, and are aware of the corresponding types as a result. Kripke compares this awareness of the types to Russellian acquaintance. (ibid pg. 271) I would add that we perceive the tokens *as* instances of their corresponding types.

Analogously, “Every time we determine a referent, we are introspectively acquainted with how the referent is determined, and that is the corresponding sense. And our introspective acquaintance with this sense gives us a way of determining it, and of referring to it, and this is the indirect sense...Each level of the hierarchy is the acquaintance-sense of the previous level.” (ibid pg. 271-2) For Kripke, this also explains why the indirect sense is immediately revelatory: it determines its referent by acquaintance, and hence no calculation is required to determine what the referent is. (ibid pg. 272)

Kripke goes on to bolster his argument that Frege was committed to what he calls “acquaintance-senses” by considering Frege’s treatment of indexicals, demonstratives, and the like, which we will discuss in the next chapter. Having now considered the other major approaches to the hierarchy of senses, let’s finally reconsider the account I defended in the last section as Frege’s own, paying careful attention to the variety

---

<sup>42</sup>There is some controversy regarding whether this was actually Frege’s view, as Kripke admits, and he is hesitant to fully endorse it as the proper interpretation. (ibid pg. 265-7) Whatever Frege’s actual view, I think it is clear that he *should have* adopted the view Kripke ascribes to him, but this isn’t the place to go into the matter.



of broader philosophical issues that have arisen in the course of our discussion.

## 2.5 Meta-Representation, Consciousness, and Judgment

Let's return to Frege's own account. Recall that on Frege's view, the indirect sense of an assertoric sentence ' $p$ ' is both the customary sense of 'the thought that  $p$ ' and a *part* of the thought customarily expressed by ' $p$ ', a part mapped by the sense of 'is true' to that same thought. Recall also that the structure of this indirect sense mirrors the structure of the thought it refers to (and is a part of), and I argued that one can maintain this mirroring structure up the hierarchy of senses by identifying the *doubly* indirect sense of ' $p$ ' with the *indirect* sense of 'the thought that  $p$ ', which will itself be mapped by the *indirect* sense of 'is true' to the indirect sense of ' $p$ '. That is, one ought to identify the doubly indirect sense of ' $p$ ' with the part of the singly indirect sense that corresponds to itself in the thought it refers to, much like how a highly detailed map of a region, within which that very map is located, might include a representation of itself as occupying a part of that region. This is the natural extension of Frege's view, and iterating it gives us an account of the structure of the infinite hierarchy of senses, as desired.<sup>43</sup>

---

<sup>43</sup>One might suspect that some circularity appears once we iterate this procedure up the entire infinite hierarchy, since how could the map contain a representation of itself so highly detailed that it contains a representation of itself which itself contains a representation of itself which...etc.? No such map could ever be constructed, one might think, even in principle! But note once again that thoughts aren't the results of temporal processes; they are abstract and atemporal. And it is important to recognize, as I argued in the last chapter, that complete thoughts are metaphysically prior to their parts on Frege's view. Thoughts have a unity and completeness that is metaphysically more fundamental than that of their parts (compare the unity of the real number line). This connects

Note that this structural account is compatible with various treatments of the recursion inherent in the hierarchy, including those of Peacocke and Kripke discussed above. One should also note the following: in Section 2 above we defined the indirect sense of a sentence ‘p’ as the customary sense of the expression ‘the thought that p’. As discussed in Section 3, we can then characterize the doubly indirect sense of ‘p’ as the part of the customary sense of ‘the thought that the thought that p is true’ that refers to the indirect sense of ‘p’, and so on. An arbitrary indirect sense can be characterized as the part of the customary sense of an appropriate expression (i.e. an expression that expresses a sense containing that sense as a part) that refers to the relevant sense at the previous level of the hierarchy.

For example, the indirect sense of ‘is tall’ could be characterized as the part of the customary sense of ‘the thought that John is tall’ that refers to the customary sense of ‘is tall’ (which is assumed to already be understood), the doubly indirect sense could be characterized as the part of the customary sense of ‘the thought that the thought that John is tall is true’ that refers to the just-characterized singly indirect sense, etc. So, we precede from an already-understood customary sense and isolate a part of it via its structure and the structure of its referent, the relevant parts of which are also assumed to be understood.<sup>44</sup>

But why hold such a view of the structure of thoughts? On this view, one could say that every thought represents itself as true, a conclusion Frege seemed happy to

---

with old debates about the unity of the proposition (not to mention lines in geometric space, or geometric space itself) that are beyond the intended scope of this chapter, so I’ll leave it at that for now.

<sup>44</sup>I am grateful to Anubav Vasudevan for pressing me to say more here.

accept. Note that this doesn't imply that any subject capable of thinking a thought is also capable of thinking *about* that thought, which we agreed with Burge above is far too strong. Some carvings of a thought a thinker grasps may be unavailable to that particular thinker (e.g. a young child) without further intellectual development and sophistication. For example, if a young child still lacks the concept "property", they may be unable to grasp the thought that John has brown hair in the way suggested by the sentence "The property of having brown hair is instantiated by John", even though (as I suggested in the last chapter) this sentence expresses that very same thought. Thinking about thoughts simply provides another example of this same phenomenon.

Let's move away from abstract thinking for a moment and consider an intimately connected domain: conscious experience. Much like you (I imagine), I am presently perceiving a region of physical space as occupied by various objects, including my own body. While typing that sentence, I considered that conscious perception itself as distinct from the region of physical space it is a perception of. I believe it is correct to conclude that my conscious experience just included a part representing that entire conscious perception.<sup>45</sup>

More broadly, whenever I think about all of the representations that make up a part of my present conscious experience, simply by virtue of doing so I'm ensuring that one of those very representations will represent that whole "space" of conscious representations. Hence, meta-level reflection on my own conscious representational experience has precisely the structure of Frege's treatment of canonical concepts of

---

<sup>45</sup>We will discuss perception further in the next chapter.

thoughts, according to my interpretation.<sup>46</sup> Furthermore, we shouldn't be surprised by structural connections between conscious perceptual experience and more abstract thinking. Surely the latter in some sense developed out of the former.<sup>47</sup>

I think we can make further progress by asking the following question: why is it cognitively useful for me to be able to represent my own conscious experience as structured in various ways? And the answer suggests itself: to aid in conscious decision-making leading to various actions I can perform, including mental actions such as making judgments, forming intentions, and the like.<sup>48</sup>

More specifically, by representing my conscious experience as structured in such a way that aspects of it represent various objects as instantiating properties and bearing relations to each other, I can *compare* how things appear to be with various other sources of information I have on how they actually are for the purposes of making a

---

<sup>46</sup>That is not to say, of course, that Frege ever considered this connection. In fact, Frege said little about the structure of self-consciousness. If Frege did notice this connection, why would he stay silent about it? This is admittedly very conjectural, but it may have been because (1) Frege wanted to sharply distinguish his position from the Kantians, and discussing a connection with the structure of self-consciousness might be seen as too much of a concession, and (2) Frege rarely discussed the structure of thoughts in detail, probably because he never developed a theory of thoughts (including their identity-conditions) to his own satisfaction. In any case, nothing I'm discussing here really hinges on whether Frege himself made the connection to self-consciousness I'm suggesting. I am grateful to Achille Varzi for pressing me to address this point.

<sup>47</sup>In fact, a similar structure will exist at the level of *reference*, provided one is willing to treat one's conscious perceptual experience as having a location in physical space along with one's body. I am currently perceiving my own body as located within the region of physical space that I'm also perceiving, and to the extent that it makes sense to think of my conscious experience as having a physical location (at the egocentric origin of my perspective, for example), my perceptual experience will also represent a whole of which its own location is a part. Perhaps this will be easier to accept if one accepts the following (still controversial) position: my conscious experience is a part of me, and I'm currently perceiving *myself* as located in physical space. If we treat the True as the actual world (broadly construed to include abstract objects as parts), which I'll defend in Chapter 4, the same structure will exist with respect to the relation between a true thought and its referent (the True), making it even less surprising that this structure is found in indirect senses as well. In any case, nothing that follows hinges on these issues.

<sup>48</sup>I'm tempted to add: why would conscious experience evolve with such a rich structure unless it was useful for the purposes of helping the agent make decisions leading to various actions, actions which will have consequences for the agent's bodily survival, reproduction, etc?

judgment regarding whether the appearances match reality. For the same reason, I need to be capable of *representing* my own beliefs as expressing various propositions in order to be capable of rationally assessing them as correct or incorrect. Without such an ability, reflective judgment would be impossible.<sup>49</sup>

Hence, the thought expressed by my thinking e.g. “I believe that John has brown hair” ought to not only represent my relation to the thought that John has brown hair; it ought to also represent the relation of that latter thought to what it is about. The point of such a meta-level reflection is to compare the latter thought to the world, to rationally assess my own beliefs as correctly or incorrectly representing the world by virtue of expressing such thoughts. *That* is why both customary senses and indirect senses are expressed in indirect contexts.

Turning now to another aspect of Frege’s account, let’s ask the following question: what is it to represent an object (or system of objects) as *structured*? It is to represent aspects of that object (or system of objects) as capable of in-principle variation relative to other aspects held fixed. What I mean by “in-principle variation” is that the fixed structure determines a class of possible objects (or systems of objects), each of which corresponds to a possible variation of the variable aspects. The object needn’t be capable of actually varying in this way.

---

<sup>49</sup>Recall Kripke’s claim that Frege was committed to an agent’s being introspectively acquainted with a sense every time they determine a referent, a sense corresponding to how the referent was determined, and the sense expressed by this acquaintance is the indirect (canonical) sense of the original sense. One can now generalize this position: every time my mind constructs a conscious representation, I have a conscious awareness of that representation as expressing a sense, an awareness that can itself be formed into a *representation* of that sense that will express the canonical sense of that sense. Compare again (my extension of) Kripke’s point about *direct* quotation contexts: I have an awareness of a word-token as being an instance of its corresponding type, an awareness that can be formed into a representation of that type that will express the sense of that word in an unembedded direct quotation context (in which the word-type refers to itself).

I believe that this notion of structure is at the heart of Frege's use of the object/function distinction to elucidate the structure of thoughts, and this is why we should hold onto this distinction. Frege had good reason to treat the sense of a predicate as a function from the sense of a term to the sense of the sentence resulting from their composition, a sense containing both as parts united into a complete thought. The thought contains a fixed aspect (the predicative sense), variation relative to which defines a class of thoughts with a structural identity (e.g. the thoughts expressed by 'John has brown hair', 'Jim has brown hair', 'Fred has brown hair', etc.). Frege's use of the object/function distinction makes this structure explicit.

This point about structure gives us another reason for preferring Frege's account: as I argued above, one can't hold onto both Frege's use of the object/function distinction to elucidate the structure of thoughts and treat indirect senses as structured in such a way that they contain their referents as parts. One *can*, however, treat their *referents* as structured in such a way that they contain their canonical senses as parts, and this is precisely what Frege did.<sup>50</sup>

With these points in mind, let's modify the notation for senses used by Burge and Peacocke by making the object/function distinction explicit: expressions for function-senses will have the form ' $\langle f \rangle ( )$ ', while expressions for object-senses will retain the form ' $\langle o \rangle$ '. So, the thought that John has brown hair can be denoted by the expression ' $\langle \text{has brown hair} \rangle (\langle \text{John} \rangle)$ ' in addition to the expression ' $\langle \text{John has}$

---

<sup>50</sup>Recall that this may have been what forced Burge to treat his first level of canonical names of senses as primitive rather than as composed from his term for the canonical sense function and ordinary terms, leading to an undesirable lack of unity in his formal language, along with a lack of unity between his language and his principles governing canonical senses. Neither Peacocke nor Kripke gave us enough information about the structure of senses to determine their own views on the matter, although Peacocke's choice of notation might suggest a position similar to that of Burge.

brown hair>’, and both expressions express the canonical concept of that thought, grasped as the result of applying a predicative sense to a complete sense in the former case and as unstructured in the latter. Frege’s theory then becomes: for any assertoric sentence ‘ $p$ ’,  $\langle p \rangle = \langle \text{is true} \rangle(\langle \langle p \rangle \rangle)$ . Our extension to the second level of the hierarchy becomes: for any thought  $\langle p \rangle$ ,  $\langle \langle p \rangle \rangle = \langle \langle \text{is true} \rangle \rangle(\langle \langle \langle p \rangle \rangle \rangle)$ .<sup>51</sup>

I trust the reader can iterate this notational procedure in the obvious way.

What about indirect contexts? As I argued above, in such contexts both the canonical concept of a thought and the thought itself are expressed, and expressed in a way that makes their connection salient. This modification of Frege’s view results in the following: ‘Mary believes that John has brown hair’ expresses the thought  $\langle \text{Bel} \rangle(\langle \text{Mary} \rangle, \langle \langle \text{John has brown hair} \rangle \rangle, \langle \text{is true} \rangle(\langle \langle \text{John has brown hair} \rangle \rangle))$ . Recall that  $\langle \text{is true} \rangle(\langle \langle \text{John has brown hair} \rangle \rangle) = \langle \text{John has brown hair} \rangle$ . So, ‘ $S$  believes that  $p$ ’ expresses the thought  $\langle \text{Bel} \rangle(\langle S \rangle, \langle \langle p \rangle \rangle, \langle \text{is true} \rangle(\langle \langle p \rangle \rangle))$ , where  $\langle \text{is true} \rangle(\langle \langle p \rangle \rangle) = \langle p \rangle$ . One can make analogous moves to handle other sorts of indirect contexts (e.g. expressions of thinking, judging, considering, etc.).

Hence, modifying Frege’s view in this way allows both the canonical concept of this thought and the thought itself to be expressed, where the latter is expressed in a way that makes salient its relation to the former. Iterating this procedure will respect Burge’s principles since at each level of embedding, the lowest-level thought will still be expressed. This also supports Kripke’s position that canonical senses of thoughts

---

<sup>51</sup>Although I’ve adopted and slightly modified a popular notation, I trust the reader won’t read the wrong mereological structure into it:  $\langle \langle p \rangle \rangle$  is a part of  $\langle p \rangle$  rather than vice versa.

are immediately revelatory: I can't grasp a canonical sense of a thought without also grasping that thought, and grasping it in a way that makes its being the referent of its canonical sense salient.<sup>52</sup>

The reader may have noticed that throughout the text I've been discussing canonical senses of *thoughts* rather than the more general case of canonical senses of *senses*. In fact, we are already committed to a view of the more general case: the canonical sense  $\langle\langle s \rangle\rangle$  of a sense  $\langle s \rangle$  is the part of the canonical sense  $\langle\langle p \rangle\rangle$  of a thought  $\langle p \rangle$  containing  $\langle s \rangle$  which refers to  $\langle s \rangle$ . Such a part will always exist (and be unique) by our arguments at the end of the third section concerning the structure of indirect senses.<sup>53</sup>

With that said, we can modify Frege's view again in order to give a more general response to Peacocke's worries about inferences than his own: let's introduce an expression 'ref<sub>o</sub>( )' that refers to the reference function from an object-sense to its referent, and an expression 'ref<sub>f</sub>( )' that refers to the reference function from a function-sense to its referent. In English, these expressions correspond to the phrase 'the referent of ( )'. Then, we can generalize Frege's view by insisting that for object-senses  $\langle s \rangle$  other than thoughts,  $\langle s \rangle = \langle \text{ref}_o \rangle(\langle\langle s \rangle\rangle)$ , and analogously for function-senses  $\langle g \rangle( )$ ,  $\langle g \rangle( ) = \langle \text{ref}_f \rangle(\langle\langle g \rangle\rangle)( )$ .<sup>54</sup>

---

<sup>52</sup>Hence I hold the further view that whenever a thought is grasped as including a canonical sense of a sense as a part, the latter sense must be grasped as a part of that thought as well. In a formal system, this could easily be axiomatized.

<sup>53</sup>In the case of a function-sense  $\langle g \rangle( )$ , I'm adopting the convention of denoting its canonical sense with the expression ' $\langle\langle g \rangle\rangle( )$ '. Note that the empty parentheses remain outside the scope of the brackets.

<sup>54</sup>Note that for ease of readability (and to avoid a formal mess), I've adopted the convention of moving the empty parentheses after  $\langle\langle g \rangle\rangle$  outside of the scope of  $\langle \text{ref}_f \rangle( )$ .

Frege himself would be unhappy with my use of identity between functions. This isn't the place



With this modification, a thought such as that expressed by ‘Mary believes that John has brown hair’ can be grasped in such a way that it includes both  $\langle\langle\text{John}\rangle\rangle$  and  $\langle\text{John}\rangle$  as parts, and furthermore in a way that makes their reference-relation salient, since  $\langle\langle\text{John has brown hair}\rangle\rangle = \langle\langle\text{has brown hair}\rangle\rangle(\langle\langle\text{John}\rangle\rangle)$  and (thanks to our modification)  $\langle\text{John has brown hair}\rangle = \langle\text{ref}_f\rangle(\langle\langle\text{has brown hair}\rangle\rangle)(\langle\text{John}\rangle) = \langle\text{ref}_f\rangle(\langle\langle\text{has brown hair}\rangle\rangle)(\langle\text{ref}_o\rangle(\langle\langle\text{John}\rangle\rangle))$ .<sup>55</sup> Thus, we now have an account under which both indirect senses and customary senses can be expressed in indirect contexts, as desired.

The reader may have noticed that although my initial example of a problematic inference included the expression ‘is false’, I’ve so far failed to discuss this expression. On the view I’ve been developing, ‘is false’ ought to refer to a function from a thought to the truth-value of its negation, and it ought to express a sense mapping the canonical sense of a thought to the negation of that thought. Hence, for any thought  $\langle p \rangle$ ,  $\langle p \rangle = \langle \text{is false} \rangle(\langle\langle\neg p\rangle\rangle)$ .<sup>56</sup> This allows us to capture the initial

---

to discuss his infamous concept horse problem, but I trust the reader bothered by such an identity can modify my notation in the appropriate way (e.g. state the condition on function-senses with empty spaces filled in). Note that I have made a slight concession to Frege’s position by specifying two reference-functions rather than one. The reader bothered by this concession is welcome to unite them.

<sup>55</sup>More specifically, this thought can be denoted by the rather complex expression ‘ $\langle\text{Bel}\rangle(\langle\text{Mary}\rangle, \langle\langle\text{has brown hair}\rangle\rangle(\langle\langle\text{John}\rangle\rangle), \langle\text{ref}_f\rangle(\langle\langle\text{has brown hair}\rangle\rangle)(\langle\text{ref}_o\rangle(\langle\langle\text{John}\rangle\rangle)))$ ’, and grasping this thought in the way suggested by this expression immediately makes salient transitions to other thoughts about John (since  $\langle\text{ref}_o\rangle(\langle\langle\text{John}\rangle\rangle) = \langle\text{John}\rangle$  by our modification), as desired. In line with my example above, it may make salient a transition to the thought that John is a good friend of mine, a thought which we could formalize in such a way that it includes  $\langle\text{ref}_o\rangle(\langle\langle\text{John}\rangle\rangle)$  as a part, thus allowing us to easily demonstrate that my example was a valid inference. For the same reason, grasping the original thought in this way also immediately makes salient transitions to other thoughts about having brown hair.

As we discussed in the previous chapter, thoughts can be grasped in ways that don’t make such transitions salient as well. Properly capturing a particular inference forces the thoughts involved to be represented as decomposed into senses in particular ways.

<sup>56</sup>Recall from the last chapter that for any thought  $\langle p \rangle$ ,  $\langle p \rangle = \langle \neg\neg p \rangle$ .

problematic inference I discussed in the Introduction without the need for additional premises, and in a way that neither the system of Peacocke nor of Burge can express, since on their accounts thoughts do not include parts referring to their negations or to themselves.<sup>57</sup>

For us, the thought expressed by ‘But the planetary orbits aren’t circles’ can be denoted by the expression ‘<is false>(<<the planetary orbits are circles>>)’ and since <<the planetary orbits are circles>> *refers* to the negation of this thought, that negation can be existentially quantified over in our formal system in precisely the manner needed to demonstrate that the inference is valid.<sup>58</sup>

We have thus seen the many benefits, both formal and philosophical, of adopting Frege’s own account of the structure of senses relevant to his hierarchy. We have also seen how to generalize his account to widen the range of transitions in thought it can be brought to bear on. We have also discussed the elegant connections between Frege’s account and the structure of conscious perceptual experience, as well as the role of

---

<sup>57</sup>Why should a thought include a part referring to its negation? I’ll let Frege answer this question: “We grasp the content of a truth before we recognize it as true, but we grasp not only this; we grasp the opposite as well. In asking a question we are poised between opposite sentences. Although it is usually only one side that is expressed when we speak, the other is still always implied; for the sense of the question remains the same if we add ‘or not?’.” (between 1879 and 1891, PW pg. 7-8) Also: “When it is a question of whether some thought is true, we are poised between opposite thoughts, and the same act which recognizes one of them as true recognizes the other as false.” (1897, PW pg. 149)

<sup>58</sup>Recall that the problematic inference was the following: “Frege believed that Copernicus believed that the planetary orbits are circles. But the planetary orbits aren’t circles. Hence, some proposition that Frege believed that Copernicus believed is false.” I leave expressing this inference in our sketch of a formal system in a way that makes it trivially valid as an exercise for the reader.

Note that extending this account to a formal system would involve axiomatizing various principles that ought to govern the hierarchy. For example, I shouldn’t be able to apply existential introduction to a sense without also applying it to its canonical sense (and vice versa), assuming that its canonical sense was also expressed. For example, in inferring “Mary believes that someone has brown hair” from “Mary believes that John has brown hair”, I have abstracted away from *both* the customary sense *and* the indirect sense of ‘John’; I couldn’t have abstracted away from one without the other. One may want to design an alternative notation to reflect this.

meta-representation in reflective judgment. I think it is clear that these considerations overwhelmingly support choosing this account of the hierarchy of senses over those of our contemporaries.

# Chapter 3

## Demonstratives and Self-Reference

### 3.1 Introduction

The content of an assertion will often depend upon the context in which it is made.<sup>1</sup>

When I assert “It is raining here now”, whether my assertion is true or false will depend upon whether it is in fact raining at the place and time of my utterance.<sup>2</sup> If

I assert “It is raining here now” at a different location or at a different time, whether

---

<sup>1</sup>As in previous chapters, I am here using *content* as an intuitive notion from which more technical notions can be isolated. Once again, a *thought* is an abstraction from the full content of an assertion to the sort of content relevant to determining whether the assertion is true or false, and thoughts will be our primary concern in this chapter.

<sup>2</sup>There are unusual cases in which the place or time relevant to determining the truth-value of the assertion might be different than the spatiotemporal location of the utterance. Perhaps the speaker knows that the utterance will be heard one day later than it is made, and also believes that it will rain the next day. In such a case, they may still utter “It is raining here now”, so long as the hearer can be expected to recognize that the speaker was referring to the time the utterance would be heard rather than the time at which it was made. One might think that in such a case, one is using the sentence in a nonstandard way in order to express a content analogous to that ordinarily expressed by “It will be raining here tomorrow”. However, the speaker might not even know when the utterance will be heard; they may simply know that whenever the utterance will be heard, it will be raining at the location at which the utterance was made (e.g. perhaps the advent of rain would *cause* the hearer to hear the utterance previously made; less exotic examples involve written text: “it is raining here now” could be written on the wall of a shelter from the rain). See (Gaifman 2008) for a treatment of these and related issues.

my assertion is true or false will depend upon whether it is raining at *that* location or *that* time. This suggests that even if we wish to isolate the sort of content relevant to determining truth-conditions, we will often need to take *context* into account.

Left at such a high level of abstraction, this may seem uncontroversial. Trouble arises when we attempt to give a systematic account of how contextual factors contribute to the truth-conditions of assertions. A natural first step is to suggest that when making a normal assertion of “It is raining here now”, the expressions ‘here’ and ‘now’ denote the place and time of utterance, respectively (but see footnote 2 regarding abnormal assertions). Furthermore, as we shall see, it seems natural to insist that ‘here’ and ‘now’ express senses that present their referents in particular ways. That is to say, the expressions ‘here’ and ‘now’ contribute more than merely their referents to the truth-conditions of assertions.

For example, ‘now’ when uttered by me and ‘noon’ when uttered by me may refer to the same time,<sup>3</sup> and yet if I have lost track of the time, I may believe that it is raining here now without believing that it is raining here at noon. Furthermore, given that the two sentences ‘It is raining here now’ and ‘It is raining here at noon’ can differ in truth-value even while the context remains fixed (e.g. if it is not in fact

---

<sup>3</sup>One might think that I can’t utter ‘now’ and ‘noon’ at the same time, since utterances take time, so it might be better to say that if I *had* uttered ‘noon’ instead of ‘now’, it *would have* referred to the same time as ‘now’ actually refers to. But note that ‘now’ doesn’t really refer to a precise point in time (where we think of a timeline as having the topological structure of the real line); it seems to refer to a short duration of time with vague boundaries (or to vaguely refer, if one prefers), where the length of this duration is also sensitive to context. Treating ‘now’ as if it refers to a precise point in time is a useful idealization. Similar remarks apply to ‘here’.

I make no attempt to present a systematic theory of vagueness in this chapter, although I will address various issues as they arise. For an elegant treatment of vagueness I refer the reader to (Gaifman 2010). Note that I speak of entities such as times of utterance as if they themselves have vague boundaries. If the reader is unhappy with this, I trust that they can translate my claims in accordance with their preferred treatment of vagueness.

noon), they must have distinct truth-conditions, from which it follows that ‘now’ and ‘noon’ make distinct contributions to the truth-conditions of assertions. So, at least, goes the natural Fregean story.<sup>4</sup>

So, ‘now’ seems to contribute more than its referent to the truth-conditions of an assertion. Roughly, it expresses a sense that presents its referent *as* the time of the utterance rather than as noon, as the time I am writing this chapter, etc.<sup>5</sup> Similarly, ‘today’ expresses a sense that presents its referent as the day containing the time of the utterance (i.e. the day containing now), while ‘yesterday’ expresses a sense that presents its referent as the day prior to today. Similar remarks seem to apply to other indexicals.

And yet, now one might worry that our account of Fregean thoughts has grown too fine-grained. There is a natural sense in which “Today it is raining here” says the same thing as “Yesterday it was raining here” uttered the next day (assuming the location of utterance remains constant).

---

<sup>4</sup>Here is a potential problem with this sort of argument that has been widely discussed: Someone might not realize that ‘doctor’ and ‘physician’ express the same sense, and hence it seems that their two utterances ‘I believe that all doctors are doctors’ and ‘I believe that all doctors are physicians’ could differ in truth-value, even though ‘doctors’ and ‘physicians’ make the same contribution to truth-conditions by assumption. Extending this example would seem to lead to the unfortunate conclusion that no two distinct words can ever express the same sense, and perhaps even that no two distinct tokens of the same word can ever express the same sense either.

In response, note that when making this sort of argument we must assume that the agent is an idealized linguistically competent speaker, otherwise they may think that a sentence expresses different truth-conditions than it actually does. In such a case, the agent might not realize that they in fact bear a particular relation (such as a belief relation) to the thought in question, since they don’t recognize which thought is being expressed by a particular sentence.

In effect, we are abstracting from the ordinary notion of belief to a relation that holds between an agent and a thought (rather than a sort of content more fine-grained than a thought). So, the agent in the purported counterexample in fact believes that all doctors are physicians; they merely fail to recognize that the corresponding sentence expresses the same thought as the sentence ‘All doctors are doctors’, and hence may be unwilling to assert the former sentence.

<sup>5</sup>That isn’t to say that the sense of ‘now’ is equivalent to that of a description such as ‘the time of this very utterance’. We’ll discuss these issues in further detail later.

And yet how could these two utterances express the same thought, when the thought expressed by the former contains the sense of ‘today’ while the thought expressed by the latter contains the obviously distinct sense of ‘yesterday’? In fact, how could anyone ever grasp the thought expressed by ‘Today it is raining here’ on any day but the day on which it was originally expressed?<sup>6</sup> Given our discussion of Frege puzzles above, the thought expressed by e.g. ‘On November 30th, 2017, it was raining here’ is a *distinct* thought.

Once it is no longer the day in which ‘Today it was raining here’ was originally uttered, it seems that a thinker can no longer grasp a sense referring to that day in the way the sense of ‘today’ referred to it (as the day containing the time of utterance), and hence the thought originally expressed by ‘Today it is raining here’ can never be grasped again. Any other expression referring to that day will express a sense representing it in a different way.

The reader may not be bothered by such failures of thought preservation. If I read ‘Today it will rain’ in yesterday’s local newspaper, I may not be able to grasp the thought originally expressed by that sentence, but I can at least refer to it and grasp a related thought, e.g. the thought that yesterday it rained.<sup>7</sup> Perhaps this is all that is required for understanding the sentence. But let’s consider additional examples. When I assert “John is wearing a brown hat”, the truth-value of my assertion depends upon the time of utterance (sometimes John wears hats; sometimes he doesn’t). It

---

<sup>6</sup>I owe this point to (Kripke 2008), which we’ll discuss below.

<sup>7</sup>This is Kripke’s (2008) response on behalf of Frege. Haim Gaifman has pointed out to me that at one point Bertrand Russell made a similar move, namely: we can describe a proposition that we are not acquainted with. This is discussed in his “Knowledge by Acquaintance and Knowledge by Description” (1910-11).

seems natural to conclude that the assertion included an implicit reference to the time of utterance. Hence, the thought expressed included a sense as a component that refers to that time.

But then how could I possibly express the same thought at a distinct time? My second utterance would express a thought including a sense as a component that refers to the time of *that* utterance. At the very least, I can no longer refer to the previous time of utterance in the same way, since it is no longer the time of my current utterance. This issue becomes pervasive once one notices that the vast majority of our thoughts and utterances are tensed. Surely I can say the same thing at more than one time!<sup>8</sup> How can a Fregean theory account for this?

Now, there are certainly cases in which it can be cognitively significant that John is still wearing a brown hat after a few seconds have passed. In such cases, it does seem that a pair of utterances of “John is wearing a brown hat” several seconds apart ought to express distinct thoughts. But there are other cases in which it seems the two utterances ought to express the same thought.

For example, consider a conversation between two people: “John is wearing a brown hat.” “I agree. John is wearing a brown hat.” Here the second speaker intends to repeat the claim made by the first. In order for this to be possible on the view we’re considering, this must be a context in which the time of utterance is the same for both utterances, and hence its duration is longer than a few seconds.

---

<sup>8</sup>But note that the “time of utterance” has vague boundaries and its duration is sensitive to context, as we discussed in footnote 3. The point remains, however, that once it is no longer the time of the original utterance, it seems that the original thought can’t be expressed. Indeed, it can’t even be grasped.



But suppose that I say “John is wearing a brown hat” and few seconds later, John removes his hat. Someone else says to me, “What you said was false! John isn’t wearing a brown hat!” Clearly something has gone wrong; the second person has ascribed the wrong truth-conditions to my original utterance. I wasn’t committed to John’s wearing a brown hat for an extended length of time; I was only committed to John’s wearing a brown hat *at the time of my utterance*, and hence in this context the time of my utterance must be less than a few seconds. Otherwise, I would have said something like “John will be wearing a brown hat for the next few seconds.” Here the duration of the time of utterance would need to be determined by features of the context *after* the utterance is actually made.<sup>9</sup>

As another example, consider the case of belief ascriptions, which we discussed in detail in the last chapter. If Mary says “John is wearing a brown hat”, it seems natural to report her belief via an utterance like “Mary believes that John is wearing a brown hat”. On a Fregean view, this belief report expresses a thought that represents Mary as bearing the belief relation to the thought that John is wearing a brown hat.<sup>10</sup> But which thought is that? Given that Mary made her utterance at a different time than mine, which thought am I reporting her as believing: the thought that Mary expressed at the time she made her utterance, or the thought that I would express if I

---

<sup>9</sup>This clearly raises various issues regarding the way the context determines the relevant duration of the time of utterance, but discussing these issues further would go beyond the intended scope of this chapter.

<sup>10</sup>I defended the view in the last chapter that belief is actually a three-place relation (in order to account for the expression of both indirect and customary senses in indirect contexts), but I’m suppressing that complication here. Incidentally, when combined with the position on truth-values defended in the next chapter I am committed to belief being a three-place relation between subject, thought, and world.

had instead uttered “John is wearing a brown hat”? These thoughts refer to different times.

But note once again that on this view I can no longer grasp such a thought once the time it was originally uttered has passed. This implies that I couldn’t be reporting Mary as believing the thought that she originally expressed, since (as we saw in the last chapter) when making such a belief report I am expressing the canonical sense of that thought, and I can’t grasp this canonical sense without also grasping the thought it refers to.

Hence I am not reporting the thought that Mary originally expressed, but rather the thought I believe she currently bears the belief relation to, which is the same thought I would have expressed if I had uttered “John is wearing a brown hat” at the time I instead made the belief report. This implies that Mary is constantly updating her beliefs, bearing the belief relation to a new class of thoughts at every moment. Again, this isn’t necessarily an objection to the position we’re considering, but it is worth noting. It seems to go against intuitions about the preservation of beliefs over time, even when we’re abstracting from the content of beliefs to the sort of content relevant to determining truth-conditions.

But there is a more serious objection. Suppose later that day I utter “Mary believed that John was wearing a brown hat.” Obviously, Mary may no longer believe this, and I’m clearly not ascribing the current belief that John is wearing a brown hat to her. I’m not ascribing the belief that John *was* wearing a brown hat to her either, since she may have given up that belief in light of new information. Instead, I’m asserting that Mary *did* believe that John was wearing a brown hat.

But on the view we're considering, I can't actually grasp the thought I'm attempting to assert. This is because that thought includes a canonical sense referring to the thought Mary originally believed, and yet I can no longer grasp that thought, and hence (as we saw in the last chapter) I can't grasp (or even express) its canonical sense either.

Again, the problem is that indirect senses are *transparent* or *revelatory* of their referents, and hence one cannot grasp an indirect sense without also grasping the sense it refers to. This implies that I can't ascribe a belief to someone without grasping the thought I'm ascribing. But this simply cannot be done on the view we're considering in a wide variety of cases. In order to make sense of such belief reports, I had better be able to grasp tensed thoughts at later times, and yet that simply cannot be done on this view of the structure of thought.

Such considerations can pull one in the direction of a treatment of thoughts under which all that matters is preservation of the appropriate *referent*, so that "Today it is raining" and "Yesterday it was raining" can express the same thought so long as 'Today' and 'Yesterday' refer to the same day in their respective utterances. But then we seem to have no answer to the Frege puzzles that pushed us in the direction of a Fregean treatment of truth-conditions in the first place. The Fregean account of truth-conditional content we've been considering seems too fine-grained, while this non-Fregean account seems too coarse-grained.

There is another set of famous objections to a Fregean account of indexicals and demonstratives, stemming from the work of (Kripke 1980) and (Kaplan 1989). While the status of these objections is controversial (and there are various Fregean re-

sponses), this won't matter for our purposes. I will sketch an objection here and later argue that Frege himself had a natural response.

Let's first discuss demonstratives. When used in a particular context, a demonstrative such as 'that' often has to be supplemented by an action (e.g. an act of pointing) in order to determine a unique referent. When I utter "That is my cell phone" while pointing at an object, the truth-value of my assertion depends upon which object I pointed at. This suggests that the act of pointing contributes to the expression of the thought. One might say that the demonstrative 'that', in combination with the act of pointing, expresses a sense that refers to my cell phone (we'll see below that this was Frege's actual view). Importantly, the very same utterance of 'that' accompanied by the very same act of pointing might have referred to a different object (e.g. the items on my desk might have been arranged differently).<sup>11</sup>

Here is an argument in the spirit of Kaplan (1989, pg. 516ff): Consider my assertion "That is *my* cell phone, but that might have belonged to someone else", where both uses of 'that' are meant to be accompanied by the same act of pointing (to my actual cell phone, let's assume). Both uses of 'that' are also meant to express the same sense (or close enough to the same sense, if identity is impossible). When I make this assertion, I am *not* claiming that someone's else's cell phone might have been the object pointed at by my action; I am instead making the claim that my cell phone might have belonged to someone else.<sup>12</sup>

---

<sup>11</sup>Here it is being assumed that such acts are not individuated in terms of the objects being pointed at.

<sup>12</sup>There *may* be unusual contexts in which I could make the former claim (or the claim that someone else's object might have been at that location) with that utterance (perhaps in a discussion about how pointing works), but let's suppose this isn't one of them. The Fregean is still obligated

Here is one way to evaluate this case: when evaluating the truth-value of this utterance, it seems I will need to consider other possible states of the world, but while doing so the object relevant to determining the truth-value remains fixed as my actual cell phone.<sup>13</sup> The fact that I might have pointed at a different object is irrelevant to determining whether my cell phone might have belonged to someone else. Hence, when considering other possible states of the world, I needn't consider other objects that I may have pointed at. Instead, I should continue considering the object that I in fact pointed at in the actual state of the world.

This suggests that such a use of 'that' (in conjunction with an act of pointing) functions as a *rigid designator*, designating the same object even when evaluated relative to a different possible state of the world.<sup>14</sup> But then the sense expressed by (this use of) 'that' and my act of pointing cannot be a sense that represents its object as whatever object I am pointing at (if any) in other possible states of the world, since this gives the wrong truth-conditions of the assertion.<sup>15</sup> But what else

---

to give the truth-conditions of the utterance I actually made.

<sup>13</sup>Note that I am *not* committed to the position that a possible worlds semantics is the correct treatment of modality. Here I merely discuss the objection in its most natural form.

<sup>14</sup>Compare a standard utterance of "It is not raining here, but it might have been." The location relevant to determining the truth-value remains fixed as the actual location of the utterance even while I'm considering other possible states of the world. The fact that it might have been raining at a different location where I might have uttered that sentence is irrelevant to determining its actual truth-value. This suggests that words like 'here' are also rigid designators.

<sup>15</sup>For example, this would seem to be the sense expressed on a view where the sense is given by the following Kaplan-style rule for determining the reference (examples of which are given in (Kripke 2008)): when someone uses the word 'that' demonstratively, they refer to the demonstrandum of the relevant demonstration.

Could one insist that a sense maintains fixed reference to its actual referent even when evaluated relative to other possible states of the world? Frege did in fact hold that senses have the same reference even when the context varies (change in reference entails change in sense), and extending this view to the modal case may seem natural. However, consider the following: "The first Postmaster General of the United States might not have been the inventor of bifocals." That utterance is true, and yet it seems false that Benjamin Franklin might not have been Benjamin Franklin. As

could the sense possibly be?

In this chapter, I'll attempt to sketch a Fregean treatment of demonstratives and indexicals that addresses these arguments. I'll begin by looking carefully at what Gottlob Frege actually said about such cases, followed by discussing how his account should be modified. As in the last chapter, I'll bolster the case for my account by considering relevant aspects of conscious experience, perception, and reflective judgment.

## 3.2 Frege on Demonstratives and Indexicals

In the last chapter we discussed Saul Kripke's paper on Frege: "Frege's Theory of Sense and Reference: Some Exegetical Notes" (2008). Besides discussing Frege's account of the hierarchy of senses, Kripke also gives an interpretation of Frege's brief remarks on demonstratives and indexicals, primarily in "Thoughts" (Frege 1918-19a). In this section I will defend various aspects of Kripke's interpretation as correct and then develop it further in light of Frege's commitment to the Multiple Decompositions Thesis. We'll consider objections in the next two sections.

First, some context. Kripke is partially responding to (Burge 1979), in which Burge accused Kripke of mistakenly equating sense with meaning in (Kripke 1980). Burge supported his alternative position by arguing that for Frege the sense associated

---

the reader probably knows, this style of argument is due to (Kripke 1980).

Perhaps the defender of such a strong fixed-reference view could insist that 'the inventor of bifocals' is not a referring expression in that utterance, but is instead an *inseparable* component of the predicate '( ) might not have been the inventor of bifocals'. This seems very strained, and it leaves one unable to explain why the expression 'the inventor of bifocals' can be replaced by other expressions in a way that systematically changes the sense expressed by the predicate.

with an indexical like ‘today’ must shift from context to context, since its referent shifts. However, the *meaning* of an indexical like ‘today’ remains constant throughout contexts, and hence senses aren’t meanings. (Burge 1979, pg. 214-15 in the 2005 reprint) Kripke responds by defending an interpretation of Frege where the sense of an indexical *doesn’t* shift from context to context.

Throughout this dissertation, I’ve made it clear that senses should not be identified with linguistic meanings, for the thought expressed by an assertion (for example) is an *abstraction* from the full content expressed by the assertion to the sort of content relevant to determining whether the assertion is true or false. Frege’s primary interest was the sort of content relevant to *logic*, not just any content expressed by a sentence in natural language.<sup>16</sup> Hence, I agree with Burge that senses shouldn’t be identified with full linguistic meanings, although they can be thought of as one aspect of meaning. What about Burge’s *argument*? Let’s consider Kripke’s response.

Burge and Kripke both base their accounts primarily on the following quotation:

(1) If a time-indication is conveyed by the present tense one must know when the sentence was uttered in order to grasp the thought correctly. Therefore the time of utterance is part of the expression of the thought. If someone wants to say today what he expressed yesterday using the word ‘today’, he will replace this word with ‘yesterday’. Although the thought is the same its verbal expression must be different in order that the change of sense which would otherwise be effected by the differing

---

<sup>16</sup>Here’s the most relevant quote again: “So one has to separate off from the content of a sentence the part that alone can be accepted as true or rejected as false. I call this part the thought expressed by the sentence.” (Frege 1906c, PW pg. 197-8)

Kripke complains that if Frege didn’t identify sense with meaning, then why didn’t he ever use a technical term for ordinary meaning? (Kripke 2008, pg. 273 in the 2011 reprint) For one, I don’t think Frege had any need for such a term, since ordinary meaning was not his primary interest. In any case, a plausible candidate term is ‘content’ (‘Inhalt’).

times of utterance may be cancelled out.<sup>17</sup> The case is the same with words like ‘here’ and ‘there’. In all such cases the mere wording, as it can be preserved in writing, is not the complete expression of the thought; the knowledge of certain conditions accompanying the utterance, which are used as means of expressing the thought, is needed for us to grasp the thought correctly. Pointing the finger, hand gestures, glances may belong here too. The same utterance containing the word ‘I’ in the mouths of different men will express different thoughts of which some may be true, others false. (Frege 1918-19a, pg. 64/358)

While Burge interprets this passage as implying that e.g. the sense of ‘today’ will vary from day to day (Burge 1979, pg. 214-5), Kripke points out the following: Frege says the time of utterance is part of the expression of the thought. This suggests the time of utterance itself expresses a sense that is part of the thought expressed; the utterance by itself fails to express a complete sense. What does the expressed sense refer to? The time of utterance. Hence, the time of utterance is being used autonomously, as an expression that refers to itself. (Kripke 2008, pg. 274)<sup>18</sup> In further support of this, note that at the end of the passage Frege says that certain conditions accompanying the utterance are used as means of expressing the thought (not our *knowledge* of certain conditions, but the conditions *themselves*).

Later in the same paper, Frege says “The words ‘This tree is covered with green leaves’ are not sufficient by themselves to constitute the expression of the thought, for the time of utterance is involved as well. Without the time-specification thus given we have not a complete thought, i.e. we have no thought at all.” (1918-19a,

---

<sup>17</sup>Here we see that Frege wanted various tensed thoughts to be preservable over time, and we can also see his suggestion for how to do so (which should come as no surprise to readers of the first two chapters of this dissertation), which we’ll discuss further below.

<sup>18</sup>Note that Kripke’s interpretative point does not hold in general: the letters that make up a word may be part of the expression of a thought, but they do not thereby express senses that are parts of that thought. I will argue in the next section that Kripke’s interpretative point is correct when applied to this particular case.



pg. 76/370) This again suggests that the time of utterance itself expresses a sense which is part of the complete thought expressed, and by virtue of doing so it specifies (refers to) a time (itself). When discussing the first person pronoun ‘I’, Frege again makes such a remark: “...by doing this he makes the conditions accompanying his utterance serve towards the expression of a thought.” (ibid pg. 66/360)

So, if I utter “It is raining in Brooklyn”, the thought expressed by the utterance is incomplete, and must be completed by a sense referring to the time of utterance.<sup>19</sup> This sense is provided by the time of utterance itself, where that time is used as an expression which refers to itself. According to Kripke, the reason a speaker can refer to the time of utterance in this way is because they are *acquainted* with it, presumably in virtue of its currently *being* that time. (2008, pg. 275) I am acquainted with the time it actually is, and this allows me to refer to it in a special way. Similarly, I am acquainted with the place I’m actually located at, and I am acquainted with myself.<sup>20</sup>

Kripke generalizes this interpretation to other indexicals and demonstratives. For example, the sense expressed by ‘here’ (used non-demonstratively) must be saturated by the sense expressed by the location of utterance used as a name of itself, the sense expressed by ‘I’ must be saturated by the sense expressed by the speaker/thinker used as a name of herself, etc. However, the sense expressed by a demonstrative such

---

<sup>19</sup>As Kripke points out (2008, pg. 275 fn. 63), the *complete* expression would have the form ‘F(t)’, where ‘F( )’ is the unsaturated utterance and ‘t’ is the completing time of utterance, treated as an expression. ‘F( )’ itself refers to a function from times to truth-values.

<sup>20</sup>Kripke is using “acquaintance” in roughly the sense of Russell. I have a special sort of epistemic access to the time it actually is. Compare the access I have to an utterance by virtue of being the one who made it: this epistemic access allows me to represent that utterance as a token instance of the relevant utterance-type, as is needed in Frege’s account of direct quotation (a context in which the words refer to themselves, each expressing a sense of autonomous designation).

I don’t think that acquaintance is necessary for treating an object as a name of itself (any prior representation of the object will do), but I mention Kripke’s position here for completeness.

as ‘that’ won’t be saturated by the sense expressed by the object demonstrated used autonomously, but rather by the sense expressed by the act of demonstration. This does seem to have been Frege’s view (see the end of the quotation above), although it *prima facie* gets the truth-conditions of various utterances wrong, as I pointed out in the Introduction. There I discussed an argument that ‘that’ is often used as a rigid designator, and does not co-refer with the description ‘the object of this demonstration’ when evaluated with respect to various possible states of the world. That might not have been the object of this demonstration, but it seems absurd to say that might not have been that (compare footnote 15). We will discuss whether an account of senses can be given that avoids this objection in a later section.

Recall Frege’s treatment of direct and indirect quotation. In an indirect context, words refer to the senses they would express outside of that context. In a direct quotation context, words refer to themselves. Furthermore, as I argued in the last chapter, every thought contains a sense as a part (the *canonical sense* of that thought) that refers to the whole thought. Frege was no stranger to self-reference.

So, when I utter “Today it is raining in Brooklyn”, the sense expressed by the utterance must be completed by a sense referring to the time of utterance. In particular, the sense of ‘Today’ is incomplete, and refers to a function from a time to the day containing it. (2008, pg. 276)<sup>21</sup> Thus Kripke can argue, against Burge, that

---

<sup>21</sup>In Frege’s earlier work he made this point more explicitly: “Words like ‘here’ and ‘now’ only acquire their full sense through the circumstances in which they are used. If someone says ‘it is raining’ the time and place of utterance has to be supplied...So the explanation for all these apparent exceptions is that the same sentence does not always express the same thought, because the words need to be supplemented in order to get a complete sense, and how this is done can vary according to the circumstances.” (1897, PW pg. 135)

However, this quotation should be taken with a grain of salt; Frege doesn’t explicitly state here that the circumstances themselves are being used to express thoughts, and hence could have easily

indexicals and demonstratives express the same sense in all contexts. What shifts with context are the senses expressed by contextual factors themselves treated as referring expressions.

What about Frege's remarks about expressing the same thought at two different times by e.g. exchanging 'yesterday' for 'today'? Kripke considers this bizarre, given that the two expressions seem to clearly present the same day in two different ways, and thus express distinct senses. Kripke goes on to conclude that the thought expressed by e.g. "It is raining today" can never be grasped again once the day has passed, and to say otherwise would be to reject Frege's basic principles. (Kripke 2008, pg. 277-8) I, of course, disagree, as does Gareth Evans (1981).

According to Evans, changing epistemic circumstances "force us to change in order to keep hold of a constant reference and a constant thought—we must run to keep still." (Evans 1981, pg. 308 in the 1985 reprint) An utterance like "Yesterday it was raining" can manifest the persistence of a belief that was expressed yesterday with the utterance "Today it is raining." (ibid pg. 309) Furthermore, we shouldn't prescribe a belief of the form "It is now  $\psi$ " to someone if they don't have the propensity as time goes on to form beliefs of the form "It was just  $\psi$  a moment ago", "It was  $\psi$  earlier this morning", etc.<sup>22</sup> A capacity to keep track of time is a precondition of temporal thought. (ibid pg. 309)

---

changed his mind in the following twenty years before "Thoughts" (1918-19a) was published. Still, it seems clear that Kripke's account is the natural way to make Frege's updated view more explicit: e.g. 'Today' refers to a function from a time to the day containing it, and expresses an incomplete sense that must be completed by a sense referring to the time of utterance.

<sup>22</sup>Evans's position might be better served by insisting that the subject is merely *preserving* their belief rather than forming *new* beliefs.

The point being, in such a context these different expressions are expressions of the same persistent belief state. Intuitively, it is the *same belief* that is being preserved over time that must be expressed in distinct ways due to the passage of time. Hence, it is plausible that Frege wanted this belief to be a relation to the same thought over time. And as I pointed out above, it is clear from passage (1) how Frege wanted to accomplish this: by insisting that there are multiple decompositions of the same temporal thought (e.g. the thought expressed by “Today it is raining” uttered in a particular context), some of which include e.g. the sense of ‘yesterday’ as a component while others include the sense of ‘today’ as a component.

Given Frege’s more general account of multiple decompositions of thought, such a move may seem like a natural way to account for the intuition that one can express the same tensed thought more than once. But it isn’t so clear how the details of such an account should be spelled out. Take the thought expressed by “Today it is raining here”, uttered at a particular place and time. According to the interpretation we’ve adopted, ‘here’ expresses a function-sense from the sense expressed by the location of utterance to a sense  $s_1$  denoting that location, while ‘today’ expresses a function-sense from the sense expressed by the time of utterance to a sense  $s_2$  denoting the day containing that time. Furthermore, ‘( ) it is raining ( )’ expresses a function-sense mapping the senses  $s_1$  and  $s_2$  to the complete thought, and it denotes a function from places and times that maps them to the True if it is in fact raining at that place and that time and the False otherwise.<sup>23</sup>

---

<sup>23</sup>Actually, it may denote a function from the place referred to by ‘here’, the duration of time referred to by ‘today’, and the time of utterance to a truth-value, since if I had said “Today it is raining here” when it had been raining earlier in the day but had stopped recently, it seems that I

But what about “Yesterday it was raining here”, uttered the next day at the same place? ‘Yesterday’ expresses a function-sense from the sense expressed by the new time of utterance to a sense  $s_3$  denoting the day *prior to* the day containing that time. But the combination of ‘here’ and the location of utterance expresses the same sense as before by assumption. Furthermore, any difference in sense between ‘( ) it is raining ( )’ uttered yesterday and ‘( ) it was raining ( )’ uttered today does not affect the part of the thought denoting a duration of time.<sup>24</sup> So in order for the second utterance to express the same thought as the first, it seems that  $s_2$  will have to be identical to  $s_3$ . But how could that be, given that they seem to present the same day in two different ways: as the day containing now in the former case and as the day prior to the day containing now in the latter case?

Actually, it is the sense of ‘today’ and the sense of ‘yesterday’ that differ, and hence we can still capture the intuition that the same day is being presented in two different ways while insisting that the *combination* of the sense of ‘today’ and the sense expressed by the time of the first utterance is the same as the *combination* of the sense of ‘yesterday’ and the sense expressed by the time of the second utterance. So perhaps we can identify  $s_2$  and  $s_3$  after all.

However, since it was clearly Frege’s view that this same thought could be expressed in two different ways, it would have said something false. I should have used the past tense and said “Today it was raining here” instead.

---

<sup>24</sup>Doesn’t the change in tense change the sense expressed? Yes, but it doesn’t change the part of the thought that refers to a duration of time, and hence we can ignore this shift here. The shift in contribution to truth-conditions made by this shift in tense is the following: in the former case it must be raining at the time of utterance for the utterance to be true, while in the latter case it must have rained at a time prior to the time of utterance for the utterance to be true (where which prior times are relevant is further constrained by other expressions such as ‘yesterday’).

pressed at many different times (presumably in a different way each time),  $s_2$  is going to have to include parts referring to *arbitrary* times of utterance, and it is going to have to be decomposable into pairs of parts in arbitrarily many ways, with one part referring to a time and the other referring to a function from that time to the relevant day. Otherwise, this sense wouldn't be expressible at various other times; its being so expressible implies that it contains components referring to *every* distinct time at which it could be expressed. That is to say, this sense is going to have to have the structure of a *timeline*, with any point along the timeline denoting a particular time relative to the rest of the timeline denoting the relevant function.

Compare again Frege's use of a line divided by a point to illustrate the function/object distinction, which I argued in the previous two chapters reveals much about Frege's conception of the structure of thought. Here, the point corresponds to a sense denoting a time while the rest of the line corresponds to a function-sense mapping that point to the complete line, which in this case is a complete sense denoting the relevant day.

One might think of the sense as a timeline with a particular interval emphasized (the interval corresponding to the day denoted by the timeline). Any point on the line corresponds to the sense that would be expressed by that time of utterance treated as a name of itself, while the rest of the line (with that point removed) corresponds to a function-sense mapping that point to the complete line, a line which always denotes

the same day (corresponding to the emphasized interval).<sup>25</sup> <sup>26</sup> For similar reasons, the sense expressed by e.g. the combination of the location of utterance and ‘here’ is going to have to have the structure of a *space*. Can such a view really be defended?<sup>27</sup> I’ll attempt to do so in a later section, but let’s first turn to various objections to the aspects of Kripke’s interpretation we’ve been relying upon. The next section isn’t essential to my overall position, but I consider the objections worth discussing.

### 3.3 Objections to Kripke’s Interpretation

Kripke’s interpretation has been criticized by Wolfgang Künne (2010) and Tyler Burge (2012). Künne agrees with Kripke (and has argued independently) that on Frege’s view persons, times, etc. can be part of the expression of a thought. However, Künne doesn’t agree that e.g. times of utterance must express independent

---

<sup>25</sup>This analogy isn’t perfect since e.g. *any* point in the interval would have to be mapped by the *same* function-sense expressed by ‘today’ to the *same* complete sense. It might be better to think of the function-sense expressed by ‘today’ itself as a timeline with an interval emphasized, where the timeline must be “anchored” (saturated) by a sense denoting a time in order to refer to a *specific* day (the day containing *that* time) as opposed to an arbitrary day. Note that senses denoting *any* time within that day are part of the complete sense, and *any* of them could saturate the sense of ‘today’ in a way that results in that same complete sense. That is, here we must say that a change in sense doesn’t necessarily need to be “cancelled out”.

Furthermore, recall that treating a time of utterance as a point on a line is an idealization. In reality, a time of utterance has a duration with vague boundaries whose length is determined by the context. Hence, it is actually a sense denoting such a *duration* that will be mapped by the sense of ‘today’ to the sense denoting the day in question.

<sup>26</sup>To express this sense on the day it denotes, one would use an expression like ‘today’, but to express this sense on a different day, one must use a different expression, since the sense of ‘today’ is a function-sense from a sense denoting a time to a sense denoting the day containing that time, and hence to express this same sense on a different day one needs to use an expression that expresses a function-sense from a sense denoting a time to a sense denoting a day other than the day containing that time.

<sup>27</sup>I doubt that Frege ever considered this in detail, but it does seem to be a consequence of his views on thought-preservation and the structure of thought. That is to say, he seems to have been committed to this position whether or not he ever actually considered it.

senses and have referents, for times are not the sorts of things that are *understood*. (Künne 2010, pg. 541-2)<sup>28</sup> But this begs the question against Kripke. Understanding an expression consists partially in knowing how it is used and how it contributes to the truth-conditions of complete sentences.<sup>29</sup> If a time of utterance is being used as an expression for itself, then it *is* the sort of thing that can be understood in this sense: it is used to refer to itself and expresses a sense that presents itself as itself: the time of utterance.<sup>30</sup>

On Künne's view, in a case like 'It is raining in Brooklyn', it is the *combination* of the time of utterance and the present-tense inflection of the verb that expresses a sense, where neither expresses a sense independently. He compares this to cases like the contribution of 'is' and 'raining' to expressing the sense of 'is raining', claiming that neither expresses a sense on its own on Frege's view.<sup>31</sup> Since the present-tense inflection is linguistic, it is the sort of thing that can be understood, and that's why the combination can express a sense that refers to the time of utterance. (ibid pg. 544)

---

<sup>28</sup>In support of this, Künne claims that the predicate 'is true' for Frege can be part of the expression of a thought without contributing a sense to that thought, and furthermore lacks a referent. (ibid pg. 541, 543) I demonstrated in the last chapter that this is based on a misunderstanding of Frege's actual view: 'is true' expresses a function-sense that maps the canonical sense of a thought to the thought itself, and thus refers to a function from a thought to its truth-value.

<sup>29</sup>Such knowledge may be (I would say usually *is*) tacit.

<sup>30</sup>That is not to say that the view isn't strange. We will discuss broader reasons for accepting it in later sections.

<sup>31</sup>Actually this isn't at all clear. Frege *does* speak of an expression such as 'a horse' as having a reference independently of the copula (e.g. 1892b, PW pg. 99), which implies that it expresses an independent sense as well. Furthermore, the 'is' of identity clearly expresses a sense (referring to the identity relation), and even the 'is' of predication can be seen as expressing a sense that refers to the relation between an object and a property it instantiates (Frege seems to assert this in (Frege 1906a, PW pg. 178)). However, there are other linguistic cases that support Künne's point: e.g. a proper name made up of multiple nonsense words.



But what contribution does the time of utterance make? At the least, it must contribute itself as the referent of the expressed sense. If I read ‘It is raining in Brooklyn’ without knowing when it was written, then I can’t recognize which thought was originally expressed, even though I perfectly understand the present-tense inflection of the verb. What I fail to understand is when the sentence was written, and this is what’s preventing me from recognizing which thought was originally expressed by it.<sup>32</sup>

The point being, the time of utterance makes an *independent* contribution to the expression of the thought. By holding the utterance ‘It is raining in Brooklyn’ fixed while varying the time of utterance, we vary the thought expressed in a systematic way. Similarly, by holding the time of utterance fixed while varying the tense, we also vary the thought expressed in a systematic way (‘it was raining’; ‘it will rain’). The time of utterance makes its own contribution to truth-conditions independently of the present tense (assuming, with Frege, that it itself makes a contribution as opposed to a distinct expression for it making a contribution). Hence, it expresses a sense independently as well. In fact, it is hard to see what else could be needed to conclude that something expresses an independent sense.<sup>33</sup> Thus, I see no reason to reject Kripke’s interpretation. One shouldn’t conclude that Frege didn’t hold a view just because one personally finds it strange.

Let’s turn to (Burge 2012). In that paper, Burge makes various arguments against

---

<sup>32</sup>Whether this thought can be *grasped* at another time will be discussed in the next section.

<sup>33</sup>Wouldn’t this argument also show that ‘a horse’ expresses a distinct sense from ‘is’ in an expression like ‘is a horse’? Indeed it would, and as I pointed out in footnote 31, Frege’s actual view was that concept-words refer to concepts independently of the copula. (Here I use *Frege’s* sense of ‘concept’ (‘Begriff’); the point may be clearer if you replace this term by ‘property’.)

Kripke's interpretation. He claims that on Frege's view, when I say "Today it is raining" on one day and "Yesterday it was raining" the next, 'Today' and 'Yesterday' express the same sense. (ibid pg. 558 in the 2013 reprint) But let's return to passage (1) above. Here Frege does say that if someone wants to express the same thought today that they expressed yesterday by saying e.g. "Today it is raining", they must instead say "Yesterday it was raining." Frege claims that the verbal expression must be different in order to cancel out the change in sense that would otherwise be brought about by the change in time of utterance.

Note that he never says that 'today' and 'yesterday' express the same sense. Indeed, he seems to say that they express *distinct* senses that combine with the senses expressed by the *distinct* times of utterance to express the same thought. This is yet another example of Frege's account of multiple decompositions of thought that I heavily emphasized in Chapter 1, an account under which there are many distinct ways to form the same thought out of component senses.<sup>34</sup> On such a view, one would naturally expect that varying one component of a thought while holding the rest of the thought fixed results in a distinct thought, and hence in order to preserve the thought one has to vary another component as well in order to cancel out the change. And this seems to be exactly Frege's point.<sup>35</sup>

---

<sup>34</sup>Both Burge and Kripke mistakenly attribute the unique decomposition view to Frege, Burge in (ibid pg. 560) and Kripke in (2008, pg. 270).

<sup>35</sup>Note, however, that we gave this intuition up in the last section. See footnote 25. In fact, any conception of the sense of 'Today' under which it is a function that can map distinct senses denoting times to the same sense denoting the day containing those times would require giving this up. I suspect Frege didn't consider a case in which I utter 'Today it was raining' multiple times on the same day, since it seems clear that he'd want there to be cases in which these utterances can express the same thought.

Burge also points out an earlier passage from Frege: “Words like ‘here’ and ‘now’ only acquire their full sense through the circumstances in which they are used.” (Frege 1897, PW pg. 135) Burge reads this passage as implying that e.g. ‘today’ cannot express a sense that denotes a function from a time to the day containing it, since it would then express its full sense in all circumstances. (Burge 2012, pg. 560-1, 590) But Frege could just as easily be read as implying that words like ‘today’ express *incomplete* senses that need to be saturated in order to refer to a duration of time. They may still express the same incomplete sense in all contexts. Furthermore, this passage was written twenty years before “Thoughts”. (1918-19a) Hence, Frege could have easily changed his mind in the interim.

Finally, Burge resists Kripke’s conclusion that for Frege times of utterance are used as part of the expression of the thought, i.e. they are used as names of themselves. While he admits that passage (1) can be naturally read in this way, he argues that the surrounding context suggests a different interpretation. (Burge 2012, pg. 591)

Burge points out that at the beginning of passage (1) Frege speaks of a time-indication being conveyed by the present tense. He interprets this as implying that in an utterance of “Today it is raining”, the present tense refers to the time of utterance. Burge then wonders why Frege would also need the time of utterance to refer to itself, since it is already denoted by the present tense. (ibid pg. 591) Note however that even on Burge’s view in an utterance like “It is raining now” there are two denotations of the time of utterance (the present tense and ‘now’). Furthermore, it isn’t clear that Frege meant that the present tense *denotes* the time of utterance; it may merely *indicate* that a time of utterance needs to be supplied in order for a complete thought

to be expressed. Compare the past or future tense, neither of which denote particular times (at least in normal cases). In any case, this isn't a very strong objection to Kripke's interpretation.

Burge suggests an alternative interpretation: "The time of utterance would be 'part of the thought expression' only in the loose sense that it is specified (hence denoted) by the thinker's use of contextual circumstances to express a mode of presentation that appropriately determines the time." (ibid pg. 592) This is a difficult reading to maintain. If the time of utterance doesn't contribute to the expression of the thought, then why did Frege say that it is part of the expression of the thought? Also, what "contextual circumstances" could Burge have in mind here that could be used to express a sense that denotes the time of utterance (other than the time of utterance itself)? Surely pointing at the time with one's finger isn't what Frege had in mind.

Burge also claims (ibid pg. 592) that in earlier work Frege said that the present tense can denote the time of utterance, citing a passage from "On Sense and Reference" (Frege 1892a) as evidence. Actually, this is unclear. What Frege actually says is "If an instant of time is to be indefinitely indicated in both the antecedent and consequent clause, this is often achieved merely by using the present tense of the verb, which in such a case however does not indicate the temporal present." (ibid pg. 43/171) His example is "When the Sun is in the tropic of Cancer, the longest day in the northern hemisphere occurs." Frege also says that we can't express the sense of the subordinate clause by simply saying "The Sun is in the tropic of Cancer", because "this would refer to our present time and thereby change the sense." (ibid

pg. 44/171) So, Frege says that the subordinate clause “When the Sun is in the tropic of Cancer...” indefinitely indicates a time, and when we say “The Sun is in the tropic of Cancer” we refer to our present time, but he doesn’t say that the present tense *itself* refers to our present time. Furthermore, that paper was written 25 years before “Thoughts”, and so Frege could have easily changed his mind about a few things.

More importantly, whether or not Burge is correct that for Frege the present tense refers to the time of utterance, it is clear that Frege *also* held that the time of utterance itself is part of the expression of the thought, for he says so explicitly in several places. Furthermore, Burge’s move simply won’t work in the case of past-tense or future-tense utterances: ‘Yesterday it was raining’ needs to be supplemented in order to express a complete thought; the past tense does *not* refer to a definite time, but instead indicates that the event in question took place prior to time of utterance. On Frege’s actual view, the sense of ‘Yesterday’ is saturated by a sense expressed by the time of utterance, and this is why the utterance refers to a particular day.

With that said, let’s turn to a general objection to any Fregean account under which e.g. ‘Today it is raining’ and ‘Yesterday it was raining’ can express the same thought.

### 3.4 Differing Conceptions of Sense

David Kaplan raised an objection to Frege’s position on thought-preservation. Someone could lose track of how many days have passed since the day in question, and as a result they might not realize whether they are expressing the same thought with

a new utterance. (1989, pg. 501 fn. 26) For example, suppose the person originally uttered ‘Today it is raining here’, lost track of how many days have passed, and then attempts to express the same thought by uttering ‘It was raining here the day before yesterday’ after three days have in fact passed. But how could an individual fail to recognize whether two thoughts are distinct or identical? Shouldn’t the identity-conditions of thoughts be transparent to a thinker of those thoughts?<sup>36</sup>

This might seem to be a general objection to any attempt to construct a Fregean theory of senses that satisfies Frege’s desire for thought preservation across shifts in context. However, whether it counts as an objection will depend upon the details of one’s conception of sense in the first place. As Tyler Burge points out (2012, pg. 556-7), Frege himself wanted senses to play various roles in theorizing about language and thought, and it isn’t clear that all of these roles can be played by the same notion.

When Frege introduced the notion of sense in “Function and Concept” (1891), he argued that the expressions ‘Morning Star’ and ‘Evening Star’ must have different senses although they have the same reference, since someone who didn’t know that

---

<sup>36</sup>Actually, Frege allowed for cases in mathematics in which senses are not fully transparent. His “Logic in Mathematics” (1914) includes a detailed discussion of mathematical definition in which Frege says, “The fact is that if we really do have a clear grasp of the sense of the simple sign, then it cannot be doubtful whether it agrees with the sense of the complex expression. If this is open to question although we can clearly recognize the sense of the complex expression from the way it is put together, then the reason must lie in the fact that we do not have a clear grasp of the sense of the simple sign, but that its outlines are confused as if we saw it through a mist.” (ibid pg. 211)

Indeed, Frege’s own project in the philosophy of mathematics consisted partially in elucidating the senses of mathematical expressions, senses that mathematicians had grasped all along, albeit unclearly. However, the examples discussed by Kaplan and others are importantly different. An idealized thinker can be expected to recognize whether the sense of a simple mathematical expression is identical to the sense of a complex mathematical expression, but they cannot be expected to recognize whether two days have passed or three. In the latter sort of case, they are merely lacking empirical information about the world, and this is not the sort of information that one should be expected to have in order to clearly grasp a sense. And if an idealized thinker clearly grasps two senses, how could they fail to recognize whether these senses are identical or distinct? So, at least, goes the objection. I am grateful to Haim Gaifman for helpful discussion of these issues.

the Morning Star is the Evening Star could believe that the Evening Star is a planet with a shorter period of revolution than the Earth without believing that the Morning Star has the same property. (ibid pg. 14/145) Let's call this the *cognitive significance* conception of sense, with the suggestion being that two thoughts are distinct if it is possible for an idealized thinker to believe the first thought as expressed by sentence 'p' without believing the second thought as expressed by sentence 'q'.<sup>37</sup>

It is this conception of sense that most obviously falls prey to Kaplan's scenarios. For example, couldn't a thinker believe the thought she expressed yesterday with the utterance 'Today it is raining here' while failing to believe the thought expressed today by 'Yesterday it was raining here', since she mistakenly thinks that two days have passed? But then how could these two thoughts be identical?

Actually, this case is a bit more complicated. For one, the person in our example still believes the thought originally expressed and has access to that thought via preservative memory; she would merely fail to *express* it properly (her actual utterance of 'It was raining here the day before yesterday' would express a distinct thought, unbeknownst to her). In effect, the person would fail to say what she really means. One might similarly argue that she does in fact believe the thought expressed today by 'Yesterday it was raining here'; she merely fails to recognize *which* thought is expressed by that utterance, and hence she may be unwilling to assent to it. So, a defender of this conception of sense may reply to the objection by admitting that even an idealized thinker may not know whether two utterances express the same

---

<sup>37</sup>This is only meant to be a sufficient condition for distinctness. It may be impossible for an idealized thinker to fail to believe a mathematical truth (depending on the details of our idealization), but that wouldn't mean that all mathematical truths are identical.

thought, and furthermore may not know how to express a particular thought.

The reader may feel that this merely dodges the real issue. Surely if thoughts are meant to capture something like cognitive significance, then it should be cognitively transparent to an idealized thinker whether two utterances express the same thought or two distinct thoughts. By assumption, an idealized thinker perfectly understands both utterances (she is ideally linguistically competent), so how could she fail to recognize the identity-conditions of the thoughts expressed by them? Then again, it seems to be a necessary condition of understanding these sorts of utterances that one knows various facts about the contexts in which they were made, and this knowledge is precisely what is lacking in Kaplan's examples.

I prefer a different response. Whatever Frege's actual view on the matter, we should not take the cognitive significance conception of sense as fundamental. The most fundamental conception of sense is the *truth-conditional* conception specified by Frege in Volume I of the *Grundgesetze* (1893, pg. 50-1) and in the passage I cited in footnote 16.<sup>38</sup> On this conception, the sense of an expression is determined by the contribution that expression makes to determining the truth-conditions of complete sentences in which it occurs (this includes the senses expressed by contextual factors), and a thought is the sort of content relevant to determining whether an assertion is true or false. This is the conception of sense most relevant to *logic*, and it is the

---

<sup>38</sup>The relevant *Grundgesetze* passage is the following: "Every such name of a truth-value *expresses* a sense, a *thought*. For owing to our stipulations, it is determined under which conditions it refers to the True. The sense of this name, the *thought*, is: that these conditions are fulfilled...Now, the simple or complex names of which the name of a truth-value consists contribute to expressing the thought, and this contribution of the individual name is its *sense*. If a name is part of the name of a truth-value, then the sense of the former name is part of the thought expressed by the latter." (1893, pg. 50-1)



conception I have been working with throughout this dissertation.

Given the truth-conditional conception of sense, what should we conclude from Kaplan's examples? Kaplan's examples show that the truth-conditions of an utterance can depend upon external factors. In particular, they can depend upon features of the context in which an utterance is made. For this reason, if a thinker lacks information about the context of an utterance, she may fail to recognize which truth-conditions are expressed by that utterance. In particular, she may fail to recognize whether two utterances have the same truth-conditions. But this possibility is perfectly compatible with a truth-conditional conception of sense.

Indeed, on the Frege interpretation defended in the previous two sections, the time and location of utterance themselves often make contributions to the truth-conditions of an utterance. In order to understand a tensed utterance, for example, an agent must be aware of the time of utterance. However, it doesn't follow that the agent must be able to locate that time along a timeline relative to other times. They may merely be aware of the time of utterance *as* the time of utterance or *as* the time it actually is rather than as e.g. ten minutes from the last utterance.<sup>39</sup>

Similarly, I may be aware of a day as the day it actually is, some time may pass, and I may again be aware of a day as the day it actually is. Whether these senses are the same or distinct depends upon how much time has passed. If I am unsure how much time has passed, I may also be unsure whether these senses are the same, although I may perfectly grasp both senses. I know precisely what contributions

---

<sup>39</sup>Again, I do not insist that the sense they grasp is equivalent to that of a description such as 'the time of utterance'.

the expression ‘the day it actually is’ and the time of utterance make to the truth-conditions of complete declarative sentences. In order to avoid this, one would have to hold the implausible view that ‘the day it actually is’ *always* expresses a distinct sense when it is uttered at a distinct time.

This position is further supported if one accepts the view that the senses expressed by places, times, and subjects are individuated in terms of the places, times, and subjects themselves. In that case, if one is unable to compare two places to each other, one will also be unable to compare the corresponding autonomous senses. In particular, one will be unable to determine whether these senses are identical or distinct. Compare a case in which each eye is presented with a distinct image, and although the images are identical, one is unable to determine whether they present distinct objects or the same object.

So, grasping the truth-conditions of a tensed utterance does not require being able to locate the duration of time referred to along a timeline relative to other times. Hence, grasping the truth-conditions of two tensed utterances does not require being able to locate the durations of time referred to relative to each other. Hence, I may be unable to recognize whether two utterances I perfectly understand in fact have the same truth-conditions, since neither utterance refers to a duration of time in a way that allows me to locate these two durations on a timeline relative to each other. Being able to do so is unnecessary for understanding.

With that said, let’s return to considering the consequences of Frege’s own view.

### 3.5 Spaces, Times, and Objectivity

What does it take to represent a time *as a time*? Note that there is a significant difference between representing a time merely as an *object* (an unusual case to be sure) and representing a time as a *time*.<sup>40</sup> This difference ought to be reflected at the level of sense.

In order to represent a time as a time and not merely as an object, one has to represent it as part of a *timeline*. It is essential to being a time that it is part of a timeline, and this ought to be reflected at the level of representation. Similarly, in order to represent a place as a place one has to represent it as part of a larger space.<sup>41</sup> These points may already suggest that the senses expressed by expressions denoting durations of time and regions of space ought to be structured like timelines and spaces, respectively.<sup>42</sup>

In further support of this position, consider perceptual experience. I am currently perceiving a region of physical space as occupied by myself and various other objects

---

<sup>40</sup>As an example of the former sort of case, one might know that Jim is thinking about something without knowing what that thing is, even though it turns out that Jim is thinking about a particular time. This knowledge would give one a way of thinking of the object of Jim's thought that doesn't represent it as a time, but merely as an object.

<sup>41</sup>As Frege himself argued in Volume II of the *Grundgesetze*, similar remarks apply to magnitudes in general: "A thing is a magnitude not in itself but only insofar as it belongs, with other objects, to a class that is a domain of magnitudes." (1903a, pg. 159) That is to say, in order for something to be a magnitude, it has to be part of a class of magnitudes that can be added to each other, subtracted from each other, etc. Representing a magnitude as a magnitude requires representing it as part of a larger system of magnitudes.

One might make similar remarks about other systems such as the system of natural numbers. Arguably, in order to represent the number three as the number three, I have to represent it as part of a larger system of natural numbers. If I don't have any grasp of this larger system, I can't think of the number three as the number it is.

<sup>42</sup>That is to say, one might think that in order for a sense to represent an object as part of a larger whole, it must include components which represent the whole. I give further arguments for this position below.

bearing various spatial relations to each other. Notice that my perceptions not only represent physical space as a space, they are also themselves located within a wider *perceptual space* in which they bear various spatial relations to each other. That is to say, my perceptual experience itself has spatial structure.<sup>43</sup> Hence, it should be no surprise that such structure is preserved in thought about durations of time and regions of space.

Let's return to the thought expressed by an utterance of 'Today it is raining here'. I argued above that one ought to be able to express this same thought at a later time, and agreed with Gareth Evans that one could believe this same thought over a period of time, any point at which it may have to be expressed in a different way ("we must run to keep still"). In fact, we shouldn't ascribe a belief with such content to people if they aren't capable of *preserving* this belief over time (although they may of course choose to give it up). Beliefs aren't the sorts of things that merely last for an instant.

Furthermore, I argued above that Frege himself wanted these sorts of thoughts to be expressible at multiple times, and that he suggested using his treatment of thoughts as having multiple decompositions in order to do so. Let's consider further what might be said in support of such a view.

When I assert that today it is raining here, I am committed to its still being true that it rained here today when considered from the perspective of tomorrow. One might say that the truth-conditions of my assertion do not depend on the temporal

---

<sup>43</sup>In support of this, note that it makes perfect sense to talk about the relative sizes of my visual perceptions independently of the sizes of their objects. My perception of one tree might take up far more space in my perceptual experience than my perception of a second tree, even though I know that both trees are the same size (and am perceiving them as such). The former tree might be closer to me.

perspective from which I am viewing the event in question. As time passes from today to tomorrow, my temporal perspective on the event necessarily shifts. The Fregean position is that I can still express the same thought with the same truth-conditions even while my temporal perspective shifts; this thought is not perspective-dependent. At the same time, any time I express this thought, a part of this thought will refer to the time of utterance. So, the shift in temporal perspective is captured at the level of thought, but this does not force a different thought to be expressed with every shift. In many cases, the shift is cancelled out by other changes in expression.

So, given that the sense expressed by the combination of the time of utterance and ‘today’ has the structure of a timeline, each point along this timeline will correspond to a particular temporal perspective on the event in question. A shift in this temporal perspective from today to tomorrow does not affect the sense expressed because the shift in temporal perspective is cancelled out by the transition from the sense of ‘today’ to the sense of ‘tomorrow’, each of which denotes a distinct function from times to days (the former from a time to the day containing it; the latter from a time to the day prior to the day containing it). The governing idea is that although one’s temporal perspective has shifted, one is still thinking of the event in question in the same way: as an event of raining that took place on a particular day at a particular place.

Let’s continue by considering an elegant generalization of this account to cases in perception. To do so, let’s briefly discuss perception in a bit more detail.<sup>44</sup>

---

<sup>44</sup>My discussion follows (Burge 2010), which includes a detailed discussion of the relevant literature in perceptual psychology.

## Perception

A perception is the paradigm case of a mental representation. A perception represents an object as being a certain way, and hence may be regarded as having a representational content. This content can be assessed as accurate or inaccurate. However, unlike judgments, perceptions are not the sorts of things that can be rational or irrational. We are not epistemically responsible for our perceptions; they generally occur automatically in response to sensory stimulation.

My current visual perception of an object as a coffee cup carries with it certain *expectations* concerning how that object will appear from other angles, how that object will behave if I interact with it, etc. These expectations plausibly operate at the subpersonal level and influence my behavior. Note that these expectations can be inaccurate: there are famous examples of objects that appear to be cubes from one angle and spheres from another. Given that my initial perception of such an object presented it as a cube, there is a natural sense of “contradiction” in which my later perception of that object as a sphere contradicts my initial perception. One may characterize the representational contents in such a way that the contradiction becomes explicit.

Since perceptions always present their objects in particular *ways*, it is natural to treat the representational contents of perceptions as the perceptual analogues of senses. Indeed, many contemporary philosophers of mind such as Tyler Burge and Christopher Peacocke extend their accounts of representational content to cover both the perceptual and the linguistic cases. The natural way of accommodating

perceptual representational contents on my view of senses is to characterize them via their contributions to *accuracy-conditions* (as opposed to truth-conditions).

With that said, let's shift our attention to the spatial case. What is the analogue of a particular duration of time in the case of space? It is a particular region of space having a particular three-dimensional geometric shape: the shift from a one-dimensional length of time to a three-dimensional region of space is due to the difference in our representations of the dimensionality of time and space.<sup>45</sup>

So, what is the analogue of the shift in temporal perspective on an event as occurring over a particular duration of time? It is the shift in spatial perspective on an object as having a particular geometric shape. A perception of an object as a cube carries an expectation (or *commitment*) regarding how that object will look from a different perspective. That is to say, if that perception is accurate, then so would be a perception of that object that represents that object in the relevant way from a distinct position in space (call this "accuracy-preservation"). If I shift my perspective on that object and it suddenly appears to be a sphere, this new perception contradicts my initial perception of it as a cube. My perception of the object as a cube is not consistent with a perception of it as a sphere from a different angle.<sup>46</sup> This is part of what it is for my perception to be an *objective* representation, a representation of an object as existing independently of my perspective on it.

---

<sup>45</sup>Note that the actual number of dimensions of physical space is irrelevant here: however many dimensions space actually has, we clearly *represent* it as three-dimensional.

<sup>46</sup>This assumes the object didn't change shape. The commitment carried by my perception might be better expressed counterfactually: if I had been in that position instead of this one, I would have still perceived the object as a cube, and furthermore its appearance as a cube would shift in a systematic way (i.e. specifying the other position determines precisely how the appearance would shift as well; e.g. moving closer to the cube shifts my perception of it in a determinate way).

An elegant way to capture this idea is to let the representational content of such a perception have the structure of a *space* with the object in question as origin: any point in the space corresponds to a particular perspective on that object from which I would still perceive it as a cube (and furthermore perceive it as the same size: change in apparent size is cancelled out by change in distance from the origin). By perceiving the object as a cube, I am committed to its still appearing to be a cube as my perspective on that object shifts. The representational content of my perception (at the level of abstraction we're considering) is perspective-independent; I can perceive the object in the same way (as a cube) even as I shift my position with respect to it, even though this shift in position corresponds to a shift in the way the object appears. I am *committed* to its appearing in a particular way from a different perspective by virtue of perceiving it as a cube in the first place. Treating the representational content itself as spatially structured captures these commitments in an elegant way.<sup>47 48</sup>

Furthermore, this treatment of the representational content of such a perception allows one to explain why a distinct perceiver located at a distinct point in space

---

<sup>47</sup>I'm leaving out a detailed discussion of issues involving vagueness that arise in such cases. I take it to be clear that there isn't any exact region in space such that we perceive an object as located precisely in that region. Hence, vagueness in represented position would need to be captured at the level of representational content (and sense, in the case of thought). Formalizing a model like this in something like Euclidean geometry would merely be a useful idealization. Strictly speaking, the content-space is permeated by vagueness.

There is an additional level of idealization here. In many cases it can take genuine effort for a normal perceiver to mentally rotate an object that they are currently perceiving as having a particular geometric shape. I take it that this would be unnecessary for an idealized thinker (otherwise, she wouldn't be perceiving it as having that shape in the first place). I am grateful to John Morrison for pressing me to address this point.

<sup>48</sup>Haim Gaifman has pointed out to me that there are similar ideas in Bertrand Russell, e.g. his "The Relation of Sense-data to Physics" (1914). Note that Russell is attempting to *define* a physical thing as the class of its appearances (ibid pg. 129), while I am merely arguing for treating the representational content of such a perception as a space of possible perspectives.



while perceiving the same object can still perceive it in the same way as I do: we both perceive the same object as a cube, and we are both committed to its appearing in a particular way from the position of the other perceiver. The difference in perspective is a result of our being located at different points in physical space, and this difference in perspective is captured at the level of content by the corresponding points in the content-space.<sup>49</sup>

The analogue of multiple decompositions in this case would be the following: the representational content has the structure of a space centered around the object in question, and the full content can be grasped as a function of any point in the space, where that point corresponds to a particular position of the perceiver relative to the object perceived (the space with that point removed is thus a function from that point to the complete space). Although the position of the perceiver shifts the same representational content is being expressed by each perception: the perception remains a perception of the object as a cube.

## **Thought**

The position we're considering on the senses expressed by utterances of e.g. 'today' and 'here' is the natural extension of this position from the realm of perception to the realm of thought. The sense of 'here' (in combination with the place of utterance) has

---

<sup>49</sup>If one really wanted to insist that by virtue of being located at different places, two people must be perceiving the cube in different ways, what would prevent the analogous move with respect to times (i.e. a difference in time implies a difference in way of perceiving)? It would then follow from the fact that no two people are ever at the same place at the same time that no two people can ever perceive the same object in the same way. Clearly there is a place in a theory of representational content for a level of abstraction at which two people can perceive an object in the same way.

the structure of a space where each point in this content-space corresponds to a perspective on the place referred to by ‘here’. The full sense is perspective-independent (I could think of that place in the same way even if I were in a different position). One can capture both the shift in perspective and the perspective-independence of the sense via *multiple decompositions: this very same sense can be grasped as a function of any point in the content-space*. As I shift my position in physical space, my perspective on the place referred to by my initial use of ‘here’ shifts, and hence I grasp the sense expressed in a different way although it remains the same sense.

Furthermore, by virtue of making an assertion like “It is raining here”, I am committed to its presently raining at that location when viewed from a different position in space. For the same reason, someone else could make an assertion at the same time and express the very same thought, even though they are located at a different position in space. Letting the sense expressed by ‘here’ and the location of utterance have the structure of a space allows us to respect both of these points, just like in the perceptual case. Hence, our Fregean position appears defensible after all.

### **3.6 Modality and Self-Reference**

In the last section we saw various reasons for accepting the Fregean position that senses referring to durations of time and regions of space can themselves be structured like timelines and spaces, respectively, and how such a position gives an elegant solution to some of the issues involving thought preservation that I raised in the Intro-

duction.<sup>50</sup> That wasn't the only problem that a Fregean treatment of demonstratives and indexicals had to overcome, however. There were also serious issues involving modality.<sup>51</sup>

Recall that the problem was that many demonstratives and indexicals seem to behave like *rigid designators*, designating the same entity even when evaluated with respect to a different possible state of the world. It was unclear how to give an account of senses that respects this point, since the senses of e.g. descriptions are notoriously nonrigid (e.g. 'the inventor of bifocals' denotes Benjamin Franklin when evaluated with respect to the actual state of the world, but may denote someone else when evaluated with respect to a different possible state of the world).

Recall Frege's actual treatment of contextual factors under the interpretation defended above: e.g. in the context of an utterance of 'Today it is raining here', the speaker uses both the time and place of utterance as expressions for themselves. By virtue of doing so, these contextual factors express senses that present themselves as themselves. This is yet another example of Frege's use of self-reference in his theory of truth-conditional content, of which the most obvious analogue is the sense expressed by a word in a direct quotation context (e.g. 'John said, "I am wearing a brown hat"'), a sense which presents that word as itself (the word is thus self-referential).

Note the following interesting fact about the sense expressed by an entity when that entity is used as a name of itself: this sense corresponds to a way of referring to

---

<sup>50</sup>We will go through these issues in more detail in the final section.

<sup>51</sup>As is well known, Frege himself was uninterested in modality. However, the modal objections considered in the Introduction have been very influential and I personally consider them worth taking seriously. Again, I make no commitment to a possible worlds semantics for modality, but we will see that the objections can be answered even given such a framework.

that entity that couldn't possibly be achieved by any other entity. Only the object itself can refer to itself in a self-referential way. Any other entity would necessarily have to refer to that object in a *different* way. Furthermore, if an object is being used to refer to itself in a self-referential way, then it cannot possibly refer to a distinct object. This implies that a sense of autonymous designation maintains fixed reference to its actual referent even when evaluated with respect to a distinct possible state of the world. Autonymous senses are *rigid designators*.<sup>52</sup>

Furthermore, the senses expressed by 'I', 'here', and 'now' will be rigid designators, since all three denote identity functions (from persons to persons, places to places, and times to times, respectively). Hence the complete senses expressed by the combination of 'I' and the speaker/thinker, 'here' and the place of utterance, and 'now' and the time of utterance will also be rigid designators (e.g. the first always denotes the result of applying the identity function to the *actual* speaker/thinker, since the sense expressed by the speaker/thinker used self-referentially is a rigid designator, as we just discussed).

What about other indexicals like 'today', 'yesterday', and 'the day after tomorrow'? The rigidity of 'today', 'yesterday', etc. will follow from the metaphysical assumption that times and places have their locations *essentially*. A particular time couldn't possibly be located at a different place along the timeline (a time is the time

---

<sup>52</sup>But couldn't a *distinct* entity refer in the same way by referring to *itself*? There is a sense in which this is referring in the same way, but this notion of a way of referring doesn't correspond to Frege's notion of "sense". Note that I can always take two objects and use them as names of themselves. By virtue of doing so they are expressing senses, and these senses must be distinct because they have distinct referents. The sense in which the two objects are referring in the same way can be captured by saying the two senses are instances of the same *type*; they have the same *structure*.

that it is, and not some other time), and a particular point in space couldn't possibly be located at a different point in space (it isn't clear what that would even mean; in any case it seems to contradict the necessity of identity). Hence, the duration referred to by a use of 'today' is the same duration even when evaluated counterfactually, as is the duration referred to by 'yesterday', etc.<sup>53 54</sup>

What about demonstratives like 'this' or 'that'? I mentioned above that Frege's own view was that the combination of e.g. an utterance of 'that' and an act of pointing expresses a sense that refers to the object pointed at. In isolation, 'that' refers to an identity function from objects to objects; its sense must be saturated by a sense referring to an object (such as the sense expressed by an act of pointing) in order to result in a complete sense. I pointed out that Frege's view has been criticized for getting the truth-conditions of various utterances wrong, since 'that might not have been the object I am pointing at' can be true while 'that might not have been that' is false. The point being, I could have pointed at a different object, and in such a counterfactual situation the sense expressed by my act of pointing would have

---

<sup>53</sup>Our language doesn't seem to include good spatial analogues of indexicals like 'today'. 'There' functions more like a demonstrative, needing to be accompanied by an act of pointing (or the salience of a location in the context of a conversation) in order to pick out a unique place. The spatial analogue of an indexical like 'today' or 'yesterday' would be an expression that picks out a particular region of space defined relative to the location of utterance (or another fixed location). Perhaps expressions like 'two meters to my left' come close.

<sup>54</sup>There is a complication here. There is a use of 'today' under which it is true that today might have been longer than it is: e.g. if the Earth rotated more slowly there might have been more hours in a day. I've been treating 'today' as if it refers to a (roughly) 24-hour duration of time, whereas one might insist that a day is defined as something like the amount of time it takes the Earth to complete a rotation.

One can capture this sense by insisting that the day itself (picked out as the day containing the time of utterance) is being used as a name of itself, thus expressing an autonomous, rigidly-designating sense. The incomplete sense expressed by 'today' will refer to a function from days to themselves, and it will need to be saturated by such an autonomous sense in order to refer to the appropriate day. Similar remarks apply to other expressions denoting particular days, years, etc.

referred to a distinct object, but ‘that’ seems to function as a rigid designator.

The reader may have already guessed my suggestion for fixing Frege’s view: simply apply the same technique I described above, a technique that Frege himself used in cases of subjects, places, and times. On this modified view, ‘that’ in isolation still refers to an identity function from objects to objects, but its sense needs to be saturated by an autonymous sense resulting from using the object in question as a name of itself. One can use the object in this way because one is already representing it in another way, e.g. via perception. On this account, demonstratives function as rigid designators, as desired. Furthermore, someone else can point to the same object while uttering ‘that’ and express the very same sense, since the sense expressed no longer depends on the details of the act of pointing.<sup>55</sup> <sup>56</sup>

This technique for ensuring rigid designation may seem like a mere technical trick. First of all, how does one use an object as a representation of itself? Secondly, are there any reasons for thinking that doing so has somehow been implicit in the way we’ve been using demonstratives and indexicals all along?

---

<sup>55</sup>Consider the following conversation (where both speakers are pointing to the same object): “That is a cube.” “I agree; that is a cube.” It seems plausible that the two speakers are expressing the same thought. While this wouldn’t follow on a view under which the sense expressed depends on the details of the act of pointing, it does follow on the view I just described. Hence, cases of intuitive thought preservation across speakers provide further support for this view.

<sup>56</sup>Can one extend this account to proper names, thereby ensuring that the senses expressed are rigid designators? One way of doing so would be to adopt the causal-historical framework described in (Kripke 1980). The idea would be that in a naming baptism, an object is used as a name of itself, where this object is first made available for thought by another representation of it (e.g. a perception). One then introduces the name in a manner intended to express the autonymous sense expressed by that object used as a name of itself, and this sense is preserved across transmissions of the name from one speaker to another. Two proper names would express distinct senses exactly if their corresponding naming baptisms were distinct (hence names in distinct languages could express the same sense, and two tokens of the same name-type could express distinct senses).

However, proper names are not my concern in this dissertation, and I offer the preceding as merely a sketch of what such an account might look like. I leave questions of its plausibility to the reader.

Regarding the first question, *any* object can in principle be used as a representation of anything you like. *A fortiori*, any object can be used as a representation of itself.<sup>57</sup> Regarding the second question, I already pointed out that we often use *words* as names of themselves (consider an utterance such as “John has four letters”). As further support, I offer the following speculative remarks on some aspects of self-consciousness.

A peculiar feature of conscious perception is that I can not only think about the objects of my perceptions, I can also think about those perceptions themselves. In the last chapter I stressed the importance of such a meta-representational ability for reflective judgment: in order to reflectively judge whether my perceptions are accurately representing reality, I need to be able to represent those perceptions as distinct from the objects they are perceptions of, and furthermore I need to represent those perceptions as expressing representational contents that I can compare with additional evidence I have about the objects being perceived. Without such a meta-representational capacity, I wouldn’t be able to judge whether reality is as it appears to be.

Let me make the following observations: (1) Here is an especially simple way of constructing a self-referential representation: simply combine a representation with what it represents. Since the former was a representation of the latter, their combination now represents itself. (2) I can usually remember my just having had a particular

---

<sup>57</sup>Consider a case in which someone is using various models arranged on a table to represent larger objects. They might want to represent a small object that is somehow relevant to the discussion, and they can do so by placing that very object on the table along with the models. In such a case, it seems natural to say that, since all the objects on the table are being used to represent an actual situation, the object in question is being used as a representation of itself.

conscious experience. Such a memory is a *representation* of that experience, and hence my mind must be constantly constructing meta-representations of my present conscious experience. (3) Conscious perception has a peculiar phenomenology: thinking *about* a present conscious perception doesn't seem very different (phenomenologically speaking) from *having* that perception (I encourage the reader to compare the two experiences).<sup>58</sup>

Observation (3) might suggest that when thinking *about* a present conscious perception, I am using that very perception as a representation of itself. Observations (1) and (2) suggest a model under which this would be the case: my present conscious experience includes a meta-representation of that very experience, a meta-representation that is combined with that experience in such a way that elements of that experience represent themselves (in addition to whatever else they may represent). This captures the self-referential phenomenology of human consciousness (more broadly, of the consciousness of beings with particular sorts of meta-representational capacities).<sup>59</sup>

Whatever one thinks of the suggested model, the more general point is that conscious perception has a peculiar phenomenology that seems well-explained by our perceptions being used as representations of themselves in the context of meta-level reflection on them.<sup>60</sup> If that is in fact the case, then using an entity as a representation

---

<sup>58</sup>If the reader wears glasses, they can probably remember having thoughts of the form “My perception of that object is very blurry”. Note that in such a context one is *using* one's perception (“that object”) and *mentioning* it (“my perception of”) at the same time. The transition from one to the other doesn't seem to make much of a phenomenological difference.

<sup>59</sup>I should mention that this model is heavily inspired by the A- and B-edition transcendental deductions of Immanuel Kant's *Critique of Pure Reason* (1781/1787). It should go without saying that I make no attempt at Kant exegesis in this dissertation.

<sup>60</sup>Note that this explains why I never have this sort of meta-representation of a conscious perception without also having that conscious perception. If the perception itself is being used to represent



of itself is ubiquitous in reflective judgment, and hence it should be no surprise that the same structure is present in thought involving words, indexicals, demonstratives, and (perhaps) proper names. Hence, there are in fact reasons for thinking that the account I developed above is more than a mere technical trick. Hence, there does seem to be a natural Fregean account of senses under which the senses of various expressions are rigid designators, as desired.

### 3.7 Preservation of Thought

In the preceding sections I sketched a model of Fregean senses of demonstratives and indexicals that avoids the modality objections and is grounded on Frege's actual view. I also discussed how this model handles some cases of thought preservation in an intuitive way. Let's discuss how this model handles thought preservation in further detail. While this isn't the last word on preservation of thought, I hope to establish that we've made significant progress in addressing the issues I raised in the Introduction.

I argued that 'Today it is raining here' uttered one day and 'Yesterday it was raining here' uttered the next can express the same thought, given that the sense denoting a duration of time is structured like a timeline that can be decomposed into a sense denoting the time of utterance and a function-sense mapping that sense to the complete timeline. Generalizing this picture, this thought can be expressed by a utterance at a distinct time so long as the sense denoting the day in question is

---

itself, this sort of meta-representation can't occur without its referent also occurring. It is immune to various sorts of error, including failing to refer.

a function *solely* of the sense expressed by the time that utterance is made used as a name of itself. So, ‘The day before yesterday it was raining here’ and ‘Tomorrow it will be raining here’ could express the same thought while ‘On my birthday it was raining here’ could not, since the sense of ‘my birthday’ is not a function of an autonomous sense denoting the time of utterance.<sup>61</sup> Analogously, a thought about a particular region of space (picked out relative to the speaker’s own location) could still be expressed after the speaker shifts their location in space.

I also argued above for a treatment of the sense of demonstratives where the sense expressed doesn’t depend on the details of the act of demonstration. As a result, multiple agents could point at the same object and utter ‘That is a cube’ while all expressing the same thought. This required a modification of Frege’s view, since according to Frege the combination of the sense of ‘that’ and the sense expressed by the act of pointing refers to the object pointed at. I noted above that this seems to get the truth-conditions of various utterances wrong since in such contexts ‘that’ seems to function as a rigid designator, and furthermore the multiple agents in my example intuitively ought to be able to express the same thought even though they are pointing at the object in different ways.

What about the thought expressed by an utterance of ‘John is wearing a brown hat’? Given that the sense expressed by the combination of the time of utterance itself and the implicit reference to the time of utterance (indicated by the present tense) has the structure of a timeline, this same thought could be expressed by a later utterance

---

<sup>61</sup>Similarly, the sense of e.g. ‘On September 16th, 2018, it was raining here’ will *not* express the same thought, since the relevant day is referred to via a calendar date rather than via a function of an autonomous sense denoting the time of utterance.

of ‘John was wearing a brown hat a moment ago’ or an earlier utterance of ‘John will be wearing a brown hat a moment from now’. Similar remarks apply to analogous cases of belief ascription and inference.

Hence, one can give a Fregean account of truth-conditional content that respects various intuitions about thought preservation across contexts and speakers, and such an account is roughly the account suggested by Frege in the first place (but missed by various scholars). As we discovered in the last chapter, it seems that Frege himself had a better treatment of these issues than many of our contemporaries.

# Chapter 4

## Presuppositions of Inference

### 4.1 Introduction

Inference is essential to rationality. An agent unable to draw conclusions is unable to reason. Arguably the purest form of inference is *deductive inference*, in which the truth of the premises guarantees the truth of the conclusion. More specifically, if inference is a transition between thoughts, where a thought is the sort of content relevant to determining whether an assertion is true or false, then the guarantee that the conclusion is true given that the premises are true can be explained solely by appeal to the *structure* of the thoughts making up the inference.<sup>1</sup> Let's begin by looking at this conception of deductive inference in a bit more detail.

This conception of deductive inference reveals part of the value of a formal language. Note the following: the sentences of the formal language correspond to

---

<sup>1</sup>I don't mean to say that such an explanation can be given without (implicitly or explicitly) appealing to various basic facts about truth, etc. The point is that nothing else about the *thoughts* involved in the inference is needed for the explanation of guaranteed truth-preservation other than basic facts about their structure.

particular structures instantiated by thoughts (each such sentence corresponds to a thought-type). Note that here I am using ‘structure’ in the sense of *form*, *pattern*, or *type* rather than the sense used by logicians (i.e. as synonymous with ‘model’). I will use ‘model’ for the latter notion. The relevant structures for present purposes are those that correspond to sentences in possible formal languages that can express the thoughts in question. For example, if the sentence ‘ $P(a)$ ’ expresses a particular thought once it is interpreted, that thought will instantiate a structure corresponding to the syntactic structure of that sentence.<sup>2</sup>

Anyway, since the sentences of the formal language correspond to particular structures instantiated by thoughts, one can thereby constrain the semantics of the formal language to only give an explanation of the truth-condition associated with a thought by appealing to a structure instantiated by that thought. This implies that one can only explain the guaranteed truth-preservation of an inferential transition from one collection of thoughts to another by appealing to structures instantiated by those thoughts, structures corresponding to sentences defined by the syntax of the formal language.

Take classical first-order logic as an example. A standard first-order language  $\mathcal{L}$  will include sentences of the form ‘ $P(a)$ ’, where ‘ $a$ ’ is a constant symbol and ‘ $P$ ’ is a one-place predicate symbol. According to the standard semantics of such a language, this sentence counts as true in a model  $\mathcal{M}$  of the language  $\mathcal{L}$  exactly if  $\mathcal{M}$  assigns

---

<sup>2</sup>Note that this doesn’t hold in general. Since a thought is an abstraction from the complete content of a declarative sentence, not just any syntactic structure of a sentence in a natural language will correspond to a structure instantiated by the thought expressed by that sentence. However, due to our focus on truth-preservation, the formal languages we’re considering have truth-conditional semantics. In effect, they are designed to express thoughts and make structural features of those thoughts transparent. Additional content has already been abstracted away.

‘ $a$ ’ an object  $a^{\mathcal{M}}$  in its domain  $D$  and assigns ‘ $P$ ’ a set of objects  $P^{\mathcal{M}}$  that is a subset of  $D$  such that  $a^{\mathcal{M}}$  is in fact an element of  $P^{\mathcal{M}}$ .<sup>3</sup> Intuitively, this sentence has been interpreted by the model in such a way that the object denoted by ‘ $a$ ’ is an element of the extension of the property denoted by ‘ $P$ ’, i.e.  $a$  instantiates  $P$ . This determines a structural feature of the truth-condition associated with any thought that instantiates the structure corresponding to ‘ $P(a)$ ’.

Now, suppose I want to demonstrate that the inference from a thought instantiating the structure corresponding to ‘ $P(a)$ ’ to a thought instantiating the structure corresponding to ‘ $\exists xP(x)$ ’ is truth-preserving. ‘ $\exists xP(x)$ ’ is true in a model  $\mathcal{M}$  of the language  $\mathcal{L}$  exactly if there is an object  $o$  in the domain  $D$  of  $\mathcal{M}$  such that  $o$  is an element of  $P^{\mathcal{M}}$ . If ‘ $P(a)$ ’ is true in  $\mathcal{M}$ , then  $a^{\mathcal{M}}$  is an element of  $P^{\mathcal{M}}$ , and hence ‘ $\exists xP(x)$ ’ is also true in  $\mathcal{M}$ . Hence, I can explain the truth-preservation of the inference from the former thought to the latter solely by appealing to structural features of their corresponding truth-conditions, structural features that are themselves determined by the structures corresponding to the two sentences of the language  $\mathcal{L}$ .

Hence, if one can demonstrate using such a formal language that if a collection of thoughts of specified types are all true, then another thought of a specified type is guaranteed to be true, one can be confident that this explanation of the truth-preservation of a corresponding deductive inference appeals solely to the structures determined by the syntax of the formal language. Provided that one has selected an appropriate formal language (on which more later), it follows that the inference is truth-preserving solely in virtue of *logical form*.

---

<sup>3</sup>For an elegant introduction to contemporary mathematical logic, see (Ebbinghaus et. al. 1994).

So, there is an intuitive sense in which deductive inference is guaranteed to be truth-preserving in virtue of form, and this sense can be made more precise via contemporary formal languages. But in order for an agent to actually *make* such an inference, they must be grasping the respective thoughts in the appropriate *way*.

Recall the conclusion of Chapter 1: thoughts can be grasped in many ways, and each way of grasping a thought makes various possible inferences salient while masking others. For example, if I grasp the thought that John has brown hair as unstructured (like an atomic sentence ' $p$ ' in propositional logic), I will be unable to recognize the possible inference to the thought that someone has brown hair. However, if I grasp the thought that John has brown hair as consisting of the function-sense  $\langle \text{has brown hair} \rangle ( )$  applied to the object-sense  $\langle \text{John} \rangle$ , I will be capable of recognizing this possible inference.

This particular structural feature, namely, being completely determined by a function-sense applied to an object-sense, is precisely the structure that corresponds to a sentence of the form ' $P(a)$ ' in a standard first-order language. And the corresponding structural feature of the determined truth-condition is: that the referent of the object-sense instantiates the referent of the function-sense.

Note that this characterization of deductive inference is available to us because of our conception of thoughts as structured entities that determine truth-conditions: on various other conceptions of "proposition" or "representational content" or what have you, this characterization could not be given. That is not to say that there isn't room for other conceptions of content (there obviously is), but it does suggest that the sort of content we've been focused on throughout this dissertation ought to have

a special place in the philosophy of logic.

But which structural features of thoughts ought to be considered *logical*? Which structural features are relevant to explaining what makes a particular inference a *deductive* inference? By choosing a particular formal language with a particular syntax and semantics to give an account of various deductive transitions between thoughts, we are *presupposing* a particular answer to this question rather than providing one. We are presupposing that the relevant structural features are precisely those that are determined by the syntax of the language.

In this chapter, I will attempt to make progress in determining which structural features of thoughts ought to count as logical. I will do so by considering various presuppositions of inference (and rational agency more generally). As in previous chapters, I will also consider connections to relevant aspects of perception and conscious experience.

## 4.2 A Fregean Conception of Truth

In Chapter 2 I argued that for Frege the sense of the predicate ‘is true’ is a function mapping the canonical sense of a thought to the thought itself (i.e. the referent of that canonical sense), which implies that it *refers to* a function from a thought to its truth-value.<sup>4</sup> I said little, however, about Frege’s famous conception of truth as an object: the True.<sup>5</sup>

---

<sup>4</sup>Recall that I argued that for Frege the thought expressed by ‘the thought that  $p$  is true’ is identical to the thought that  $p$ . Importantly, this is *not* the same thought as would be expressed by ‘the *sentence* ‘ $p$ ’ is true’.

<sup>5</sup>See for example (Frege 1891, pg. 13/144).



Part of the reason Frege treated truth as an object was mathematical elegance: Frege wanted the referent of a complete expression to be a function of the referents of its significant parts. So for example the referent of ‘the capital of Germany’ is a function of the referent of ‘Germany’ and the referent of ‘the capital of ( )’; in particular, the function referred to by ‘the capital of ( )’ maps Germany to Berlin. Hence, Frege also wanted the referent of a sentence to be a function of the referents of its significant parts. Furthermore, by letting sentences refer to truth-values, one can reduce several compositionality principles to a single principle, namely: if in a proper name, a component, which is a proper name, is substituted by a co-referential name, the reference of the whole name remains the same.<sup>6</sup> Importantly, even if we treat complete sentences as having referents, we must still distinguish sentences from names in general. As Frege pointed out, only complete sentences can be *asserted*.

Frege also argued that (1) we are only interested in the referents of parts of a sentence when we are interested in its truth-value,<sup>7</sup> and (2) the truth-value of a sentence does in fact functionally depend upon the referents of its parts (e.g. replace a significant expression in a sentence by a distinct expression with the same referent: the thought expressed may change but the truth-value will not).

In further support of (2), note that when we consider a thought  $\langle p \wedge q \rangle$ , we

---

<sup>6</sup>Otherwise, one would need four principles: one governing the reference of proper names containing other proper names, one governing the reference of proper names containing sentences, one governing the truth-value of sentences containing sentences, and one governing the truth-value of sentences containing proper names. I am grateful to Haim Gaifman for making these points.

<sup>7</sup>Frege’s example in “On Sense and Reference” is “Odysseus was set ashore at Ithaca while sound asleep.” If we read this sentence as part of a work of fiction, we don’t care whether ‘Odysseus’ actually has a referent; we only care about the thought expressed. As soon as we are interested in the truth or falsity of the sentence, we are interested in whether ‘Odysseus’ has a referent and whether that referent was in fact set ashore at Ithaca while sound asleep. (1892a, pg. 32-3/162-3)

can replace the thought  $\langle p \rangle$  by any other thought  $\langle r \rangle$  with the same truth-value while preserving the truth-value of the whole thought. While the thought  $\langle r \rangle$  may drastically differ from the thought  $\langle p \rangle$  in other ways, so long as its truth-value remains the same, so will the truth-value of the whole thought.<sup>8</sup> Furthermore, the truth-value of a sentence expressing a complete thought is going to be an *object* for Frege because an object is “anything that is not a function, so that an expression for it does not contain any empty place.” (1891, pg. 18/147) Since a complete sentence does not contain an empty place, its referent is an object.

In further support of Frege’s position, one might point to the overwhelming usefulness of truth-functions in the practice of logic, mathematics, computer science, etc. A truth-function maps truth-values to truth-values. Hence, it seems to treat truth and falsity as objects.

The reader may find these arguments unconvincing. Appeals to mathematical elegance and usefulness to mathematical practice may not be enough to conclude that sentences are the sorts of things that have referents at all, much less that they refer to the True and the False.<sup>9</sup>

To make further progress on these issues, let’s step back and consider why it is so useful to represent something as an object. What role is the notion of an object

---

<sup>8</sup>There are additional well-known Slingshot arguments that if sentences have reference, all true sentences must have the same referent and all false sentences must have the same referent. See (Church 1956) for a classic example.

<sup>9</sup>In fairness to Frege, and as Tyler Burge points out (1986, pg. 86 in the 2005 reprint), ‘reference’ is really a term of art for Frege. As Burge puts it, “ ‘The denotation’ of a sentence is whatever is most fruitfully seen as functionally dependent on the denotations of its parts.” (loc. cit.) Still, Frege’s case would be strengthened by an argument that sentences have referents even under an intuitive notion of “reference”.

playing in our cognitive lives?

Recall in the last chapter I argued that the representational content of a perception of an object as having a particular geometric shape can be elegantly modeled as a space of possible perceptions of that object from various angles. This is partially because my perception of an object as a cube (for example) carries commitments to how that object would appear from various other positions in space. That perception is not consistent with a perception of that same object as a sphere from a different angle. This is part of what it is for my perception to be an *objective* representation, a representation of something as existing independently of my perspective on it.

Note that this space of possible perceptions has quite a bit of structure to it. For example, by perceiving an object as a cube I am committed to its appearance shifting in a determinate way as I shift my position in space. For this reason, I can *expect* that as I move closer to the object, its appearance will increase in size (i.e. take up more of my visual field).

I claim that this is part of the cognitive value of representing something as an object: this gives me a way of linking various representations in accordance with particular rules, representations linked as being of the *same thing*. It is in virtue of perceiving the cube as an object that I can *expect* it to appear a particular way as I shift my position in space with respect to it. Otherwise, I'd have no reason to expect a current region of my visual field to have anything in common with a region of my visual field a few moments from now. My expectation is grounded on my representing a thing out in the world that I'm shifting my perspective on.<sup>10</sup>

---

<sup>10</sup>I owe this point to Immanuel Kant. For example: "Suppose that we inquire what new character

Object-representation also grounds my expectation that other perceivers will have (or are having) similar experiences to my own. If I wasn't perceiving that cube as a thing out in the world, then how could I expect someone else in the same room to have an experience at all similar to mine? If I found out that I was hallucinating, for example, I'd have no reason to expect you to have a similar experience to my own.

As another example, I'm currently perceiving my coffee cup. It is by virtue of representing it as an object that I can expect to have various types of experiences as I pick up the coffee cup, throw it against the wall, etc. I have a general idea of how much my coffee cup weighs and how much force it would take to break it. That is to say, object-representation also grounds my expectations that various potential actions of mine will have particular consequences.

Returning to the issue at hand, in what sense can an assertion be usefully thought of as referring to an object? Not only that, but given that the True is meant to be the *intended* referent of our assertions (it is the aim of judgment), in what sense can all of our assertions be usefully thought of as intending to refer to the *same* object?

When I make a sincere assertion, I am attempting to describe the world as it actually is. If I were instead reading a work of fiction aloud (and this was known

---

is given to our presentations by the *reference to an object*, and what is the dignity that they thereby obtain. We then find that this reference does nothing beyond making necessary the presentations' being combined in a certain way and being subjected to a rule; and we find, conversely, that only through the necessity of a certain order in the time relation of our presentations is objective signification conferred on them." (1781/1787, B242-3)

Kant, of course, would not agree with many aspects of my position, and I make no attempt at a general comparison here. As I mentioned in the last chapter, Haim Gaifman has pointed out to me that there are similar ideas in Bertrand Russell, e.g. his "The Relation of Sense-data to Physics" (1914). Note that Russell is attempting to *define* a physical thing as the class of its appearances (ibid pg. 129), while I am merely arguing that the *cognitive value* of representing an object is that it allows one to link various representations as of the same thing in accordance with particular rules.

to my audience), I wouldn't be attempting to describe the actual world. Similarly, an actor on a stage usually isn't attempting to describe the actual world with their utterances.

This suggests treating the True as the *actual world*: the totality of that which exists.<sup>11</sup> The value of seeing the True as the same intended referent across various assertions is that I am attempting to describe the very same world. It is by virtue of representing the *same* world over time that I can expect my future experiences to have anything in common with my past experiences, and furthermore that I can expect your experiences to be anything like my own.

For example, suppose I think to myself "In the past, the sun has always risen. Therefore, the sun will rise tomorrow." The fact (let's suppose) that this pattern of inference is generally truth-preserving<sup>12</sup> *presupposes* that the following day I experience the *same* world as I've been experiencing in the past. Otherwise, I'd have no reason to expect my experiences the following day to be anything like my past experiences. Seeing the world as lawlike presupposes that you're continuing to see the same world.

So, just like the more general case of objects, truth-preserving transitions between thoughts will be grounded on rules that link those thoughts, and these links are themselves grounded on treating those true thoughts as having the *same* referent. That is to say, truth-preserving inference between thoughts maintains fixed reference to the actual world while varying its representation in accordance with particular

---

<sup>11</sup>I am grateful to Anubav Vasudevan and Sidney Felder for first suggesting to me the idea of treating the True as the actual world.

<sup>12</sup>This inference is not, of course, guaranteed to be truth-preserving.

rules. In making such an inference, one is continuing to describe the world veridically, and the rules governing such an inference are grounded on its being the *same* world that one is continuing to represent.<sup>13</sup>

Note that Frege himself seemed to have held this view when he first introduced the True as an object. Consider the following quotation from “On Sense and Reference”:

Judgements can be regarded as advances from a thought to a truth-value. Naturally this cannot be a definition. Judgement is something quite peculiar and incomparable. One might also say that judgements are distinctions of parts within truth-values. Such distinction occurs by a return to the thought. To every sense attaching to a truth-value would correspond its own manner of analysis. However, I have here used the word ‘part’ in a special sense. I have in fact transferred the relation between the parts and the whole of the sentence to its [referent], by calling the [referent] of a word part of the [referent] of a sentence, if the word itself is a part of the sentence. This way of speaking can certainly be attacked, because the total [referent] and one part of it do not suffice to determine the remainder, and because the word ‘part’ is already used of bodies in another sense. A special term would need to be invented. (1892a, pg. 35-6/164-5)

If the referent of a word in a true sentence is always a part of the True, then the True includes as parts the totality of that which can be referred to. While Frege does say that this is a special use of the word ‘part’ (‘Teil’), the sense in which he says that it is special is that a particular entity and the True as a whole do not suffice to determine the remainder of the True, and furthermore that we already think of bodies as having parts in a different sense than a particular entity can be considered

---

<sup>13</sup>Note that this would allow me to explain why inferences like the following are unsound: “Sir Arthur Conan Doyle lived in England in the 19th century. Sherlock Holmes lived in England in the 19th century. Therefore both Sir Arthur Conan Doyle and Sherlock Holmes lived in England in the 19th century.” This inference is unsound because the first premise intends to describe the actual world while the second premise does not (it might be said to describe a fictional world), and hence reference is not being preserved from premises to conclusion.

a part of the True. Neither point undermines this conception of the True, and the fact that Frege used the word ‘part’ at all remains suggestive.

But didn’t Frege define the True in the *Grundgesetze* merely two years later as a particular abstract object, namely, as the extension of his horizontal function,<sup>14</sup> and furthermore didn’t he say that this definition was nearly arbitrary? This is a common misconception. In the *Grundgesetze* Frege actually treats the True *as primitive* and defines functions such as the horizontal in terms of it (the horizontal maps the True to the True and anything else to the False). The discussion of the extension of the horizontal function is in the context of explaining what *extensions* (and value-ranges more generally) are supposed to be. He is stipulating the definition of that extension as the True rather than stipulating the definition of the True as that extension. His point is that there are other extensions that we could have instead defined as the True.

This is especially clear in a footnote near the end of that section, in which Frege considers (and ultimately rejects) defining *any* object  $\Delta$  as the extension of the concept “being identical to  $\Delta$ ”. Frege says, “Such a stipulation is possible for every object that is given to us independently of value-ranges, for the same reason that we have seen for truth-values.” (1893, pg. 18 fn. 1) Hence, for Frege the truth-values are given independently of value-ranges and are being used to define a pair of value-ranges rather than vice versa.

Still, if Frege did hold a conception of the True as something like the actual world,

---

<sup>14</sup>Recall that an *extension* is the value-range (or course-of-values) of a function that maps all of its arguments to truth-values. Frege calls such a function a *concept* (1893, pg. 8), although I will not follow this terminology.

he seems to have given it up in his later work. For example, in “Thoughts” Frege says, “So it seems likely that the content of the word ‘true’ is *sui generis* and indefinable.” (1918-19a, pg. 60/353)<sup>15</sup>

What about the False? Given that the True is the intended referent of assertion, the False corresponds to a particular sort of reference failure. In Frege’s system, for example, an assertion will refer to the False when each of its significant components succeeds in referring (names refer to objects, predicates refer to properties, etc.) but the assertion fails to describe the way those referents are related in the actual world (e.g. that object doesn’t actually instantiate that property). Hence, we can think of the False as an abstract object introduced in order to capture this sort of failure to refer to the actual world. Given that ‘John’ refers to John but John doesn’t actually have brown hair, it isn’t actually necessary to treat ‘John has brown hair’ as having a referent, although doing so can allow for a more elegant formal system. In particular, it allows Frege to maintain his thesis that the reference of a complete expression (such as a declarative sentence) depends on the reference of its significant parts.<sup>16</sup>

---

<sup>15</sup>Frege’s *argument* for this conclusion fails. Frege claims that any attempt to define truth will have to specify certain characteristics, but in application the question would always arise whether it were *true* that those characteristics were present, and hence we would be going around in a circle. (loc. cit.) But the question would be whether those characteristics were present. If they were, then we would have a truth, and otherwise we wouldn’t. It would also be *true* that those characteristics were present, but we wouldn’t need to establish this before we could establish that those characteristics were present. It would also be true that it was true that those characteristics were present, but this wouldn’t need to be established first either. The order of explanation goes in the other direction.

Perhaps Frege’s point could be better made by insisting that we must *presuppose* a notion of truth in order for judgment to make sense. But this wouldn’t imply that truth is *sui generis* and indefinable. Our presupposed notion of truth might have a determinate structure that we can uncover; it would merely follow that our judgments about that structure themselves presuppose it.

<sup>16</sup>Unlike Frege, I’m happy to let only true thoughts have referents, strictly speaking. This would lead to a more complex system, since (for example) a true thought could then include a thought as a component that lacks a referent. The value of introducing the False in a formal system is obvious.



Hence, there are reasons for accepting Frege's position that all assertions have a special object, the True, as their intended referent. By treating the True as the actual world, we can explain what all sincere assertions have in common. Furthermore, our assertions have the *same* intended referent because we're trying to describe the *same* world over time, as is presupposed by our expectation that our future experiences will have anything in common with our past experiences or that your experiences will have anything in common with mine.

### 4.3 Atomic Form

In the last section I argued that the cognitive value of our representing something as an object is that it allows us to link various representations as being of the same object, and in particular to link them in accordance with particular rules. This grounds my expectations regarding my future experiences, and my expectations that your experiences will be (or are) similar to mine.

The ability to link representations in this way is a necessary precondition of rational agency. If I couldn't link my representations with representations of expected future states of an entity or system in accordance with rules, I wouldn't be able to make a rational choice between options. I wouldn't be able to reason about the consequences of choosing one option rather than another. For example, my reasoning about whether I'd prefer the consequences of pressing button A or button B on a device presupposes an ability to *predict* the consequences of each choice (even if such prediction is merely probabilistic). One could of course also reason about differing

states of a system, but note that in doing so one is treating the system itself as an object.

Being able to *learn* from experience puts further constraints on the structure of experience and thought. In order to learn from experience, I not only need to represent particular things; I also need to represent *patterns* or *structures* that multiple things can have in common. For example, in order to learn that strawberries are delicious, I need to represent particular strawberries as instances of a more general pattern that I can then recognize other particular things as instances of (in this case let's summarize this pattern as "being a strawberry"). If I didn't represent a particular strawberry as falling under a general, repeatable pattern, I would have no reason to expect a distinct strawberry to also taste delicious. That is to say, I wouldn't recognize the next strawberry as having anything in common with the previous strawberry, and hence would have no reason to conclude that it was probably also delicious.

Note that I'm *not* saying that this sort of knowledge (e.g. the deliciousness of strawberries) couldn't be "hard-wired" into the mind; I'm instead saying that *if* such information is learned, *then* it is learned by virtue of representing particular things as instantiating patterns. In fact, even if such information were "hard-wired", in order to *apply* that information I'd need to represent a particular thing as an instance of the pattern that my information is about.

Similarly, being able to learn how systems of multiple objects are going to behave depends upon being able to represent multiple objects as instantiating relational patterns. For example, my *learning* how certain pairs of objects will behave after colliding depends upon my representing those objects as instances of a relational

pattern that I can then represent other pairs of objects as also instantiating. Consider the classic case of learning to predict the behavior of billiard balls based on their relative motion.

The point holds generally: in order to learn from experience, an agent must make something like an object/property/relation distinction. Note that it follows that this distinction is *unlearnable*: this distinction itself cannot be learned because any learning already presupposes it. That is, this distinction is *a priori* in the Kantian sense of arising from the mind itself.<sup>17</sup>

Is it possible for there to be a rational agent who doesn't learn from experience, and furthermore doesn't need to make an object/property/relation distinction at all? Note that such an "agent" would be unable to apply knowledge *to* experience either, since this would also require recognizing particulars as instances of general patterns, namely, the patterns of which the agent has knowledge. Note that both points apply to the agent's own "inner" experiences as well.

Furthermore, rational agency requires inferential transitions between thoughts governed by particular rules. Even a "rational agent" bearing no rational connection to their own experiences would have to treat thoughts as *structured* in order to make rule-governed inferential transitions between them. Such rules apply at the level of general patterns instantiated by particular thoughts (e.g. if a thought instantiating pattern A is true, then so is the corresponding thought instantiating pattern B).

---

<sup>17</sup>Of course, this distinction could still have arisen as the result of natural selection stemming from interaction with an external world. Furthermore, interaction with the external world (and particular sorts of experiences) may be necessary for a child's mind to develop in such a way that this distinction is made. Compare Kant: "But even though all our cognition starts *with* experience, that does not mean that all of it arises *from* experience." (1781/1787, B1)

But treating thoughts as structured in this way is making a distinction between particulars and patterns instantiated by those particulars after all. So, something like the object/property distinction is necessary for the possibility of rational agency.<sup>18</sup>

Let's consider a potential counterexample.<sup>19</sup> The lambda calculus doesn't make a distinction between terms and predicates, but rather treats every basic syntactic unit as a term. One might think that one could learn to reason with the lambda calculus in such a way that one's reasoning isn't relying on a distinction between particulars and patterns instantiated by those particulars. More generally, one might learn to reason with a radically nonstandard logic through experience with applying particular syntactic rules for transitioning between expressions.

In response, note the distinction between reasoning *within* a formal system and reasoning *about* a formal system. Even if the formal system I've adopted doesn't make something like a term/predicate distinction, *I* may still need to make such a distinction when I reason *about* the expressions of that system. For in order to apply a rule for transitioning between expressions, I need to recognize those expressions as instances of general expression-*types* that the rule can be applied in virtue of. And hence I'm making something like an object/property distinction at the meta-level after all.

Even if I am able to reason *within* a formal system rather than merely *about* it (and note that it isn't at all obvious that one can in fact do so with a highly nonstandard

---

<sup>18</sup>Analogous points apply to relations. For example, one needs to *relate* thoughts as premises and conclusion in order to make an inferential transition between them. The relevant relation is specified by the particular inference rule.

<sup>19</sup>I am grateful to Haim Gaifman for pressing me to address this example.

formal system), in order to apply inference rules I still need to distinguish particular contents from general *patterns* or *types* those contents instantiate.<sup>20</sup> Once again, an inference rule must be applied to particular contents *in virtue of* patterns those contents instantiate. Hence, something like an object/property/relation distinction is unavoidable in reasoning.<sup>21</sup>

The point being, when reasoning *about* a formal system I must actually *represent* something like an object/property/relation distinction, and when reasoning *within* a formal system I must make inferential transitions *in virtue of* that distinction, and hence in either case my mind must make the distinction.

I believe that we should conclude that this distinction is a *logical* distinction, and the corresponding atomic form of a thought is a *logical* form. Given the centrality of this distinction to reasoning, it should be unsurprising that thoughts are structured in such a way that allows one to distinguish between and maintain fixed reference to various objects, properties, and relations while making inferential transitions. In effect, this distinction corresponds to a structure that one would have to put in place if one were designing a rational mind, and in that sense can be thought of as fundamental to the structure of thought. I think that such fundamental structure deserves to be called logical structure.<sup>22</sup>

---

<sup>20</sup>Here I use ‘content’ as a generic expression meant to capture whatever corresponds at the level of actual reasoning to the primary syntactic units (e.g. sentences) of the formal system.

<sup>21</sup>This argument may remind the reader of Quine’s classic argument in “Truth by Convention” (1936) that the truths of logic cannot be fixed by convention, since one will need general conventions to apply to the infinite number of logical truths, and logic will be needed to derive particular conclusions from the general conventions, which would lead to an infinite regress. Many of my arguments below are also in this spirit.

<sup>22</sup>Note that I am not attempting to *define* logical form in terms of rationality. It isn’t clear that one could do so without circularity. I am grateful to Justin Clarke-Doane for emphasizing this point.

First-order languages often include a special logical relation-symbol interpreted as *identity*. Can we give a similar argument to justify its standard treatment as a logical relation? Note that there is an extensive literature on whether we ought to treat identity as a logical relation, and I cannot do justice to it here.<sup>23</sup> Arguing for my position in detail would require another dissertation. Here I can only make the first moves in what deserves a much longer discussion.

I argued above that linking distinct representations as of the same object is necessary for the possibility of rational agency. Could it be learned from experience? That is, could experience teach you to recognize an object as the *same* again if you had no conception of sameness?

I claim that it could not. I argued above that a distinction between particulars and patterns instantiable by multiple particulars is necessary to learn from experience. I claim that representing a particular requires linking that representation with *possible* representations of the same particular. In effect, in order to represent a particular one must locate one's representation within a space of possible representations structured in accordance with the properties one is representing that particular as instantiating. Compare my example of perceiving a cube: a particular perception of a cube includes commitments regarding how that cube would appear from various other positions in space (including the positions of distinct observers), and for this reason I argued in the previous chapter that the content of such a perception (at the level of abstraction

---

<sup>23</sup>Two classic positions that I will not be addressing are Wittgenstein's attempt to dispense with identity in the *Tractatus* (1922) and Peter Geach's notion of relative identity (e.g. (Geach 1972)). In the case of Wittgenstein, there is no symbol for identity in his language because of his metaphysical conception of objects. The natural way of formalizing Geach's position is to use a many-sorted language in which the identity symbol is relativized to predicates.

we're considering) is best treated as a space of possible perceptions of that cube. I claim that the same structure applies to object-representation more generally.

Given that representing a particular requires locating that representation within a space of possible representations of that particular, and given that identity corresponds to linking these possible representations as of the same particular, representation of a particular is impossible without possessing a primitive conception of identity. That is to say, one can't represent an object without having a primitive conception of identity, a conception corresponding to an ability to link representations as of the same object.<sup>24</sup>

Hence, the atomic forms we've considered correspond to distinctions and concepts that are both unlearnable and necessary for the possibility of rational agency. In the next section we'll consider whether the classical logical connectives also satisfy these two conditions.<sup>25</sup>

---

<sup>24</sup>The reader may disagree. It should be clear that the more central point for my purposes is that structural features of thoughts that correspond to distinctions and conceptions both necessary for the possibility of rational agency and unlearnable have a strong claim to be considered logical forms. Personally, I think that if a primitive conception of identity could in fact be learned from experience, then it shouldn't count as a logical relation.

<sup>25</sup>First-order languages often include *function-symbols* among the possible constituents of atomic formulas. While I do believe that the connection between functions and rational agency is worth pursuing (I believe it hinges on the possibility of acting in accordance with rules governing the transition from an initial state to a goal state, although this is insufficient to capture the general notion of a function), for now I'll merely point out that a function is a special sort of relation, and hence an atomic formula such as ' $R(f(a), b)$ ' can be translated as "the unique object bearing the  $f$ -relation to object  $a$  bears the  $R$ -relation to object  $b$ ." Uniqueness can then be captured via identity, conjunction, negation, and universal quantification, the latter three of which we'll consider in the next section.

## 4.4 Logical Connectives

In this section I'll consider various arguments that the classical first-order logical connectives, like the object/property/relation distinction, are unlearnable and necessary for the possibility of rational agency. I do not find all of these arguments compelling, especially the arguments concerning the material conditional and negation. In particular, many of the arguments we'll consider are quite controversial; considering them in full would take another dissertation. For example, we lack the space to discuss logical pluralism (the view that there is more than one correct logic, see (Beall and Restall 2006)), although the reader may have surmised that our focus on structurally-guaranteed truth-preservation and the fundamental structure of a rational mind rules out many versions of the view.

The point of this section is not to convince the reader that the classical first-order logical connectives do in fact correspond to logical structure under the conception of logic I have been working with throughout this chapter; it is instead to suggest how one might begin arguing for such a conclusion in order to give the reader some idea of the value of this conception of logic.<sup>26</sup>

Lewis Carroll, in his well-known paper “What the Tortoise said to Achilles” (1895), constructed a dialogue in which Achilles attempts to force the tortoise to accept that  $Z$  is true given that he accepts that  $A$ ,  $B$ , and  $C$  are true, where  $C$  is: if  $A$  and  $B$  are true then  $Z$  must be true. The tortoise responds that he could reject the hypothetical

---

<sup>26</sup>In particular, I don't believe that the arguments of this section would convince either the anti-realist or the philosopher who rejects the link between the possession-condition of a concept and an ability to apply it to relevant cases, among others. As we will see below, the overall argument we're considering may also hinge on viewing standard connectives as more “primitive” than non-standard counterparts.



*D*: if *A* and *B* and *C* are true then *Z* must be true. Achilles then asks him to add *D* as a premise, and the whole process repeats indefinitely.

Some philosophers have argued that this dialogue shows that the inference rule *Modus Ponens* is *unlearnable*.<sup>27</sup> However, it isn't clear what *learnability* means when talking about inference rules. Beginning to make *Modus Ponens* inferences due to divine intervention or brain surgery doesn't seem to count as *learning* it, but what about being embedded in a community that tends to make assertions following a pattern associated with *Modus Ponens*, and adopting a practice of making assertions in accordance with that pattern?

I will focus on a more specific claim: one cannot come to recognize *Modus Ponens* as truth-preserving and make inferences governed by it that are ultimately based on that recognition, since in order to do so one would already need to be making inferences governed by *Modus Ponens*.<sup>28</sup> Note that I am not insisting that this recognition be conscious; one may come to recognize an inference rule as truth-preserving subpersonally.

Anyway, if one didn't already implicitly accept the pattern of inferring '*q*' from '*p*' and 'if *p* then *q*' as truth-preserving then one certainly couldn't come to recognize that pattern as truth-preserving by coming to recognize that 'if *p* and if *p* then *q* then *q*' is always true, since adding this as an additional premise wouldn't help you to actually draw the conclusion that '*q*' is true under the given assumptions, an inference which

---

<sup>27</sup>This is Saul Kripke's diagnosis. See (Padro 2015) for discussion. I don't think this is a proper interpretation of Carroll, although it may still be a reasonable conclusion to draw from the dialogue.

<sup>28</sup>This is in fact my suggestion for how we should understand *learning* an inference rule, although I will avoid speaking of learning inference rules in the rest of this section and will instead focus on the more specific claim.

would still require you to apply the very rule that you are supposed to be coming to appreciate as truth-preserving.<sup>29</sup>

Is there some other method for coming to recognize that *Modus Ponens* is truth-preserving, given that recognizing that ‘if  $p$  and if  $p$  then  $q$  then  $q$ ’ is always true won’t work? For example, one might think that one could come to accept that every instance of *Modus Ponens* is truth-preserving in a way that needn’t involve accepting a conditional of that form as always true.

But note that *applying* this knowledge would require an application of Universal Instantiation (i.e. inferring that a *particular* instance of *Modus Ponens* is truth-preserving). And it is difficult to see how Universal Instantiation could be applied in this case without resulting in a conditional of the form “if this transition is an instance of *Modus Ponens*, then it is truth-preserving”, the application of which would require an application of *Modus Ponens* after all.<sup>30</sup> Furthermore, applying the knowledge that *Modus Ponens* is truth-preserving would require a transition from a meta-level representation of *Modus Ponens* as truth-preserving to an object-level inference governed by *Modus Ponens*.

The reader may feel that these sorts of transitions could be made without any use

---

<sup>29</sup>Note that I said “*implicitly* accept”. In actual practice an individual may first explicitly consider whether this inference rule is truth-preserving while studying a truth-table in an Introduction to Logic course. The point is that in reasoning through the truth-table they are applying this very inference rule and implicitly presupposing that it is truth-preserving. See (Peacocke 2008, Chapter 4) for an extended discussion of implicit conceptions.

The point that one cannot give a reduction of the classical logical connectives via truth-tables due to circularity is well-known. However, the truth-tables can still be *illuminating*. Similar remarks may apply to the correspondence theory of truth. I owe these latter points to Haim Gaifman.

<sup>30</sup>And note that *that* application of *Modus Ponens* couldn’t be recognized as truth-preserving on pain of infinite regress. The point being, one couldn’t recognize an instance of *Modus Ponens* as truth-preserving without already presupposing it.

of *Modus Ponens*. I conjecture that *any* attempted recognition of truth-preservation and ultimate adoption based on that recognition would involve coming to recognize that *if* something is true, *then* something else is true, and hence to rely on an application of *Modus Ponens* after all. Note that this point isn't essential to the overall argument we're considering in this section, which aims in part to establish that the structure of thoughts corresponding to the material conditional deserves to be called logical form. Given that the material conditional is definable solely in terms of conjunction and negation, this conclusion would follow from the soundness of the arguments below.

Turning to the material conditional, I claim that one couldn't possess a notion of the material conditional without treating *Modus Ponens* inferences as truth-preserving. So, the claim is that it is essential to a possessing a notion of the material conditional that the agent treats a transition to a thought that  $q$  from a thought that  $p$  and a thought that if  $p$  then  $q$  as a potential truth-preserving inference.<sup>31</sup> A notion of the material conditional is essentially tied to a particular sort of inference, and hence if one is unable to make the corresponding inference, then one is either failing to grasp the truth-conditions associated with a material conditional (and thus actually working with a distinct notion) or somehow failing to link truth-conditional information to potential inferences. In either case one fails to possess the notion.<sup>32</sup>

---

<sup>31</sup>More carefully: the agent's mind presents the inference as a potential truth-preserving inference. If the agent is in the grip of a particular philosophical theory, they may be unwilling to actually draw the inference. Compare: an agent is enjoying a perception of an object, a perception that presents that object as instantiating various properties, but the agent chooses not to take her perception at face-value, perhaps due to background information or philosophical theory.

<sup>32</sup>So, it is essential to possessing a notion of the material conditional that one be able to make the relevant inferences with it. Compare: it is essential to possessing a particular concept that one

Given that one couldn't come to recognize that *Modus Ponens* is truth-preserving (since one would need to already be presupposing this), and given that possessing a notion of the material conditional essentially involves treating *Modus Ponens* inferences as truth-preserving, it isn't clear that a notion of the material conditional is the sort of thing that could be *acquired*, since it seems that in order to do so one would need to come to recognize that *Modus Ponens* is truth-preserving.<sup>33</sup> <sup>34</sup> Is it necessary for the possibility of rational agency?

Rational agency essentially involves being able to reason. And reasoning essentially involves drawing conclusions from premises. Drawing conclusions from premises essentially involves treating the truth of the conclusion as a consequence of the truth of the premises, i.e. treating them as instantiating a consequence relation. Hence, one could draw the desired conclusion by arguing that the most primitive consequence

---

be able to apply it to the relevant cases.

Note that I am *not* arguing that what makes the concept of the material conditional the concept it is is its role in particular sorts of inferences (i.e. I'm not giving individuation-conditions); I am instead arguing that in order to *possess* a notion of the material conditional one must be able to make the relevant inferences with it. So, this account is perfectly compatible with e.g. Peacocke's (2008) treatment of concepts as individuated by their fundamental reference rules, as we discussed in Chapter 2.

<sup>33</sup>Note that this argument does *not* generalize to all concepts whose possession-conditions essentially involve treating particular inference rules as truth-preserving. One could acquire both the concept and a practice of using the inference rule without in any sense needing to acquire one before the other. This argument essentially depends on being unable to recognize the inference rule as truth-preserving without already making inferences governed by that inference rule.

One might alternatively argue along the lines of the following paragraph in the main text that one couldn't acquire a notion of a new consequence relation without relying on one's antecedent understanding of some other consequence relation, and one couldn't understand any consequence relation at all unless one already understood what is arguably the most primitive consequence relation: the material conditional. I am hesitant to endorse this argument, but I state it here for the reader's consideration.

<sup>34</sup>An intuitionist logician, for example, needn't disagree. The point is that if one possesses a notion of the (classical) material conditional, then one treats *Modus Ponens* inferences as truth-preserving. The point is *not* that if one treats *Modus Ponens* inferences as truth-preserving, then one possesses a notion of the (classical) material conditional.

relation is the material conditional. The idea would be that an ability to draw conclusions from premises presupposes an ability to relate those premises to the conclusion via a material conditional.<sup>35</sup>

In support of this, note that the truth of a material conditional merely guarantees that a transition from the antecedent as premise to the consequent as conclusion will not in fact be a transition from a truth to a falsity (i.e. its truth leaves you unable to determine which of the additional three rows of its truth-table is correct). So the claim is that a rational agent must be able to relate premise(s) and conclusion in this way. While a rational agent may grasp a more demanding consequence relation (e.g. where the premise(s) and conclusion must be *related* in some appropriate way), the claim is that one cannot grasp a more demanding consequence relation without also grasping this primitive consequence relation.<sup>36</sup> Again, note that since the material conditional is definable solely in terms of conjunction and negation, even if the reader disagrees with this point they may still be committed to accepting the material conditional as a logical connective due to the arguments below.

Turning now to conjunction, here is one way to argue that a basic ability corresponding to conjunction is necessary for the possibility of rational agency: a rational agent needs to be able to combine multiple representations of the same entity together into a united representation. Recall that I argued above that true thoughts refer to

---

<sup>35</sup>Such an argument may remind the reader of the Deduction Theorem. Note that I am *not* collapsing the important distinction between a logical connective and a logical consequence relation. I am instead arguing that relating a premise and a conclusion via a material conditional is treating them as instantiating a weak (and non-logical) consequence relation, as I explain further in the main text below.

<sup>36</sup>An intuitionist, of course, will disagree. As I stated above, I myself do not find this argument compelling. Much more would need to be said in order to resolve such a dispute.

the True: the actual world. Hence, conjunction corresponds to an ability to combine multiple representations of the actual world into a larger picture of the world. A rational agent must be able to link the conclusions they draw with the premises they drew them from into a larger picture of the world.<sup>37</sup>

Could a notion of conjunction be acquired? Possessing a notion of conjunction essentially involves being able to infer the truth of ‘ $p$  and  $q$ ’ from the truth of ‘ $p$ ’ and the truth of ‘ $q$ ’. Consider an agent who doesn’t already possess a notion of conjunction. Imagine trying to teach them when the notion ought to be applied.

Lewis: “Do you accept that  $p$  is true?” Saul: “Yes.” Lewis: “Do you accept that  $q$  is true?” Saul: “Yes.” Lewis: “So this is a case in which you should also accept that  $p$  and  $q$  is true.” Saul: “Wait, what? What sort of case is it?” Lewis: “A case in which you both accept that  $p$  is true and you accept that  $q$  is true!” Saul: “Wait, *what* sort of case?”

The point being, you couldn’t learn to recognize a case in which to apply a notion of conjunction without already possessing it, and it is essential to possessing such a notion that one be capable of applying it to cases. These considerations suggest that a notion of conjunction both cannot be acquired from experience and is necessary for the possibility of rational agency.

What about negation? Like conjunction, it isn’t clear that one could acquire a notion of negation because it seems that one couldn’t learn how to recognize a case in which to apply it without already possessing it (I encourage the reader to construct

---

<sup>37</sup>Alternatively, one might argue that rational agents must be able to draw a conclusion from multiple premises, and furthermore they must be able to deliberate about the consequences of multiple conditions holding.

the analogous dialogue: “What do you mean by *false*?”). However, whether it is necessary for the possibility of rational agency seems more controversial.

Here is one argument: some conditions hold and some conditions fail to hold. Indeed, things couldn't have been otherwise. A rational agent must be capable of what I'll call *binary deliberation*: reasoning about both the consequences of some condition holding and the consequences of that condition failing to hold. An obvious case of this is reasoning about both the consequences of performing an action and the consequences of failing to perform that action. This ability presupposes a notion of negation.

Alternatively, one might argue that one cannot grasp a thought without being capable of grasping its negation, since one can't recognize the conditions under which a sentence would be true without also being capable of recognizing the conditions under which that sentence would be false, and recognizing the truth-conditions of a complete sentence is the paradigm case of grasping a thought: namely, the thought the sentence expresses.<sup>38</sup>

However, one might think that one could possess these abilities without possessing the *classical* notion of negation, but rather by possessing some non-classical notion (e.g. *intuitionistic* negation). So, as in the case of the material conditional, this argument may hinge on the classical notion of negation being the most *primitive* notion in the sense that already possessing the classical notion is necessary for possessing

---

<sup>38</sup>Note that this latter argument is controversial: one may think that one could understand a sentence such as ‘That object is green’ without having any understanding its negation. Personally, I don't think one can possess a concept of a color without possessing concepts of other colors as well and recognizing that nothing can be entirely one color and entirely a distinct color at the same time.

any other notion of negation.

I won't consider the other standard propositional connectives. Note that it is a well-known result that *any* truth-functional connective can be defined in terms of conjunction and negation.<sup>39</sup> Hence, provided one accepts that any concept definable solely in terms of logical concepts is itself a logical concept, one will be committed to accepting the other standard propositional connectives as logical connectives as well.<sup>40</sup>

Turning finally to the universal quantifier, one couldn't possess a corresponding notion without being able to make inferences governed by Universal Instantiation. Could one come to recognize Universal Instantiation as truth-preserving? Consider the following argument due to Kripke: Suppose an agent doesn't already recognize Universal Instantiation inferences as truth-preserving. They then decide to accept on your authority as a logician that every instance of Universal Instantiation is truth-

---

<sup>39</sup>To see why, note that one can define the disjunction ' $p \vee q$ ' as ' $\neg(\neg p \wedge \neg q)$ '. Now, consider the truth-table of an arbitrary truth-functional connective. For any row in which the application of the connective results in a formula with the truth-value True, construct a corresponding conjunction consisting of the subformulas specified by the truth-table as taking the truth-value True in that row and the negations of the subformulas specified by the truth-table as taking the truth-value False in that row. After considering every row and constructing the relevant conjunctions, construct a formula consisting of the disjunction of all the constructed conjunctions. The resulting formula will be true in circumstances in which the application of the connective results in a formula with the truth-value True and false otherwise.

For all technical details I refer the reader to (Ebbinghaus et. al. 1994). Note that this sketch of a proof assumes that the semantics for negation correspond to the classical truth-table. While there is clearly room for other notions of negation, the claim would be (as I stated in the previous paragraph in the main text): a rational agent must at least *possess* the classical notion (compare the suggestion that the material conditional is a *primitive* consequence relation discussed above). As I stated at the outset of this section, these arguments will not convince e.g. the anti-realist. We lack the time or space for an extended discussion of the relevant issues.

<sup>40</sup>One might think that combining logical concepts into a new concept counts as learning, and furthermore isn't necessary for rational agency. Note that one can still recover all of first-order logic solely via conjunction, negation, and universal quantification. Furthermore, the extension of the account of logical form to concepts definable solely in terms of the concepts discussed is quite natural.



preserving. However, when confronted with a particular case of Universal Instantiation, they remain skeptical.

The point being, drawing the conclusion that this particular instance of Universal Instantiation is truth-preserving from the premise that every instance of Universal Instantiation is truth-preserving is itself an application of Universal Instantiation.<sup>41</sup> Hence, one couldn't come to recognize that Universal Instantiation is truth-preserving, since one would already need to presuppose this in one's own inferential practice. But then it isn't clear that one could *acquire* a notion of the universal quantifier, possession of which essentially involves treating Universal Instantiation as truth-preserving, and hence acquisition of which would seem to require coming to recognize Universal Instantiation as truth-preserving.

Is Universal Instantiation necessary for the possibility of rational agency? Note that applying any rule whatsoever involves proceeding from the general case to a particular case: e.g. this pattern is subsumed under that one, hence this particular instance of the former pattern is also an instance of the latter pattern. Part of the nature of a rule is to be generally applicable to a variety of cases (i.e. *all* cases of a particular type). Hence, *applying* a rule will involve proceeding from the general case to the particular case. But this seems to be the essence of Universal Instantiation.

Put another way, it is essential to rational agency that one be able to draw conclusions about a particular case based upon a general rule. Drawing such a conclusion involves treating the particular case as instantiating a pattern specified by the general rule, and making an inference about the particular case due to its instantiating that

---

<sup>41</sup>See (Padro 2015) for an extended discussion of Kripke's argument.

pattern. Hence, it seems that this inference will crucially hinge on an application of Universal Instantiation. Since rational agency essentially involves applying rules to particular cases, and since Universal Instantiation essentially involves the universal quantifier, the universal quantifier is necessary for the possibility of rational agency as well.

The reader may feel that an ability corresponding to the application of some other quantifier could be used to apply rules to particular cases. I suspect that like the case of the material conditional above, the overall argument we're considering in this section may hinge on the universal quantifier being a *primitive* quantifier in the sense that possessing a notion of it is necessary for possessing a notion of any quantifier at all. Personally, I find this line of argument in the case of the universal quantifier to be quite compelling, but am less convinced by the claims of primitiveness in the cases of the material conditional and negation.

So, one can bring forth various considerations in favor of the standard connectives of classical first-order logic being either both unlearnable and necessary for the possibility of rational agency or definable solely in terms of connectives satisfying those two conditions. If one accepts these considerations, I believe that one should conclude that the corresponding structural features of thoughts are *logical* forms, and inferences hinging solely on these structural features are *deductive* inferences.

## 4.5 Conclusion

We have thus considered various reasons why structural features of thoughts corresponding to the sentences of a standard first-order language should be considered *logical* forms, and hence inferences governed by rules whose truth-preservation can be explained solely in terms of these structural features ought to be considered *deductive* inferences. Are there other structural features of thoughts that are also unlearnable and necessary for the possibility of rational agency?

In Chapter 2 I argued that a distinction between representations and meta-representations *of* those representations is necessary for the possibility of reflective judgment (e.g. a judgment that reality is as it appears to be). However, according to the account developed in this chapter, it doesn't follow that the structure corresponding to the hierarchy of senses is *logical* structure. This is because a capacity for reflective judgment is not a necessary condition of rational agency. One could imagine a rational agent who makes rational transitions between thoughts without being capable of *representing* those thoughts as distinct from the objects, properties, and relations the thoughts are about. Such an agent would have no need to *compare* a representation of something with the thing represented, but would instead merely *transition* between representations in rule-governed ways. Indeed, small children and various animals may think in this way.<sup>42</sup>

Still, reflective judgment is a central aspect of the sort of reasoning that we engage in, and hence the structure corresponding to the hierarchy of senses is a natural

---

<sup>42</sup>I owe this point to Tyler Burge, who may nonetheless disagree that such mental structure *is* required for *reflective* judgment. See Burge (2010) for a thorough treatment of these issues.

extension of logical structure. One would expect there to be various other natural extensions of logical structure as well.

# Conclusion

This dissertation has been focused on the sort of content or meaning relevant to determining whether what we say or think is true or false. I have argued that we ought to characterize such content (i.e. *thought*) at a level of abstraction at which various sentences with differing syntactic structure can express the same thought, and furthermore the same thought will instantiate many distinct logical forms. Grasping a thought in a particular way will make a variety of possible inferences salient while masking others. I have demonstrated that this was in fact Gottlob Frege's actual view and have argued for it primarily in three ways: (1) reflecting on cases, (2) showing how it resolves various puzzles concerning the hierarchy of senses, demonstratives, and indexicals, and (3) drawing connections to plausible views on the structure of consciousness, perception, judgment, objectivity, and truth.

Over the course of this dissertation I have drawn various conclusions about the structure of thought. I have argued that we can best characterize such structure using Frege's function/object distinction, where (for example) one can grasp the thought that the number two is prime as consisting of the result of applying the function-sense expressed by 'is prime' to the object-sense expressed by 'the number two' as

argument, where both this function-sense and this object-sense are parts of that very thought. As a particular case of this I have argued that every thought contains its own canonical sense as a part, a sense which is mapped by the function-sense expressed by the truth predicate to that complete thought. I have also shown that Frege himself was committed to this position on the structure of thought.

I have also argued that the senses expressed by spatial and temporal indexicals, in combination with places and times of utterance, express senses that themselves have the structure of spaces and timelines, respectively. I have shown that this elegantly captures one's commitments to the perspective-independence of the truth-conditions of one's assertions about the spatiotemporal world while also (due to multiple decompositions) capturing the contribution of one's perspective to one's way of thinking of that world. Again, I have also shown that Frege was committed to such a view.

I have also considered logical forms of thoughts in more detail, arguing that the sort of structural features of thoughts that must exist in order for rational agency to be possible and could not be acquired ought to be considered logical forms. In effect, these features correspond to the structure that one would have to put in place if one were designing a rational mind, and in that sense can be thought of as fundamental to the structure of thought.

As stated above, I have drawn connections to various other philosophical positions over the course of this dissertation. The independent plausibility of those positions lends further support to my outlined position on the structure of thought, although none of them are essential to my position.

Let me briefly mention some of those positions: (1) that part of the cognitive value

of being able to represent our own conscious experience as structured is that it aids in conscious decision-making leading to various actions (including mental actions such as judging or inferring), (2) that our capacity for meta-representation is linked to our capacity to reflectively judge whether reality is as it appears to be, (3) that the belief relation is in fact a three-place relation between a subject, a thought, and the world, (4) that one's perception of an object as having a particular geometric shape carries commitments to how that object would appear from various positions in space, (5) that conscious perceptions are used as representations of themselves in the context of conscious meta-level reflection on them, (6) that the cognitive value of representing something as an object is that it allows one to combine various representations as of the same object in accordance with particular rules, and (7) that something like an object/property/relation distinction is necessary for the possibility of rational agency.

Each of these positions warrants further discussion and I hope to address them elsewhere. Furthermore, much more could be said about the structure of thought. In chapter 2 I gesture at a formal system meant to elucidate the hierarchy of senses in accordance with the positions that I have adopted. I hope to develop such a formal system in the future. In chapter 4 I point in the direction of a general investigation of the logical forms of thoughts, an investigation which I plan to continue.

Self-reference has played a fundamental role in many parts of this dissertation. I believe that at present we do not have a good philosophical account of the nature of self-reference, although we are well-aware of the many paradoxes that arise in connection with it. I intend to contribute to this ongoing discussion.

While I have drawn many conclusions about the level of sense, I have said rela-

tively little about the level of reference or the nature of the relation between sense and reference. Doing justice to such a topic would require drawing quite general conclusions about the relation between mind and world. Here I have merely taken the first steps in what I expect to be a lifelong project.



# Bibliography

- [1] Awodey, Steve, Erich Reck, and Gottfried Gabriel (2004), *Frege's Lectures on Logic: Carnap's Student Notes, 1910-1914*, Chicago, IL: Open Court.
- [2] Beaney, Michael (1997), *The Frege Reader*, Malden, MA: Blackwell.
- [3] Beall J. and G. Restall (2006), *Logical Pluralism*, Oxford: Oxford University Press.
- [4] Bell, David (1996), "The Formation of Concepts and the Structure of Thoughts", *Philosophy and Phenomenological Research*, 56(3): 583-96.
- [5] Burge, Tyler (1979), "Sinning Against Frege", reprinted in: Burge (2005), pp. 213-39.
- [6] Burge, Tyler (1986), "Frege on Truth", reprinted in: Burge (2005), pp. 83-152.
- [7] Burge, Tyler (2005), *Truth, Thought, Reason: Essays on Frege*, Oxford: Oxford University Press.
- [8] Burge, Tyler (2009), "Five Theses on *De Re* States and Attitudes", in: J. Almog and P. Leonardi (eds.), *The Philosophy of David Kaplan*, Oxford: Oxford University Press, 2009, pp. 246-316.
- [9] Burge, Tyler (2010), *Origins of Objectivity*, Oxford: Oxford University Press.
- [10] Burge, Tyler (2011), "Self and Self-Understanding: The Dewey Lectures", reprinted in: Burge (2013), pp. 140-226.
- [11] Burge, Tyler (2012), "Living Wages of Sinn", reprinted in: Burge (2013), pp. 556-93.

- [12] Burge, Tyler (2013), *Cognition Through Understanding: Philosophical Essays, Volume 3*, Oxford: Oxford University Press.
- [13] Bermudez, Jose (2001), “Frege on Thoughts and their Structure”, *Logical Analysis and History of Philosophy*, 4: 87-105.
- [14] Carroll, Lewis (1895), “What the Tortoise said to Achilles”, *Mind*, 4 (14): 278-80.
- [15] Church, Alonzo (1956), *Introduction to Mathematical Logic*, Princeton: Princeton University Press.
- [16] Currie, Gregory (1985), “The Analysis of Thoughts”, *Australasian Journal of Philosophy*, 63(3): 283-98.
- [17] Davidson, Donald (1965), “Theories of Meaning and Learnable Languages”, reprinted in: Davidson (1984), pp. 3-15.
- [18] Davidson, Donald (1968), “On Saying That”, reprinted in: Davidson (1984), pp. 93-108.
- [19] Davidson, Donald (1984), *Inquiries into Truth and Interpretation*, Oxford: Oxford University Press.
- [20] Dummett, Michael (1981), *The Interpretation of Frege’s Philosophy*, Cambridge, MA: Harvard University Press.
- [21] Ebbinghaus, H.D., J. Flum, and W. Thomas (1994), *Mathematical Logic*, New York, NY: Springer.
- [22] Evans, Gareth (1981), “Understanding Demonstratives”, reprinted in: Evans (1985), pp. 291-321.
- [23] Evans, Gareth (1985), *Collected Papers*, Oxford: Clarendon Press.
- [24] Frege, Gottlob (1879), *Begriffsschrift*, translated by T. Bynum as “Conceptual Notation”, in: T. Bynum (ed.), *Conceptual Notation and Related Articles*, Oxford: Oxford University Press, 1972, pp. 101-203.
- [25] Frege, Gottlob (between 1879 and 1891), “Logic”, translated by P. Long and R. White, in: Frege (1979), pp. 1-8.

- [26] Frege, Gottlob (1880-81), “Boole’s Logical Calculus and the Concept-Script”, translated by P. Long and R. White, in: Frege (1979), pp. 9-46.
- [27] Frege, Gottlob (before 1884), “Dialogue with Pünjer on Existence”, translated by P. Long and R. White, in Frege (1979), pp. 53-67.
- [28] Frege, Gottlob (1884), *Die Grundlagen der Arithmetik*, translated by J. L. Austin as *The Foundations of Arithmetic*, Northwestern: Basil Blackwell, 1980.
- [29] Frege, Gottlob (1891), “Function and Concept”, translated by P. Geach, in: Frege (1984), pp. 137-56.
- [30] Frege, Gottlob (1892a), “On Sense and Reference”, translated by M. Black as “On Sense and Meaning”, in: Frege (1984), pp. 157-77.
- [31] Frege, Gottlob (1892b), “On Concept and Object (Preliminary Draft)”, translated by P. Long and R. White, in: Frege (1979), pp. 87-117.
- [32] Frege, Gottlob (1892c), “On Concept and Object”, translated by P. Geach, in: Frege (1984), pp. 182-94.
- [33] Frege, Gottlob (1892-95), “Comments on Sense and Reference”, translated by P. Long and R. White as “Comments on Sense and Meaning”, in: Frege (1979), pp. 118-25.
- [34] Frege, Gottlob (1893), *Grundgesetze der Arithmetik: Band I*, translated by P. Ebert and M. Rossberg as *Basic Laws of Arithmetic: Volume I*, Oxford: Oxford University Press, 2013.
- [35] Frege, Gottlob (1895), “A Critical Elucidation of Some Points in E. Schröder, *Vorlesungen über die Algebra der Logik*”, translated by P. Geach, in: Frege (1984), pp. 210-28.
- [36] Frege, Gottlob (1897), “Logic”, translated by P. Long and R. White, in: Frege (1979), pp. 126-51.
- [37] Frege, Gottlob (1903a), *Grundgesetze der Arithmetik: Band II*, translated by P. Ebert and M. Rossberg as *Basic Laws of Arithmetic: Volume II*, Oxford: Oxford University Press, 2013.

- [38] Frege, Gottlob (1903b), “On the Foundations of Geometry: First Series”, translated by E.-H. W. Kluge, in: Frege (1984), pp. 273-84.
- [39] Frege, Gottlob (1906a), “On Schoenflies: *Die Logischen Paradoxien der Mengenlehre*”, translated by P. Long and R. White, in: Frege (1979), pp. 176-83.
- [40] Frege, Gottlob (1906b), “Introduction to Logic”, translated by P. Long and R. White, in: Frege (1979), pp. 185-96.
- [41] Frege, Gottlob (1906c), “A Brief Survey of My Logical Doctrines”, translated by P. Long and R. White, in: Frege (1979), pp. 197-202.
- [42] Frege, Gottlob (1914), “Logic in Mathematics”, translated by P. Long and R. White, in: Frege (1979), pp. 203-50.
- [43] Frege, Gottlob (1915), “My Basic Logical Insights”, translated by P. Long and R. White, in: Frege (1979), pp. 251-2.
- [44] Frege, Gottlob (1918-19a), “Logical Investigations: Part I: Thoughts”, translated by P. Geach and R. H. Stoothoff, in: Frege (1984), pp. 351-72.
- [45] Frege, Gottlob (1918-19b), “Logical Investigations: Part II: Negation”, translated by P. Geach and R. H. Stoothoff, in: Frege (1984), pp. 373-89.
- [46] Frege, Gottlob (1919), “Notes for Ludwig Darmstaedter”, translated by P. Long and R. White, in: Frege (1979), pp. 253-57.
- [47] Frege, Gottlob (1923-26), “Logical Investigations: Part III: Compound Thoughts”, translated by P. Geach and R. H. Stoothoff, in: Frege (1984), pp. 390-406.
- [48] Frege, Gottlob (1979), *Posthumous Writings*, H. Hermes, F. Kambartel, and F. Kaulbach (eds.), Chicago, IL: The University of Chicago Press.
- [49] Frege, Gottlob (1980), *Philosophical and Mathematical Correspondence*, B. McGuinness (ed.), translated by H. Kaal, Chicago, IL: The University of Chicago Press.
- [50] Frege, Gottlob (1984), *Collected Papers on Mathematics, Logic, and Philosophy*, B. McGuinness (ed.), Oxford: Basil Blackwell.

- [51] Gaifman, Haim (2000), “What Gödel’s Incompleteness Result Does and Does Not Show”, *The Journal of Philosophy*, 97(8): 462-70.
- [52] Gaifman, Haim (2008), “Contextual Logic with Modalities for Time and Space”, *The Review of Symbolic Logic*, 1(4): 433-58.
- [53] Gaifman, Haim (2010), “Vagueness, Tolerance and Contextual Logic”, *Synthese*, 174: 5-46.
- [54] Garavaso, Pieranna (1991), “Frege and the Analysis of Thoughts”, *History and Philosophy of Logic*, 12: 195-210.
- [55] Geach, Peter (1972), *Logic Matters*, Oxford: Basil Blackwell.
- [56] Geach, Peter (1975), “Names and Identity”, in: S. Guttenplan (ed.), *Mind and Language*, Oxford: Clarendon Press, 1975, pp. 139-58.
- [57] Hale, Bob and Crispin Wright (2001), *The Reason’s Proper Study: Essays Towards a Neo-Fregean Philosophy of Mathematics*, Oxford: Clarendon Press.
- [58] Heck, Richard and Robert May (2011), “The Composition of Thoughts”, *Nous*, 45(1): 126-66.
- [59] Jourdain, Philip E. B. (1912), “The Development of the Theories of Mathematical Logic and the Principles of Mathematics: Gottlob Frege”, in: Frege (1980), pp. 179-206.
- [60] Kant, Immanuel (1781/1787), *Critique of Pure Reason*, translated by W. Pluhar, Indianapolis: Hackett Publishing Company, 1996.
- [61] Kaplan, David (1968), “Quantifying In”, *Synthese*, 19(1/2): 178-214.
- [62] Kaplan, David (1989), “Demonstratives”, in: J. Almog, J. Perry, and H. Wettstein (eds.), *Themes from Kaplan*, New York: Oxford University Press, 1989, pp. 481-563.
- [63] Kemmerling, Andreas (2011), “Thoughts without Parts: Frege’s Doctrine”, *Grazer Philosophische Studien*, 82: 165-88.
- [64] Kripke, Saul (1980), *Naming and Necessity*, Cambridge: Harvard University Press.

- [65] Kripke, Saul (2008), “Frege’s Theory of Sense and Reference: Some Exegetical Notes”, reprinted in: Kripke (2011), pp. 254-91.
- [66] Kripke, Saul (2011), *Philosophical Troubles: Collected Papers, Volume I*, Oxford: Oxford University Press.
- [67] Künne, Wolfgang (2010), “Sense, Reference and Hybridity”, *Dialectica*, 64: 529-51.
- [68] Levine, James (2002), “Analysis and Decomposition in Frege and Russell”, *The Philosophical Quarterly*, 52: 195-216.
- [69] Padro, Romina (2015), *What the Tortoise Said to Kripke: the Adoption Problem and the Epistemology of Logic* (Doctoral dissertation). Retrieved from CUNY Academic Works.
- [70] Peacocke, Christopher (1996), “Entitlement, Self-Knowledge, and Conceptual Redeployment”, *Proceedings of the Aristotelian Society*, 96: 117-58.
- [71] Peacocke, Christopher (2008), *Truly Understood*, Oxford: Oxford University Press.
- [72] Penco, Carlo (2003), “Frege: Two Theses, Two Senses”, *History and Philosophy of Logic*, 24: 87-109.
- [73] Quine, W.V. (1936), “Truth by Convention”, in: *Philosophical Essays for A N Whitehead*, New York: Longmans, Green, 1936, pp. 90-124.
- [74] Quine, W.V. (1956), “Quantifiers and Propositional Attitudes”, *The Journal of Philosophy*, 53(5): 177-87.
- [75] Quine, W.V. (1960), *Word and Object*, Cambridge: The MIT Press.
- [76] Russell, Bertrand (1910-11), “Knowledge by Acquaintance and Knowledge by Description”, reprinted in: Russell (1918), pp. 165-83.
- [77] Russell, Bertrand (1914), “The Relation of Sense-Data to Physics”, reprinted in: Russell (1918), pp. 113-41.
- [78] Russell, Bertrand (1918), *Mysticism and Logic*, London: Longmans, Green.

- [79] Textor, Mark (2009), “A Repair of Frege’s Theory of Thoughts”, *Synthese*, 167: 105-23.
- [80] Wittgenstein, Ludwig (1922), *Tractatus Logico-Philosophicus*, translated by C.K. Ogden, New York: Harcourt, Brace.